Assignment 1 - IEEE Visualization 2008 Design Contest

By: **Shathin R Rao [MT2020067]**

Repo: [Shathin/cs732-data-visualization]

Abstract

This report describes the tools and methodology used for visualizing different scalar fields from the data set provided as part of the **IEEE Visualization 2008 Design Contest**. The data was obtained from a numerical simulation of the first stars in the galaxy. The report then attempts to make certain inferences from the visualizations generated.

Introduction

Scientists have performed three-dimensional radiation hydrodynamical calculations of ionization front instabilities in which multifrequency radiative transfer is coupled to eight species primordial chemistry. They have simulated ten scalar fields - **Particle Density**, **Gas Temperature**, and **8 chemical species** (H, $H+$, $H-$, He, $He+$, $He++$, H_2, H_2+). The technique attempted for the visualization involve contour mapping (contour fill). The methodology section elaborates on the chosen technique for each field. In the last section, an attempt is made to make certain infererences from the generated visualization

Tools

This sections describes the libraries chosen for the visualization and the briefs about the scripts written using the libraries to generate the visualizations. **Python** was used as the programming language for generating the visualization.Python was chosen primarily because Python is a very mature language when it comes to the domain Data Analysis and hence can deal with large datasets.

Libraries

The set of libraries used for generating the visualizations involved -

- `numpy` → (Library Home) **NumPy** is the fundamental package for scientific computing in Python. It is a Python library that provides a multidimensional array object, various derived objects (such as masked arrays and matrices), and an assortment of routines for fast operations on arrays, including mathematical, logical, shape manipulation, sorting, selecting, I/O, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation and much more.
- `pandas` → (Library Home) **pandas** is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language. It provides methods to read data from files and transform them as needed.
- `matplotlib` → (Library Home) **Matplotlib** is a comprehensive library for creating static, animated, and interactive visualizations in Python.

Scripts

All the scripts mentioned below have been written in a way to enable fast iteration of the data i.e., with very minimal changes to the code, new visualizations can be generated.

- `dataOps.py` → Contains methods that reads and processes the data
- `formulae.py` → Contains the different scalar formulae that can be applied on the fields in the dataset
- `scalarFields.py` → A configuration file that defines each of the given scalar fields and the derived fields. Each field is a dictionary containing different attributes that define the properties of the field.
- `scalarVisualization.py` → The python script that reads and visualizes the desired scalar field from the data set. The file contains scripts for both static and animated visualizations.

Data

This section briefly describes the given scalar fields and the derived fields that was made use of for the visualization.

Scalar Fields

- **Gas Temperature (**$Kelvin$**)** → A *given* scalar field, defines the temperature of the gas in the region.
- **Total Particle Density (** $\#\ of\ particles\ /\ cm^3$**)** → A *given* scalar field, defines the number of particles (of all the chemical species) in a given region.
- **Hydrogen Number Density (** $\#\ of\ particles\ /\ cm^3$**)** → A *derived* scalar field, defines the number of Hydrogren particles in a given region.
- **Hydrogen+ Number Density (** $\#\ of\ particles\ /\ cm^3$**)** → A *derived* scalar field, defines the number of (positively) ionized Hydrogen particles in a given region.
- **Hydrogen- Number Density (** $\#\ of\ particles\ /\ cm^3$**)** → A *dervied* scalar field, defines the number of (negatively) ionized Hydrogen particles in a given region.

Computing Derived Fields

The formule used for computing the derived fields are present in the `formulae.py` file.

Total Density

Before computing the selected derived fields, the total density is to be computed. The **total density** (g/cm^3) is a function of the mass abundance of all the chemical species (Xi_x) and the total particle density tpd. Following the formula used for computing it -

$$\rho = \frac{tpd * mh}{Xi_H + Xi_{H+} + 0.25 * (Xi_{He} + Xi_{He+} + Xi_{He++}) + Xi_{H-} + 0.5 * (Xi_{H_2} + Xi_{H_2+})}$$

Where mh is the Mass of Hydrogen i.e., $mh = 1.38066 * 10^{-24}g$

H, H+ and H- Number Density

Hydrogen, Hydrogen+ and Hydrogen- Number densities (nd_x) are calculated using the following common formula -

$$nd_x = Xi_{Hx} * \frac{\rho}{mh}$$

Where Xi_{Hx} is Xi_H, Xi_{H+} or Xi_{H-} is the mass abudance of H, H+ and H- chemical species, ρ is the total density and mh is the Mass of Hydrogen

---

Visualization and Inferences

This section describes the methodology used for the visualizations and the inferences made from the visualizations.

Choosing Timesteps

Out of the entire dataset available only the following timesteps were chosen - 0, 1, 2, 3, 6, 9, 14, 19, 29, 49, 69, 99, 129, 159, 189.

The reason for choosing such a timestep that isn't uniformly distributed is because upon trying to visualize the entire dataset i.e., all the timesteps it was observed that most of the action happens quickly during the initial few timesteps and then propagates slowly i.e., not much happens between contiguous timesteps.

Choosing data slice

The Z axis was kept constant at $Z = 124$ as it will provide a value in the center of the data set. Also, Z being the slowest changing axis it becomes convenient to keep it fixed.

Visualizations

Gas Temperature

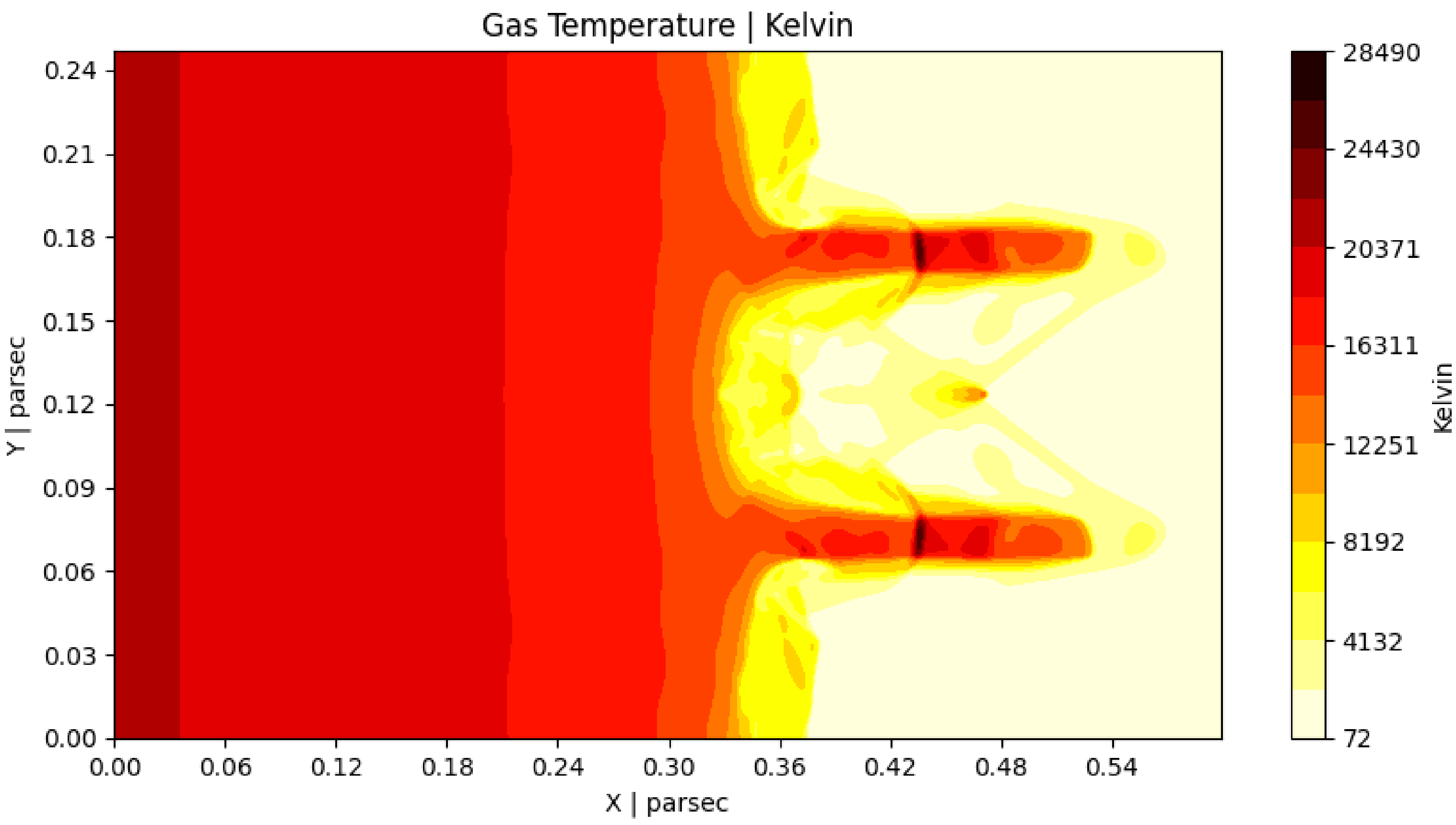

Figure 1: Gas Temperature at timestamp 99 [12.5 kYear]

▼ **Technique** - *Contour Fill*

Contour Fill (Mapping) was chosen to group the regions of (approximately close) temperature together.

▼ **Color Map** - *Hot [Reverse]*

When one thinks of temperature, one generally thinks of the color red and hence the **Hot** sequential color map was chosen.

▼ **Normalization** - *Min-Max Normalization*

Since the unit of the temperature data provided is in **Kelvin**, no negative values are required to represent temperature. Hence a normal min-max normalization proves sufficient.

▼ **Colorbar Scale** - *Linear*

Since the values aren't extremely large, a linear colorbar was chosen.

▼ **Values Utilized** - *Gas Temperature* $[Kelvin]$

The values utilized from the data set was only the Gas Temperature field. The X and Y co-ordinates were computed based on the information given in the contest webpage.

▼ **Basic Inferences**

- By looking at the above frame, it is clear that the ionization front very rapidly heats up the temperature of the gases in a region. This is evident by looking at the width of the regions closer to the front of the ionization front. The width of regions towards the left are much wider than the the regions towards the right i.e., closer to the ionization front.
- Another thing that can be observed is that the temperature continues to rise even after the ionization front passes a region and is much higher than the temperature at the ionization front.

Total Particle Density

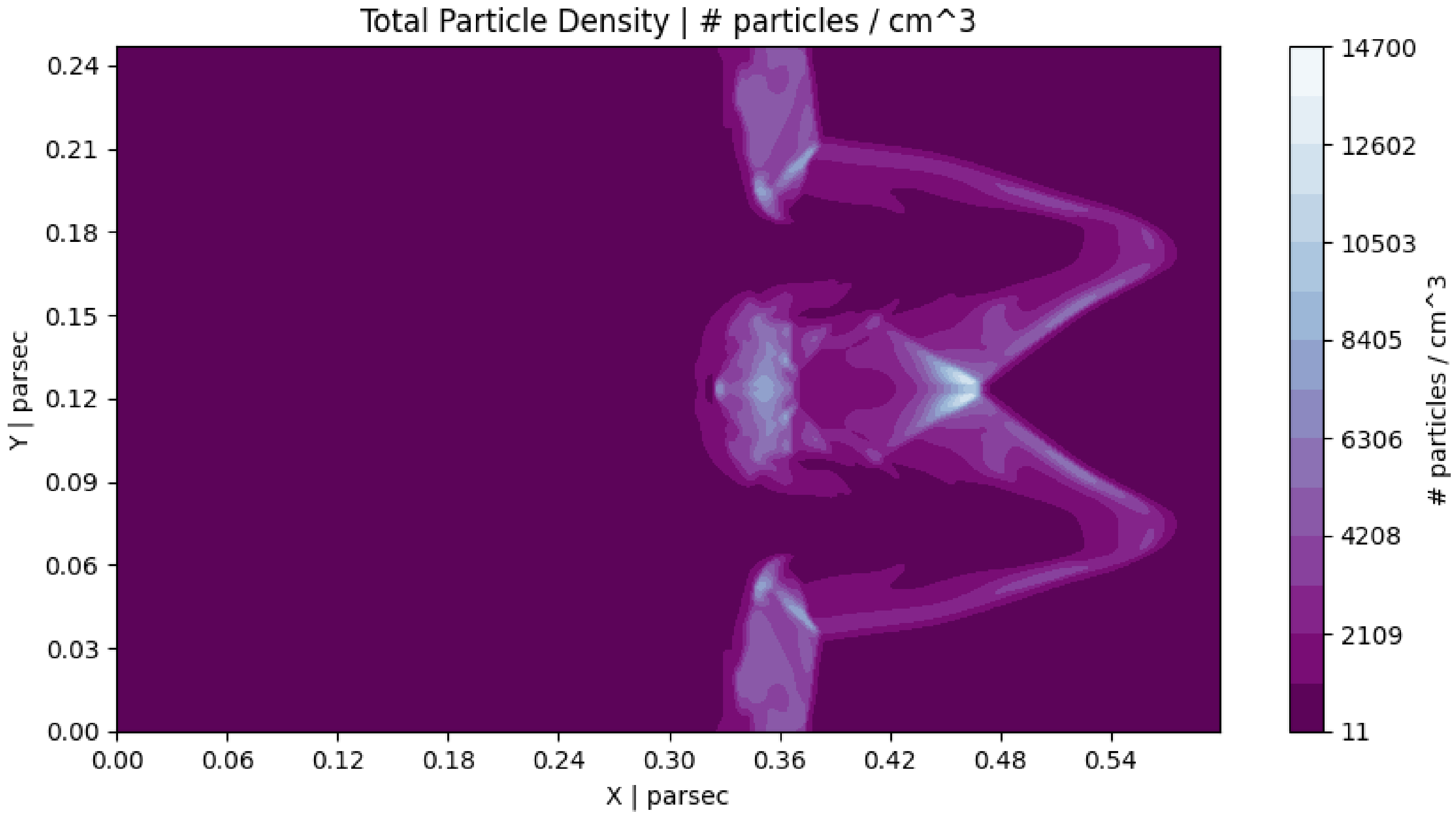

Figure 2: Total Particle Density at timestamp 99 [12.5 kYear]

▼ **Technique** - *Contour Fill*

Contour Fill (Mapping) was chosen to group the regions of (approximately close) particle (number) density together.

▼ **Color Map** - Blue-Purple (BuPu) *[Reverse]*

This color map was chosen since it was perceptually uniform color map. The reverse of the color map was chosen since region is very narrow and using the non-reversed color map makes it hard to see the regions.

▼ **Normalization** - *Min-Max Normalization*

Since the unit of the data provided is in **number of particles per** cm^3, no negative values are required to represent total particle density. Hence a normal min-max normalization proves sufficient.

▼ **Colorbar Scale** - *Linear*

Since the values aren't extremely large, a linear colorbar was chosen.

▼ **Values Utilized** - *Total Particle Density* [$\# \ of \ particles \ / \ cm^3$]

The values utilized from the data set was only the Total Particle Density field. The X and Y co-ordinates were computed based on the information given in the contest webpage.

▼ **Basic Inference**

- By looking at the above frame it is evident that more particles are concentrated in front of the ionization front leading to the observation that the ionization front has a compressing effect of the gas particles in the region.

H Number Density

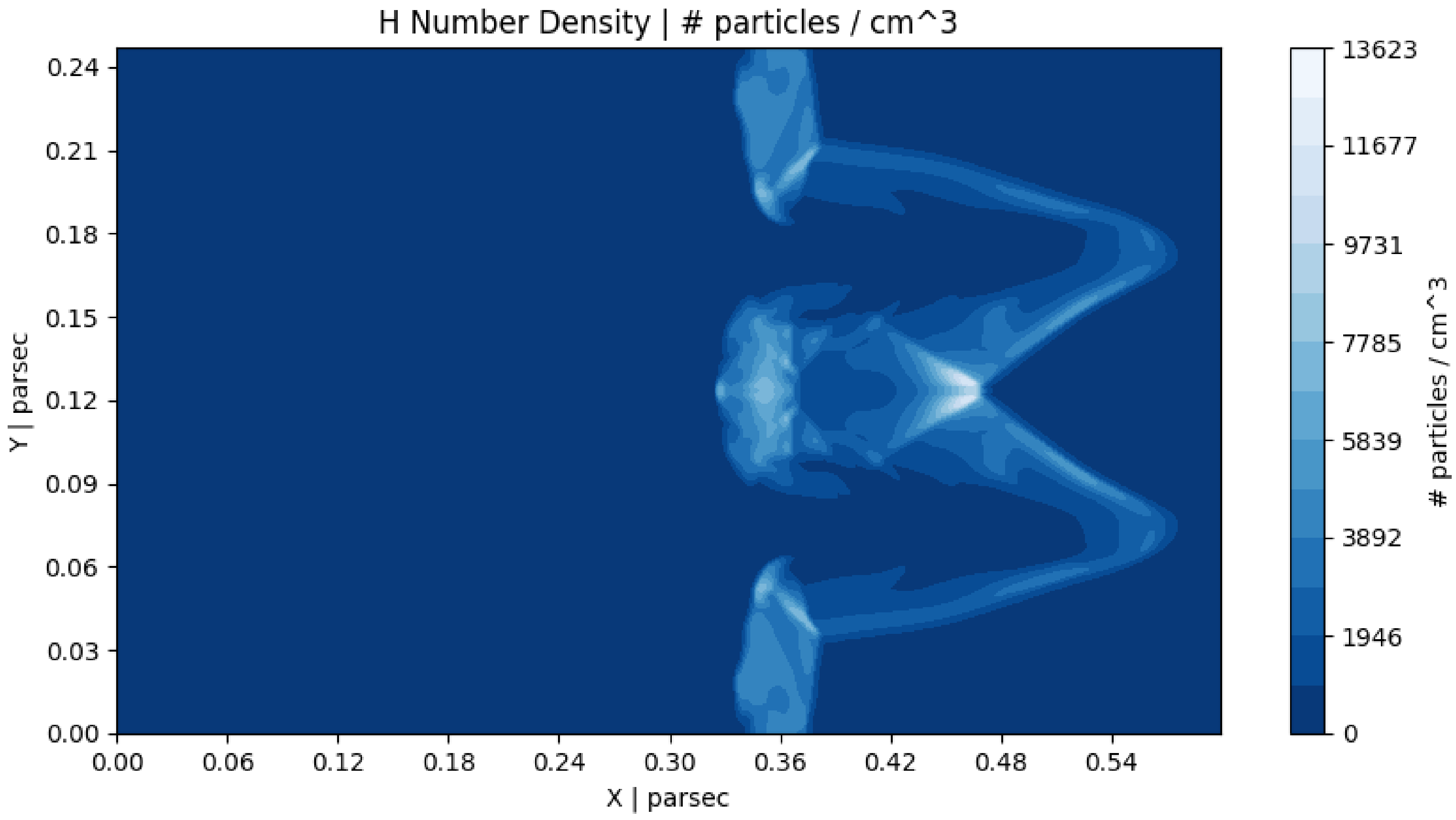

Figure 3: Hydrogen Number Density at timestamp 99 [12.5 kYear]

▼ **Technique** - *Contour Fill*

Contour Fill (Mapping) was chosen to group the regions of (approximately close) Hydrogen particle (number) density together.

▼ **Color Map** - *Blues [Reverse]*

This color map was chosen since it was perceptually uniform color map. The reverse of the color map was chosen since region is very narrow and using the non-reversed color map makes it hard to see the regions.

▼ **Normalization** - *Min-Max Normalization*

Since the unit of the data provided is in **number of particles per** cm^3, no negative values are required to represent Hydrogen Number density. Hence a normal min-max normalization proves sufficient.

▼ **Colorbar Scale** - *Linear*

Since the values aren't extremely large, a linear colorbar was chosen.

▼ **Values Utilized** - *Total Particle Density* [$\# \ of \ particles \ / \ cm^3$]

This is a derived field and hence to compute this **total particle density** field and all the **chemical species** fields had to be used for the computation. The X and Y co-ordinates were computed based on the information given in the contest webpage.

▼ **Basic Inference**

- By looking at the above frame it is evident that more particles are concentrated in front of the ionization front leading to the observation that the ionization front has a compressing effect of the gas particles in the region.

H+ Number Density

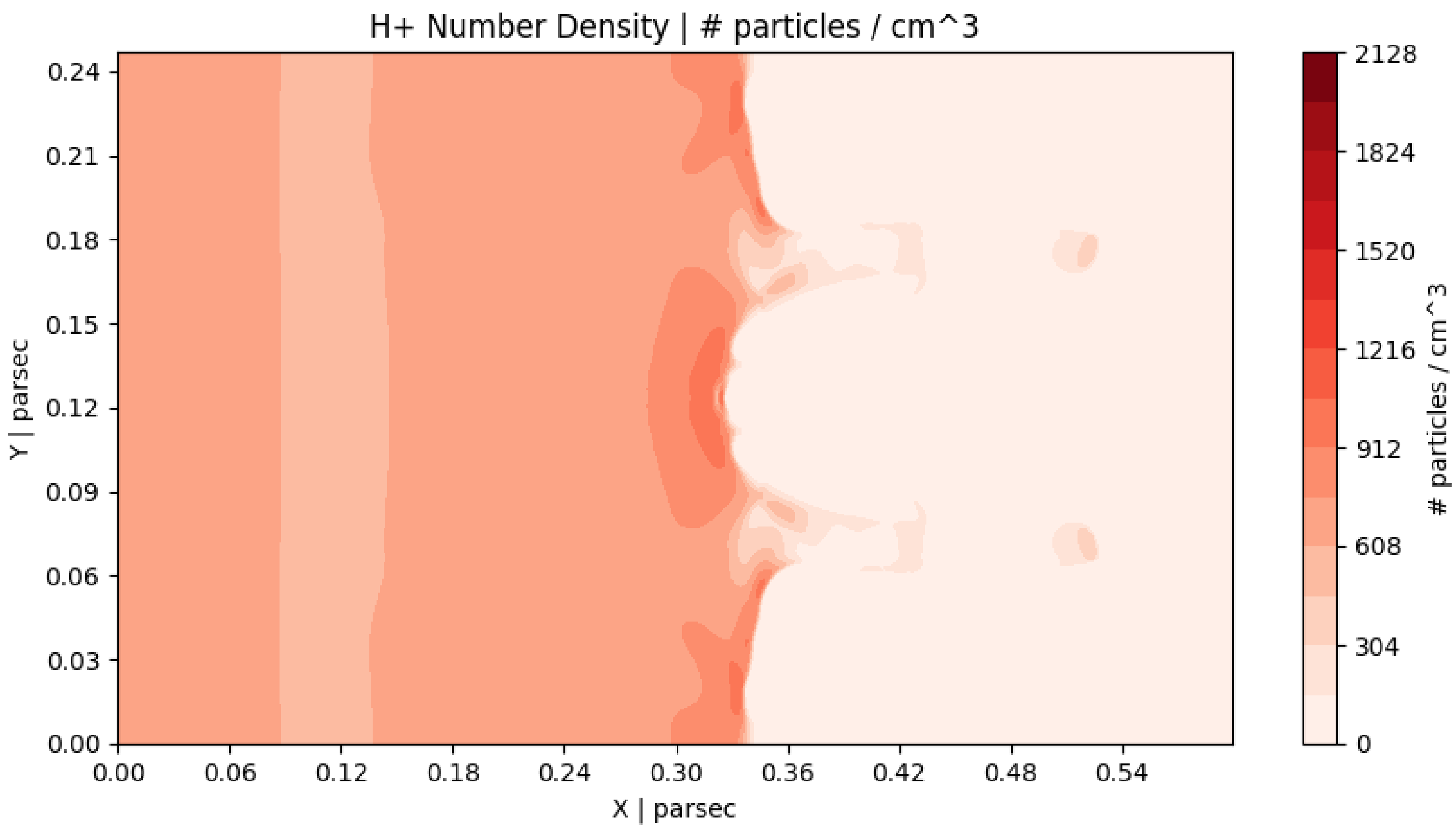

Figure 3: Hydrogen+ Number Density at timestamp 99 [12.5 kYear]

▼ **Technique** - *Contour Fill*

Contour Fill (Mapping) was chosen to group the regions of (approximately close) positively ionized Hydrogen particle (number) density together.

▼ **Color Map** - *Reds*

This color map was chosen since it was perceptually uniform color map. Red was chosen because ionized gas is very hot.

▼ **Normalization** - *Min-Max Normalization*

Since the unit of the data provided is in **number of particles per** cm^3, no negative values are required to represent Hydrogen+ Number density. Hence a normal min-max normalization proves sufficient.

▼ **Colorbar Scale** - *Linear*

Since the values aren't extremely large, a linear colorbar was chosen.

▼ **Values Utilized** - *Total Particle Density* $[\# \ of \ particles \ / \ cm^3]$

This is a derived field and hence to compute this **total particle density** field and all the **chemical species** fields had to be used for the computation. The X and Y co-ordinates were computed based on the information given in the contest webpage.

▼ **Basic Inference**

- By looking at the above frame it is evident that more particles are concentrated in behind of the ionization front leading to the observation that the ionization front leaves behind a trail of hot ionized gas.

H- Number Density

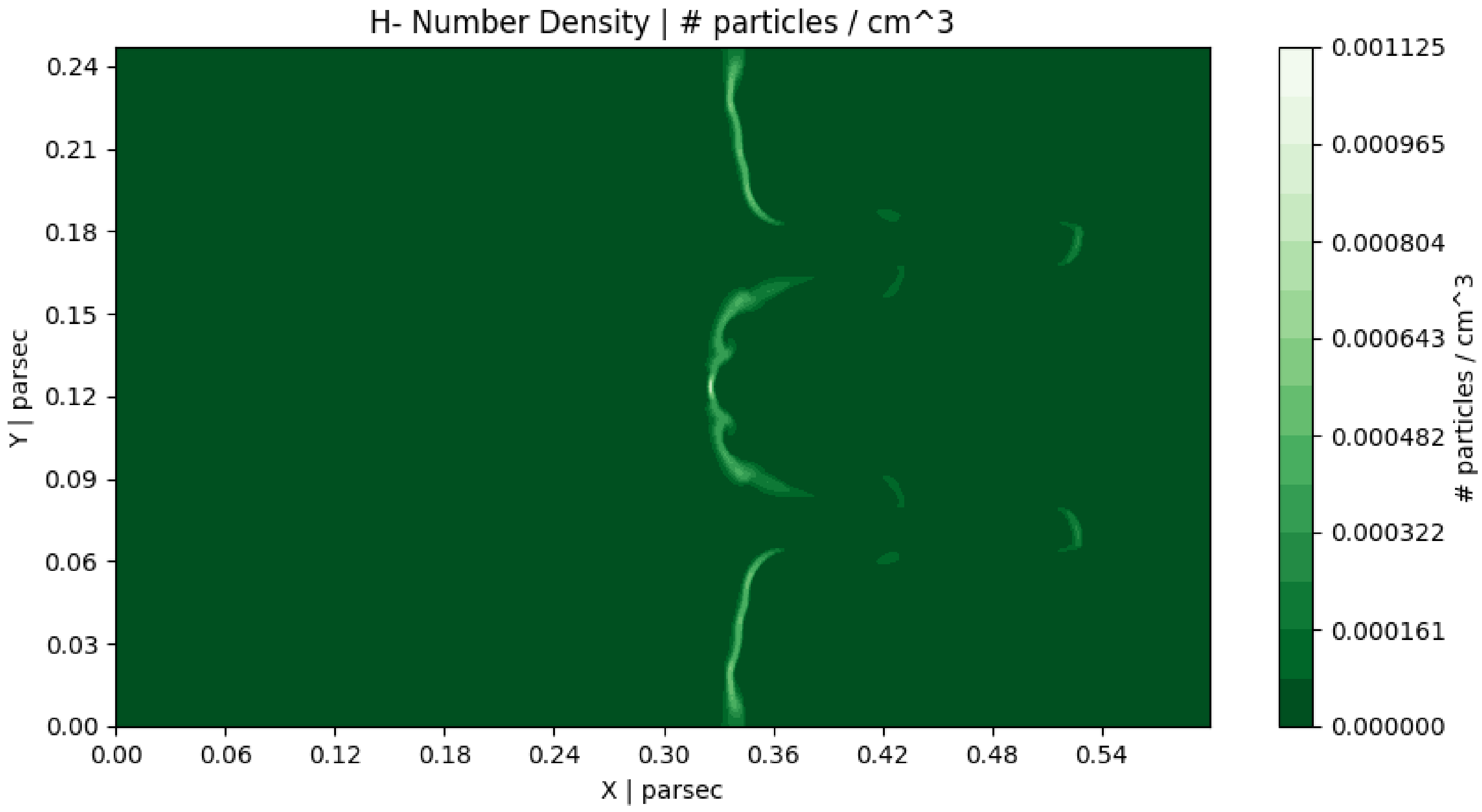

Figure 4: Hydrogen- Number Density at timestamp 99 [12.5 kYear]

- **Technique** - *Contour Fill*

  Contour Fill (Mapping) was chosen to group the regions of (approximately close) negativel ionized Hydrogen particle (number) density together.

- **Color Map** - Greens *[Reverse]*

  This color map was chosen since it was perceptually uniform color map. The reverse of the color map was chosen since region is very narrow and using the non-reversed color map makes it hard to see the regions.

- **Normalization** - *Min-Max Normalization*

  Since the unit of the data provided is in **number of particles per** cm^3, no negative values are required to represent Hydrogen- Number density. Hence a normal min-max normalization proves sufficient.

- **Colorbar Scale** - *Linear*

  Since the values aren't extremely large, a linear colorbar was chosen.

- **Values Utilized** - *Total Particle Density* [$\# \ of \ particles \ / \ cm^3$]

  This is a derived field and hence to compute this **total particle density** field and all the **chemical species** fields had to be used for the computation. The X and Y co-ordinates were computed based on the information given in the contest webpage.

- **Basic Inference**

  - By looking at the above frame it is evident that particles are concentrated closer to the ionization front leading to the observation it is extremely unstable and doesn't remain for long.

Combined Inferences

Looking at the 5 chosen field together, we can observe how one affects the other. Following are three frames chosen out of the timesteps used for the visualization.

Time Frame 1 - 0.7 kYear

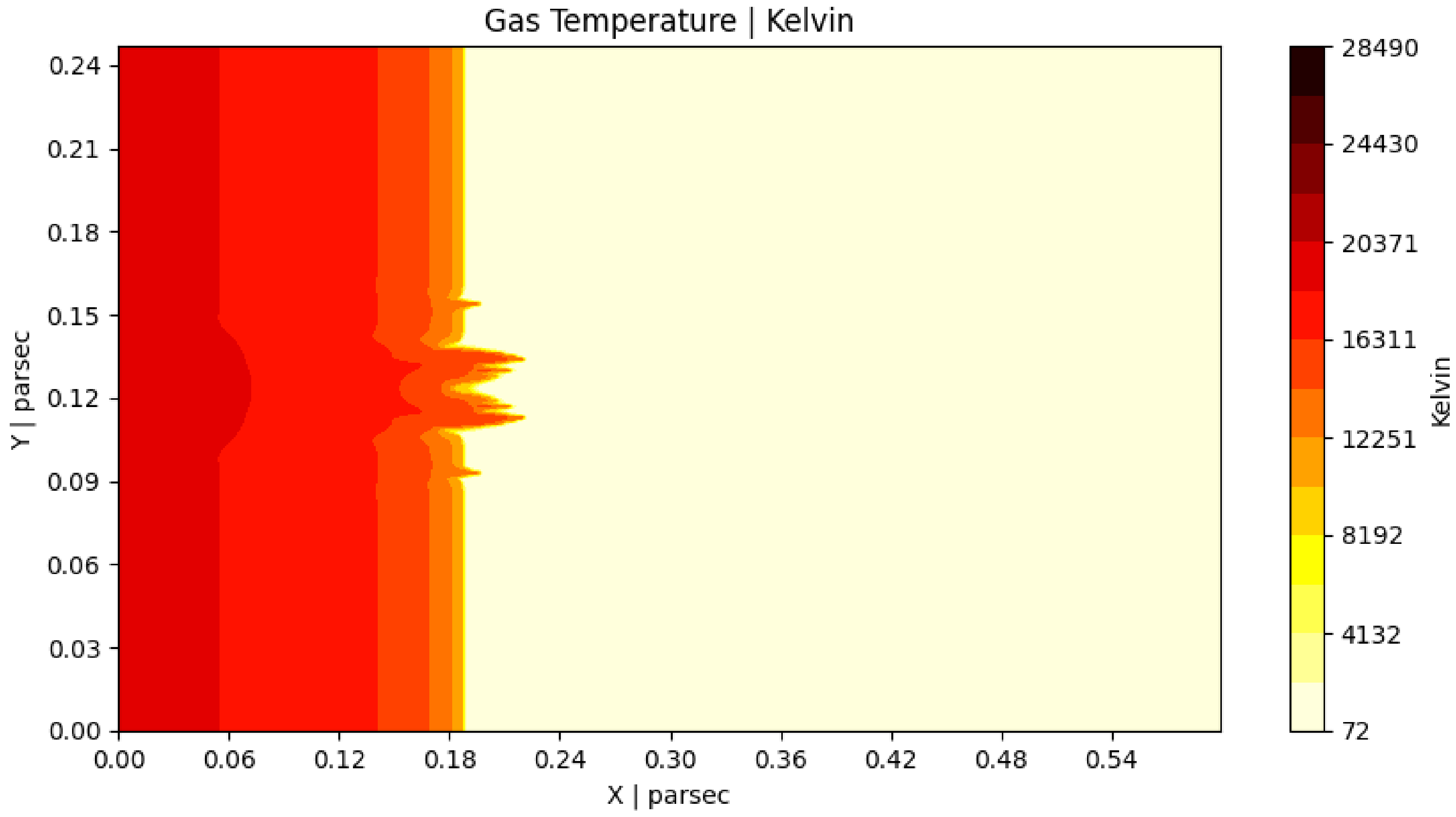

Figure 5.1: Gas Temperature at timestamp 06 [0.7 kYear]

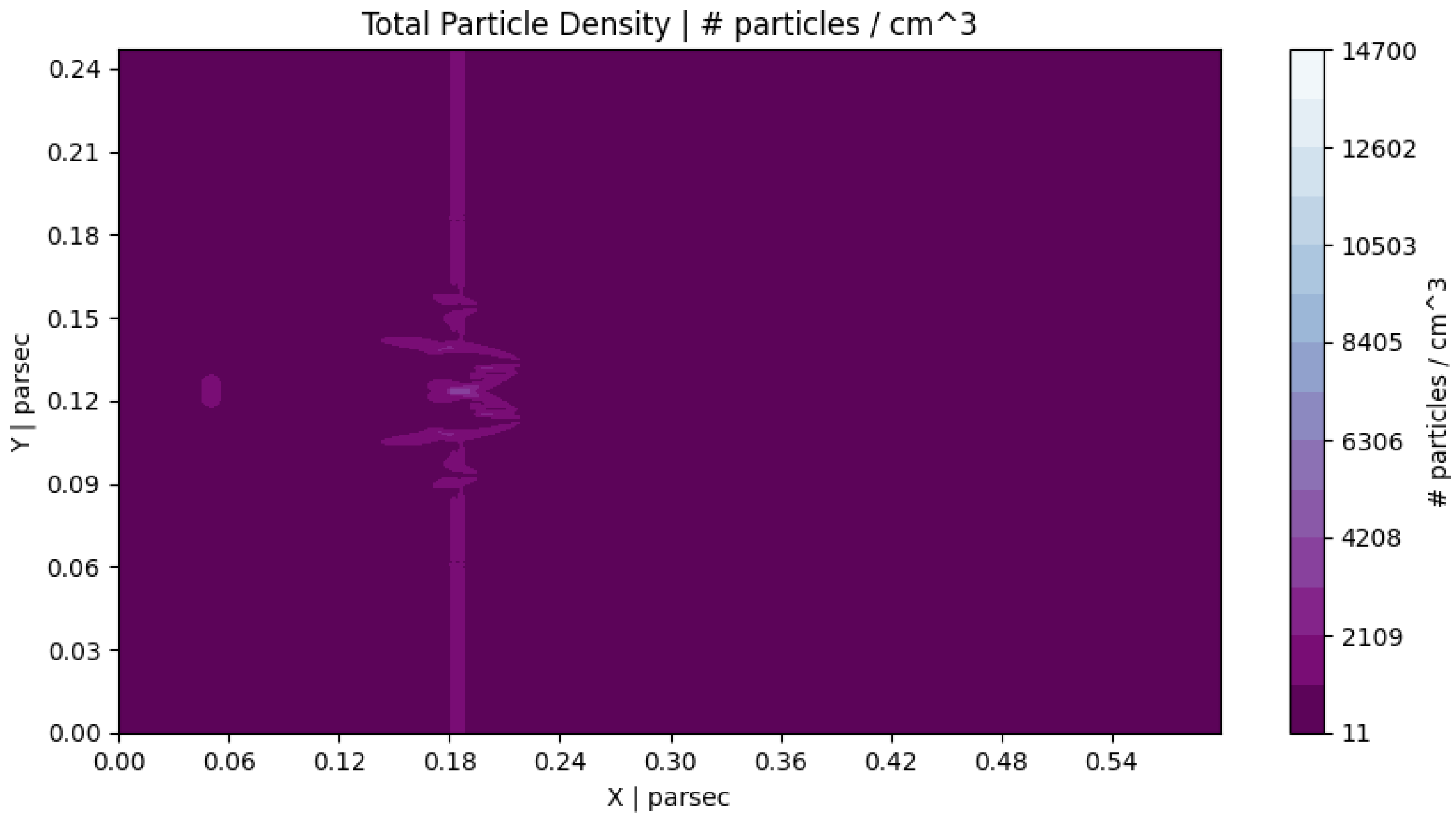

Figure 5.2: Total Particle Density at timestamp 06 [0.7 kYear]

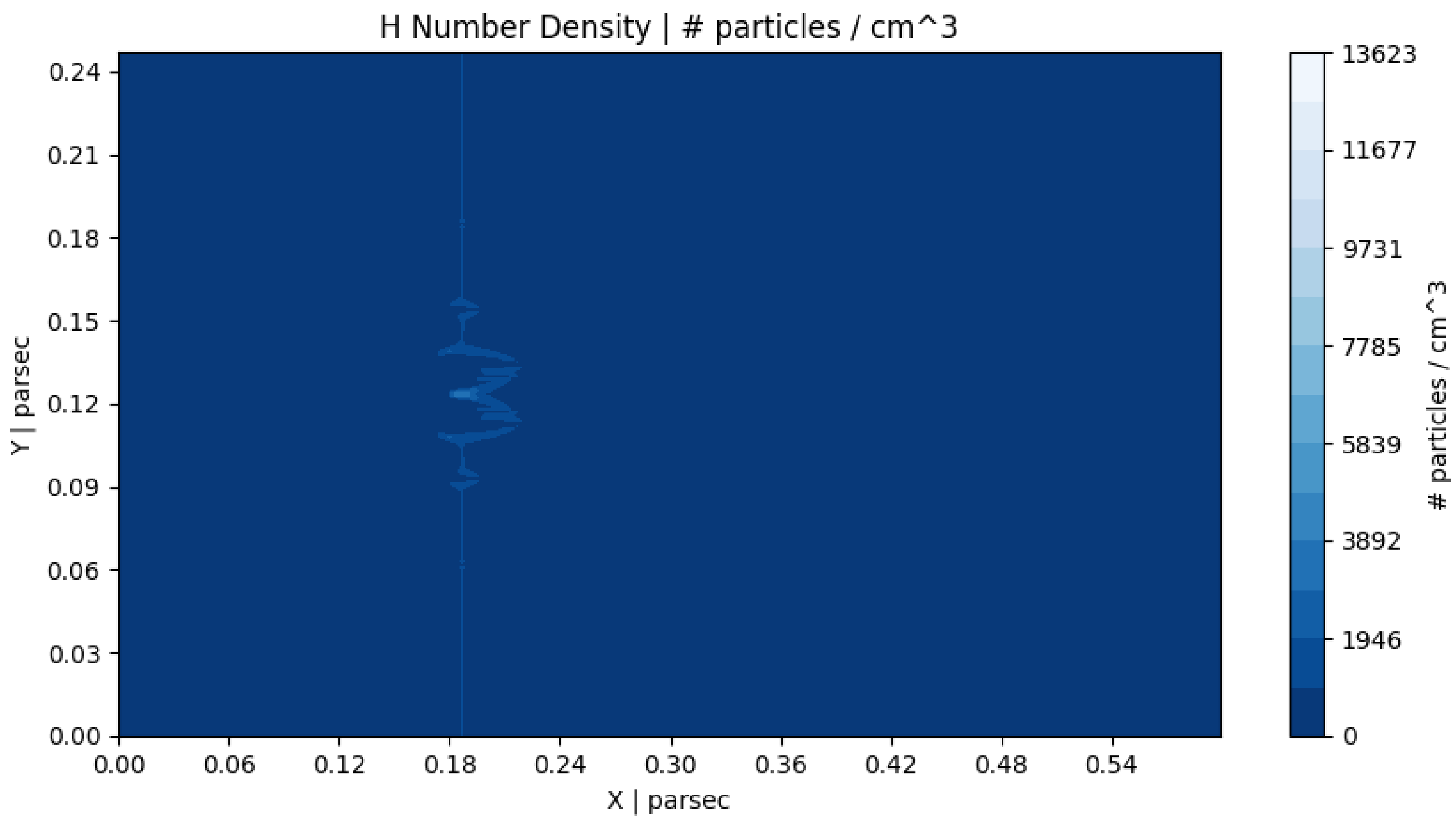

Figure 5.3: Hydrogen Number at timestamp 06 [0.7 kYear]

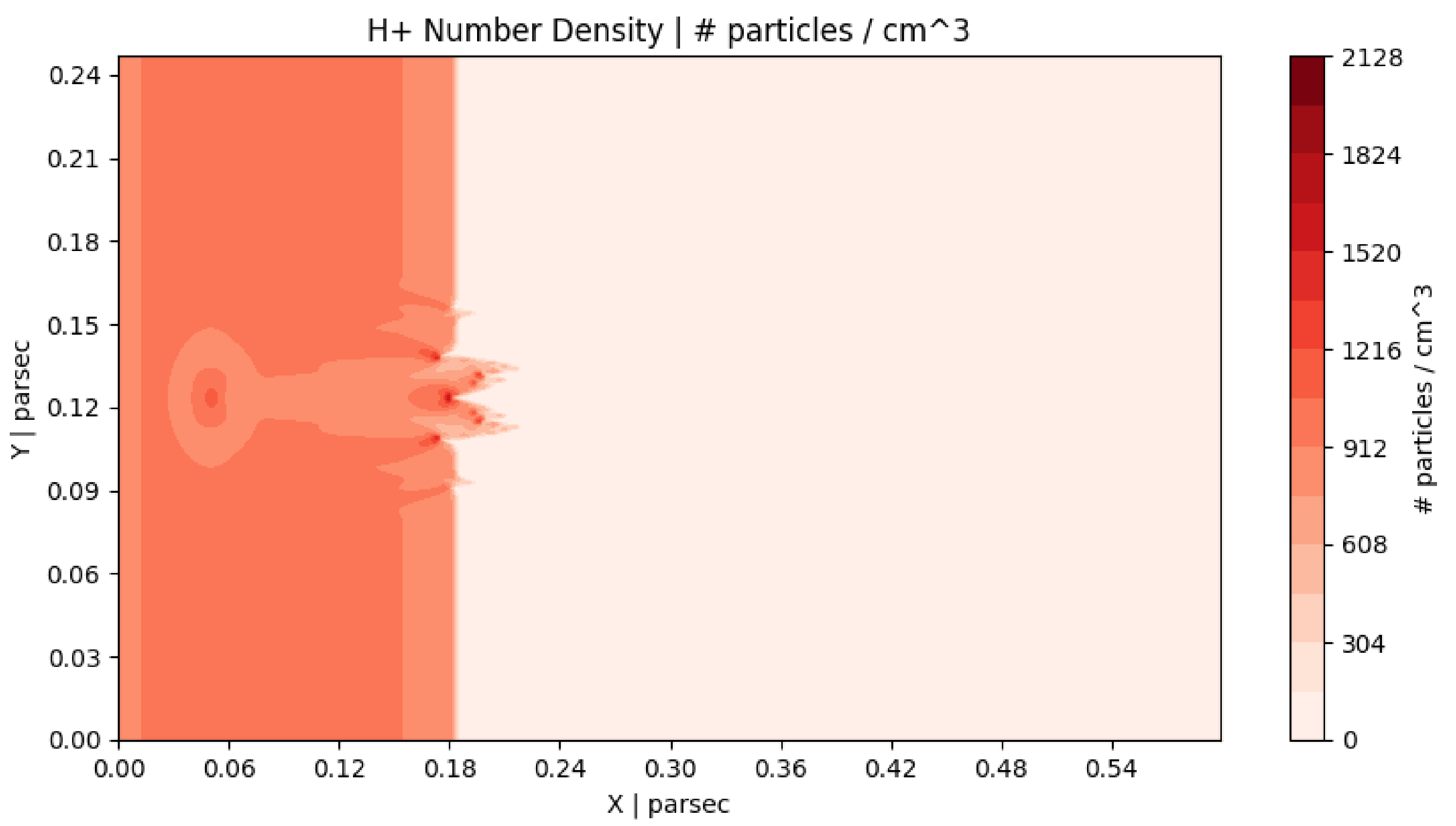

Figure 5.4: Hydrogen+ Number at timestamp 06 [0.7 kYear]

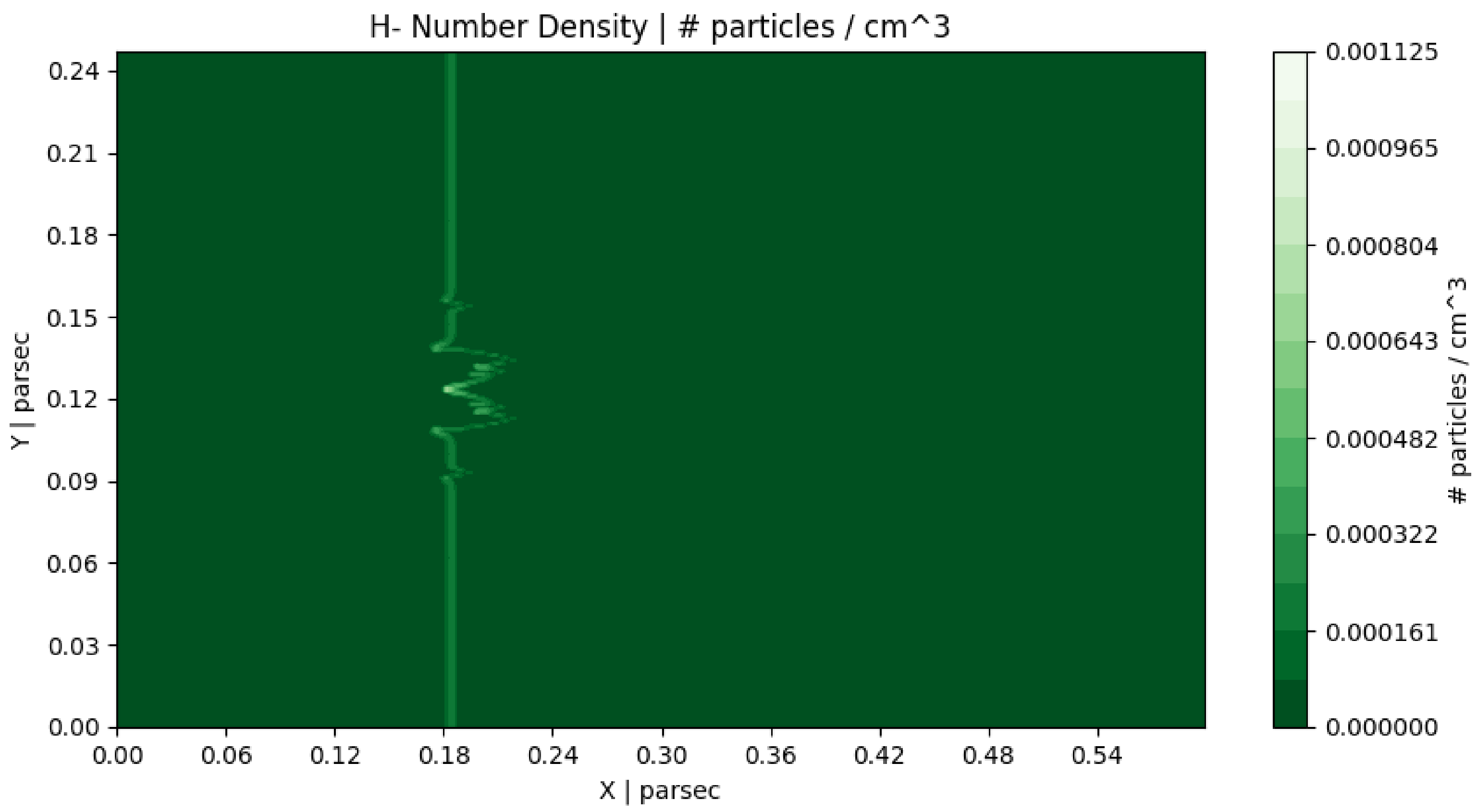

Figure 5.5: Hydrogen- Number at timestamp 06 [0.7 kYear]

Time Frame 2 - 4.6 kYear

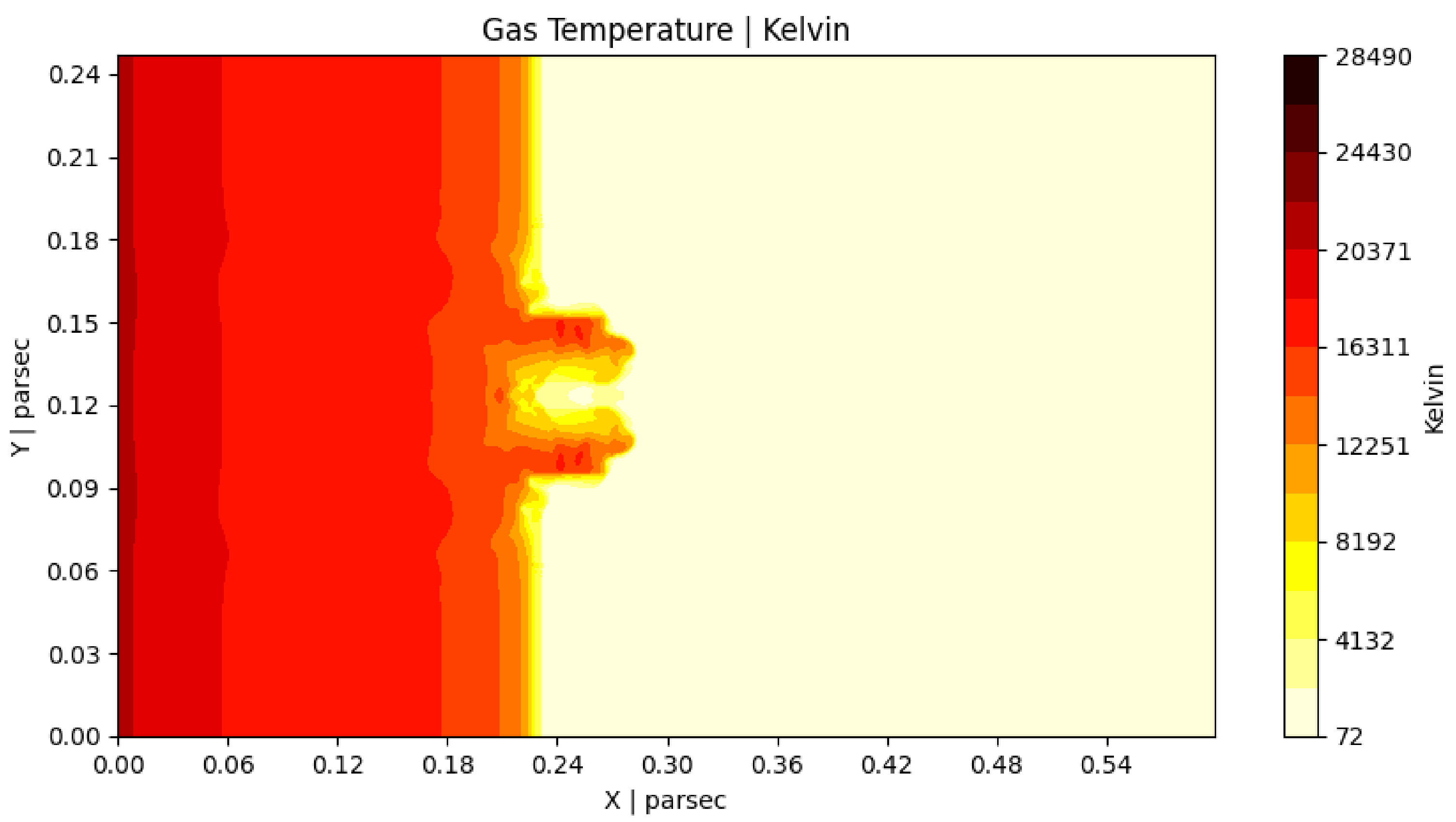

Figure 6.1: Gas Temperature at timestamp 29 [4.6 kYear]

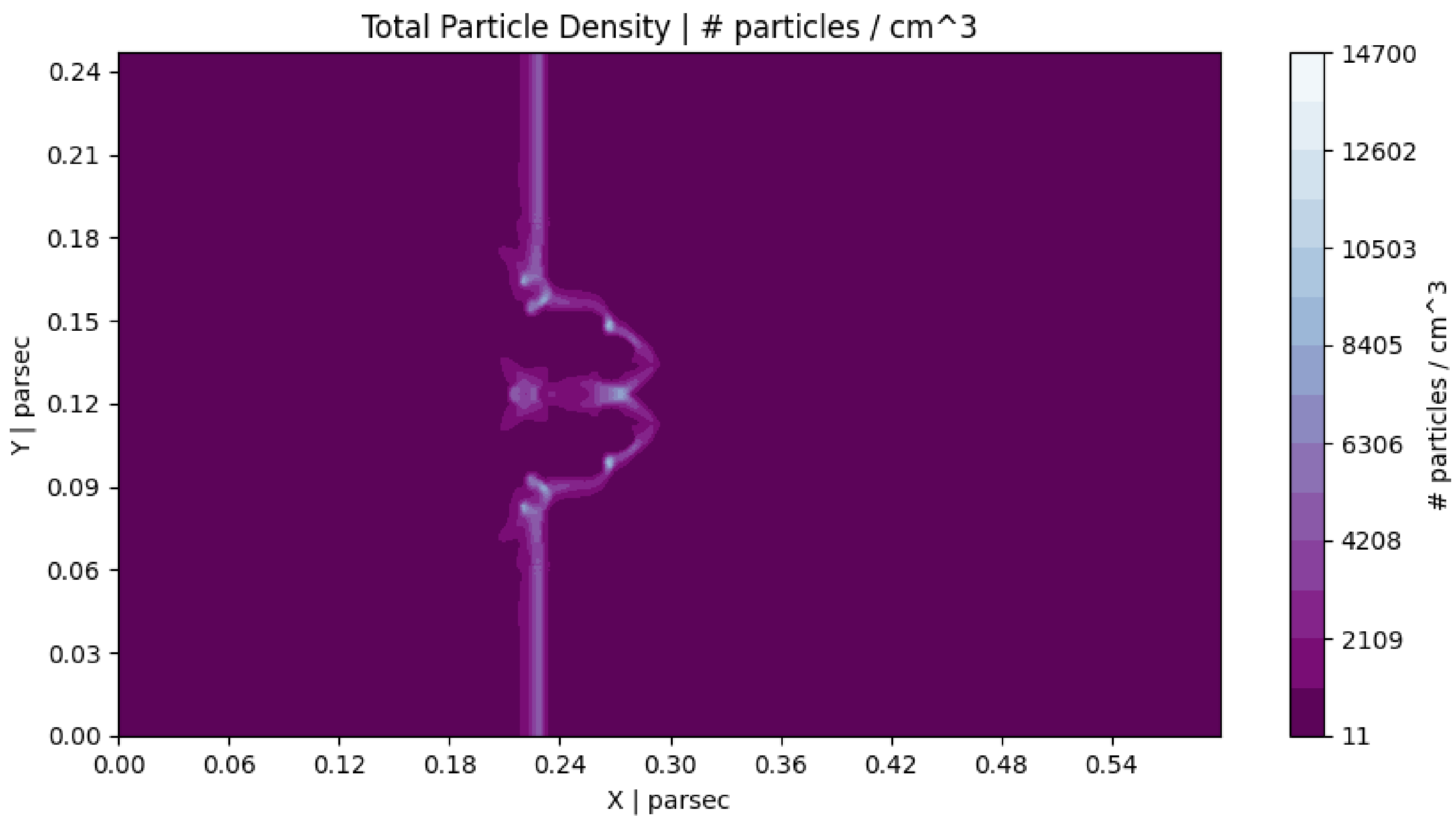

Figure 6.2: Total Particle Density at timestamp 29 [4.6 kYear]

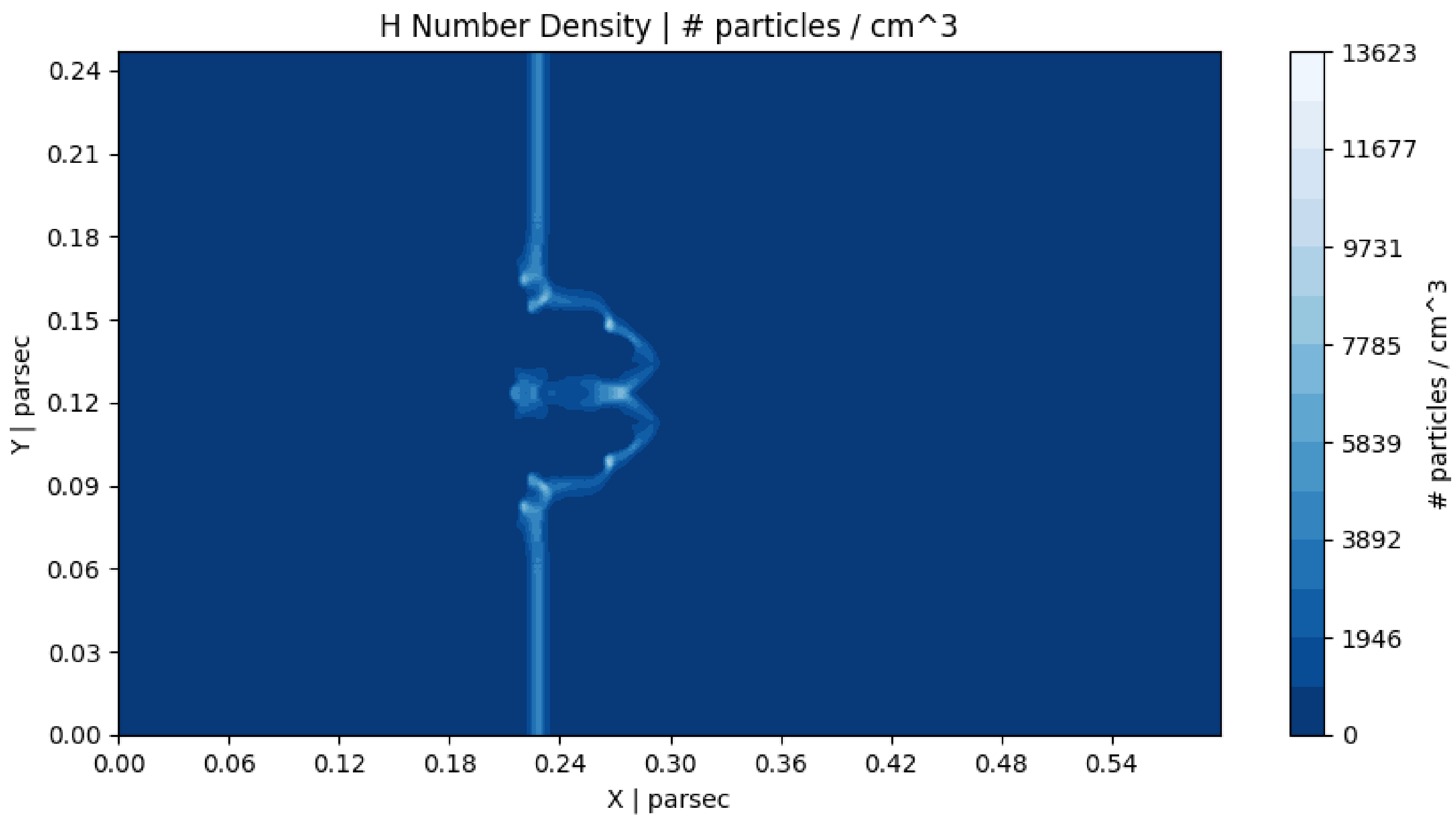

Figure 6.3: Hydrogen Number Density at timestamp 29 [4.6 kYear]

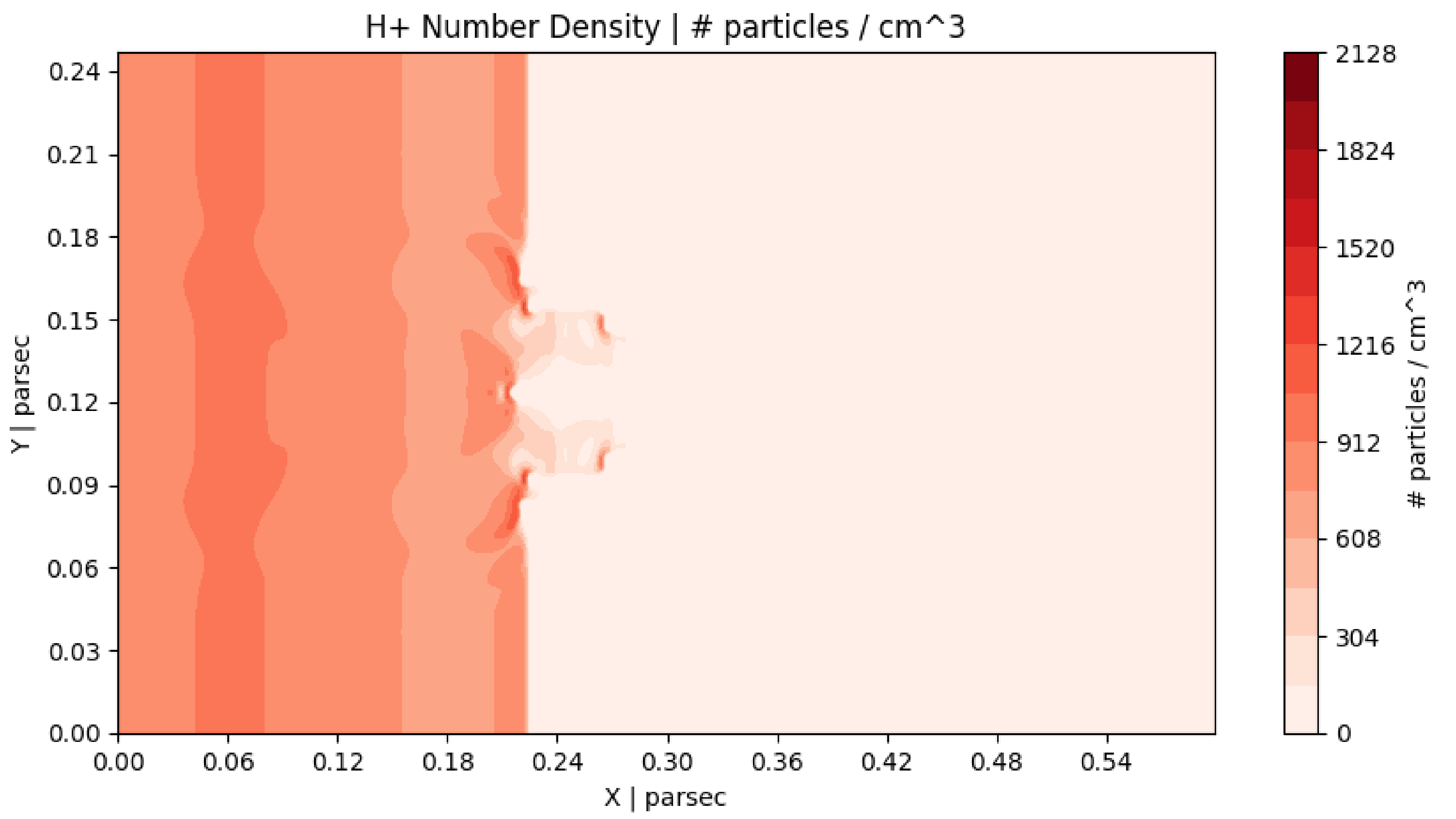

Figure 6.4: Hydrogen+ Number Density at timestamp 29 [4.6 kYear]

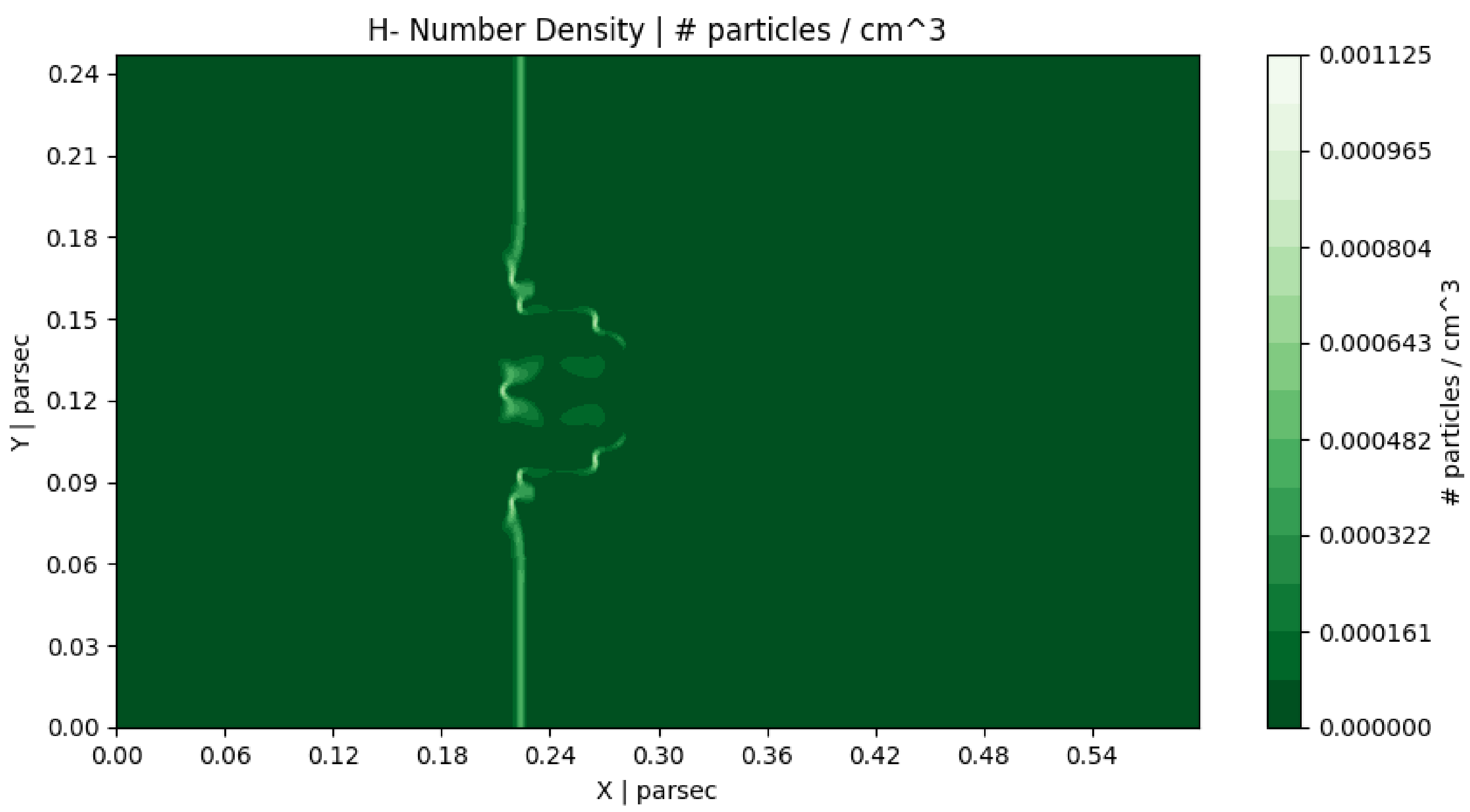

Figure 6.5: Hydrogen- Number Density at timestamp 29 [4.6 kYear]

Time Frame 3 - 12.5 kYear

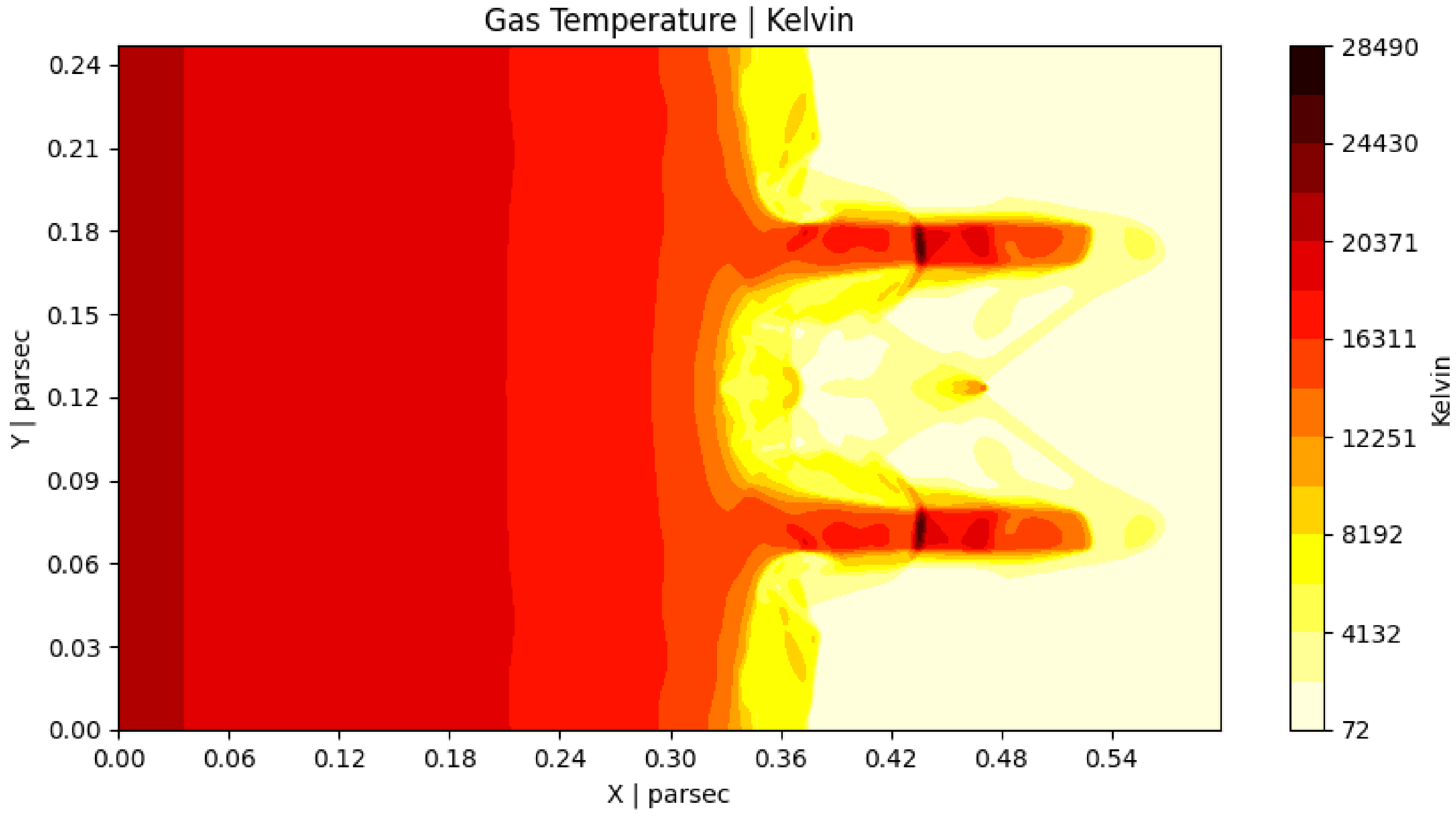

Figure 7.1: Gas Temperature at timestamp 99 [12.5 kYear]

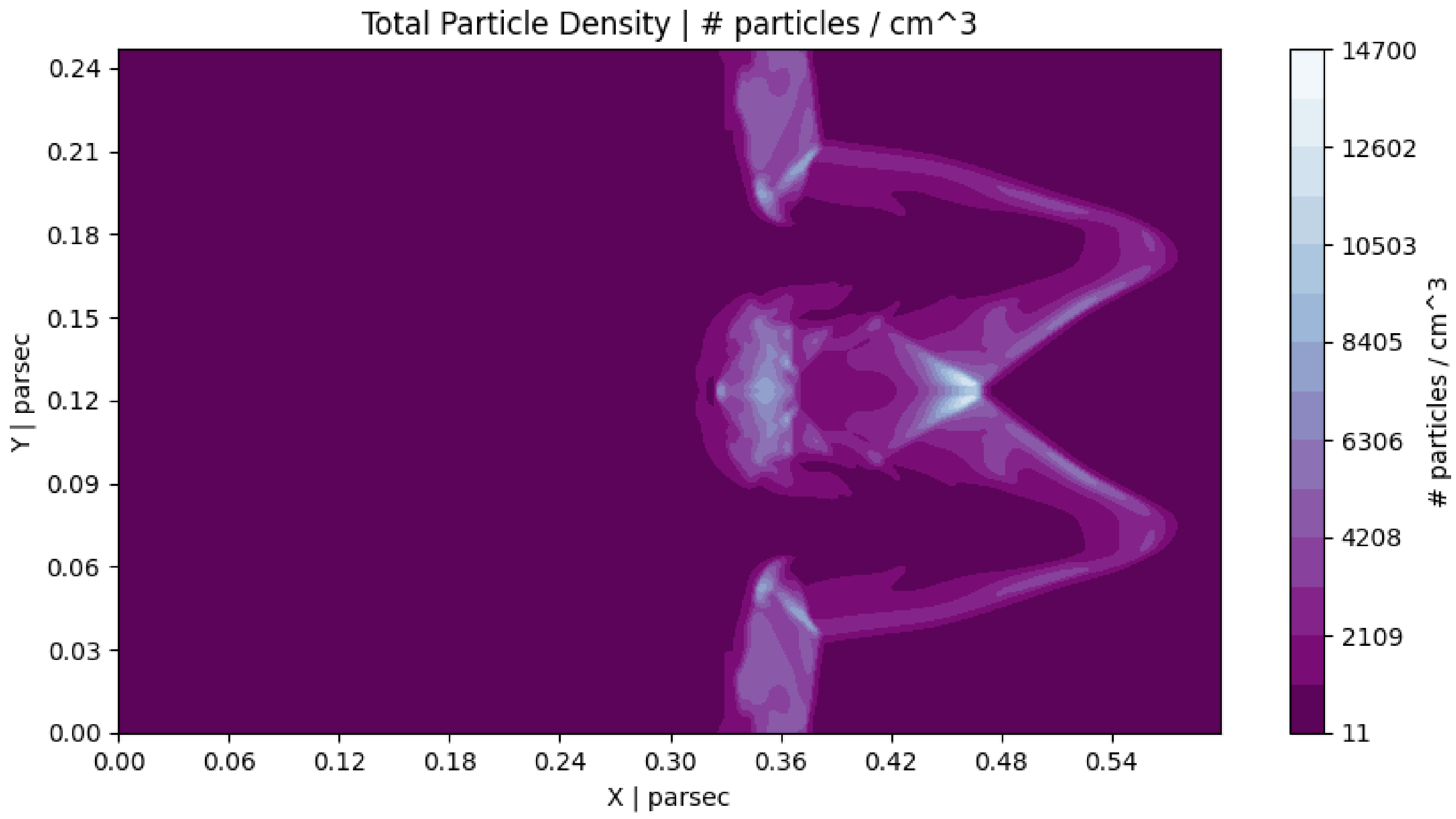

Figure 7.2: Total Particle Density at timestamp 99 [12.5 kYear]

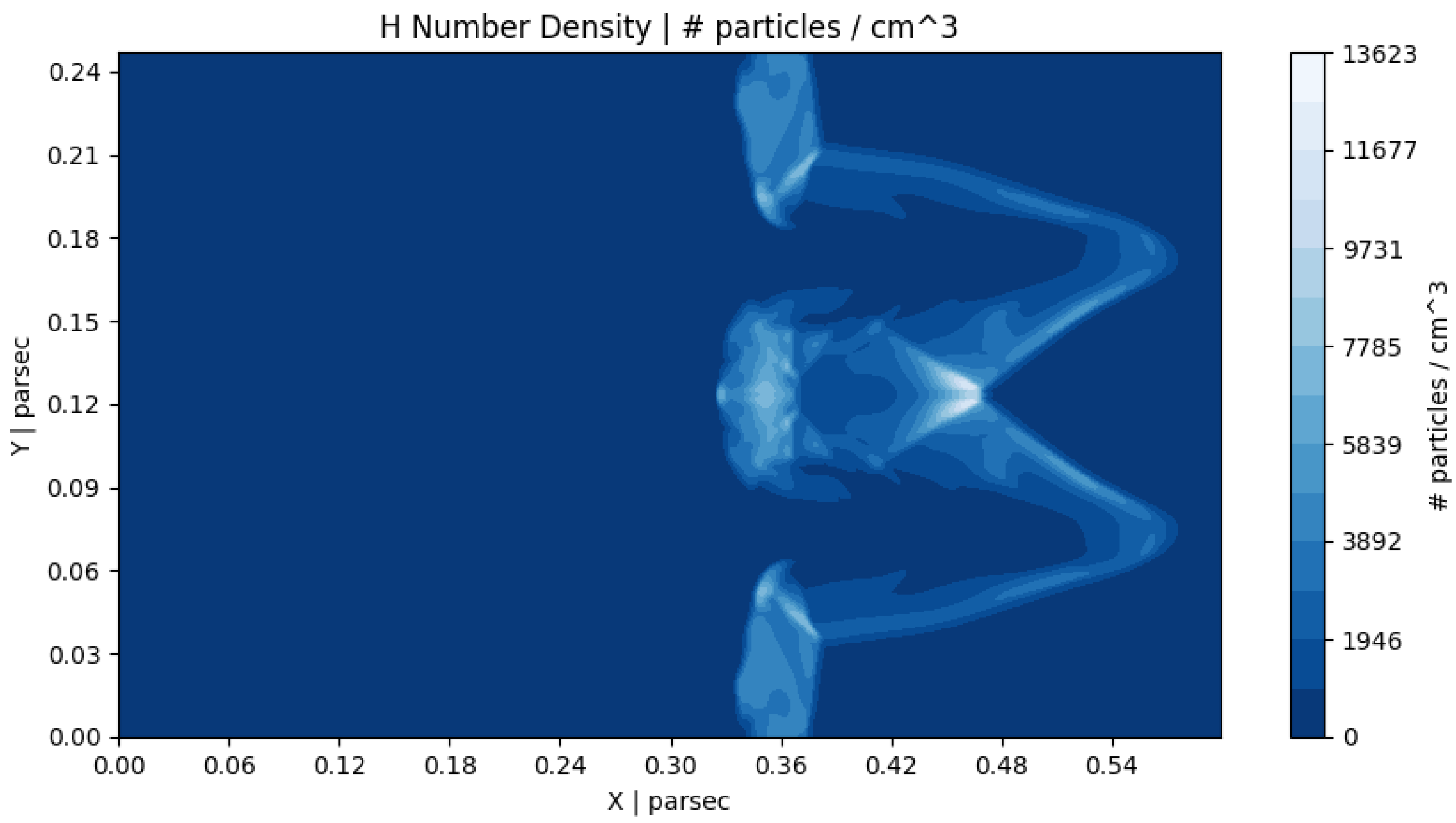

Figure 7.2: Hydrogen Number Density at timestamp 99 [12.5 kYear]

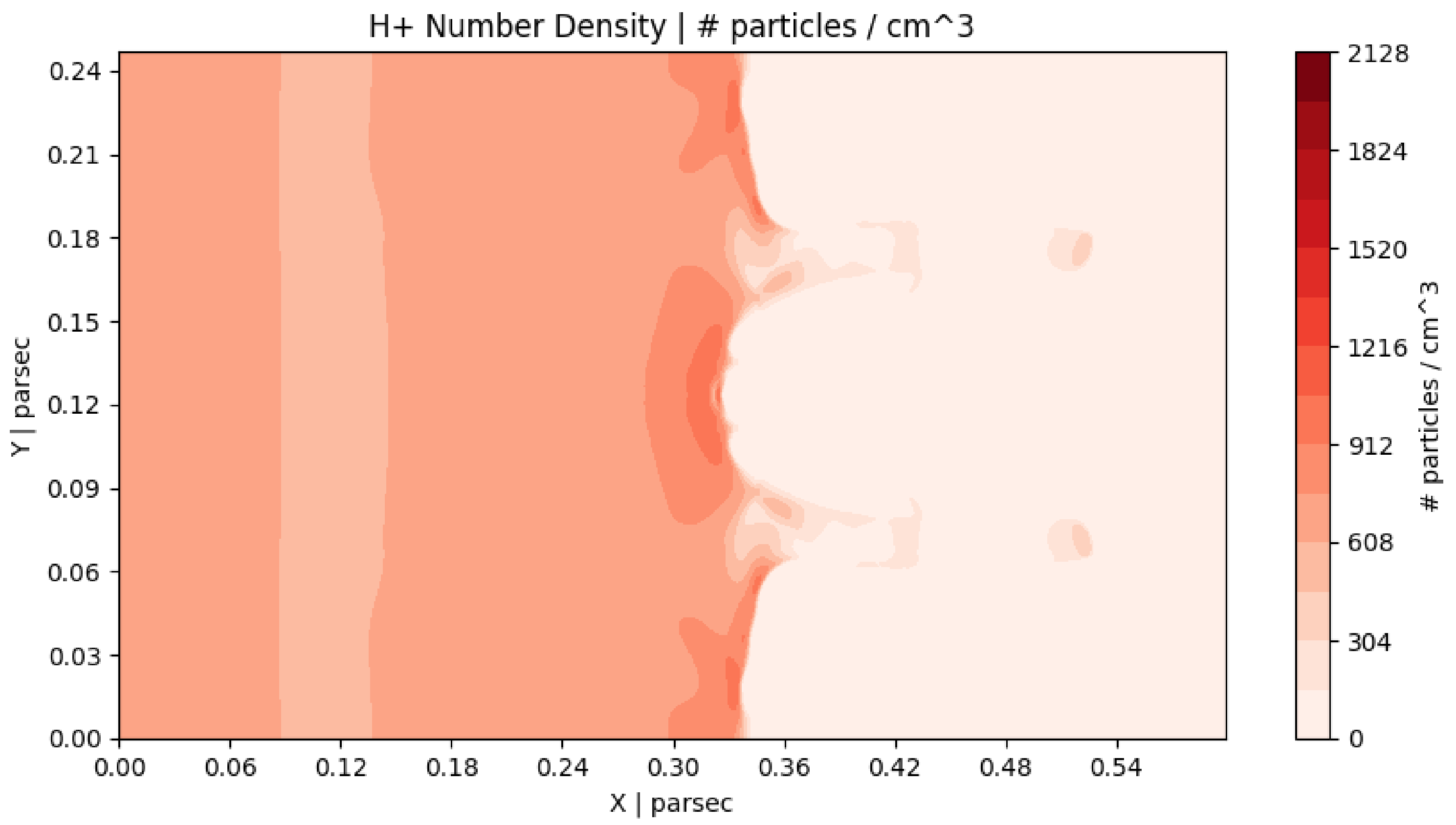

Figure 7.3: Hydrogen+ Number Density at timestamp 99 [12.5 kYear]

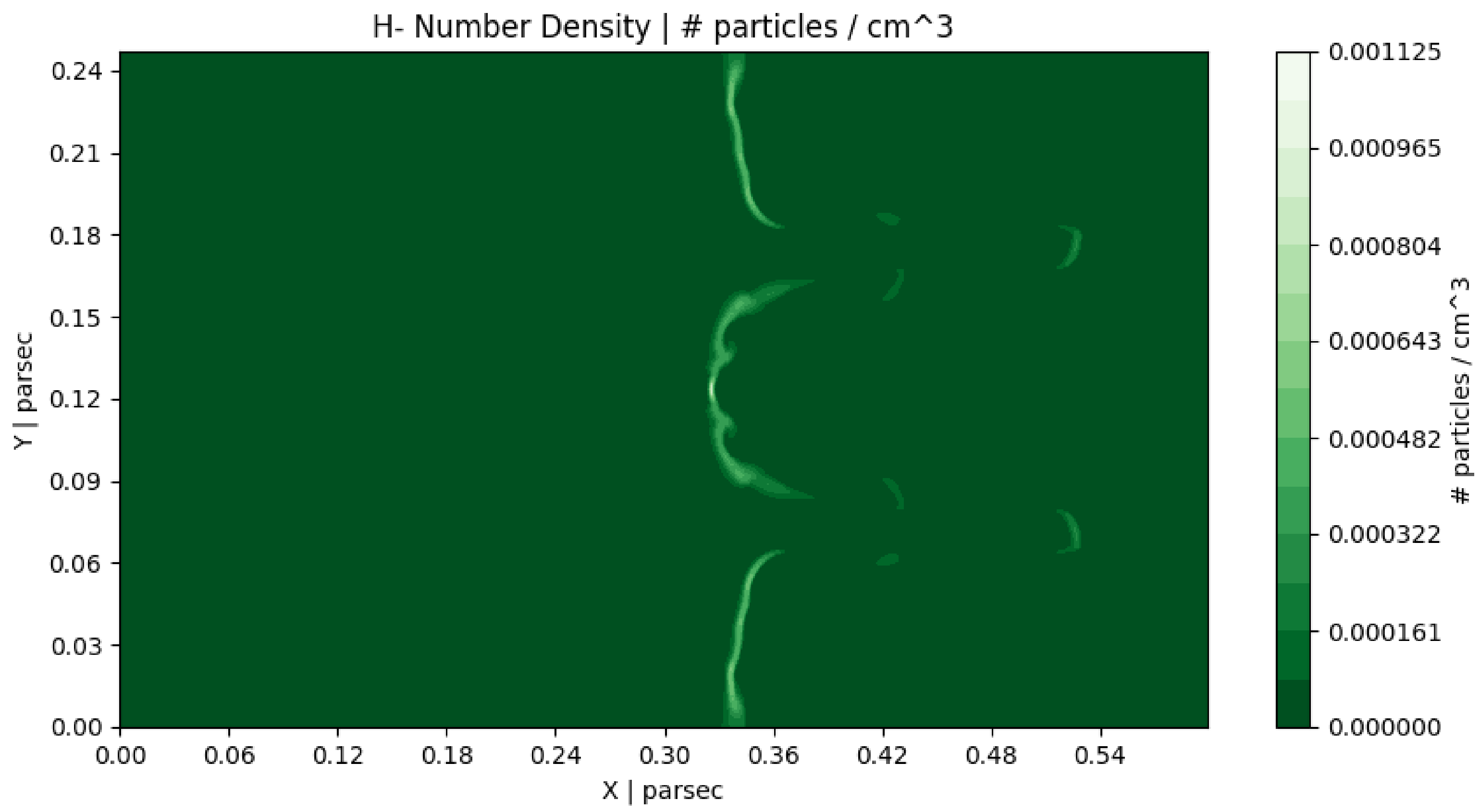

Figure 7.4: Hydrogen- Number at timestamp 99 [12.5 kYear]

Following inferences can be made looking at the above visualizations -

- The ionization front rapidly heats up a region which then continues to heat up even after the ionization front passes on.
- The number of particles are more near the ionization front meaning that the ionization front compresses the gases near the front.
- Looking at the plots of the H, H+ and H- number densities, the following can be said -
  - Not all the Hydrogen particles are converted to Hydrogen+ ions and Hydrogen-. This can be said by looking at the maximum number of particles in each graph.
  - The conversion rate of Hydrogen to Hydrogen+ is far greater than Hydrogen to Hydrogen-.
  - The Hydrogen- ions only exist for a small period of time and only exist right behind the ionization front. Hence are very unstable compared to H+ ions.
- Looking at the Gas Temperature, H+ and H- Number Density plots we can say the the temperature of H+ ions is far greater than the temperature of H- ions.

Conclusion

In conclusion, the report has successfully elaborated on the visualizations performed and the tools & methodology used to perform the required tasks.