Visual Analysis of Structure Formation in Cosmic Evolution

Karsten Schatz* Christoph Müller* Patrick Gralka* Moritz Heinemann* Alexander Straub*
Christoph Schulz* Matthias Braun* Tobias Rau* Michael Becher* Patrick Diehl† Dominic Marcello†
Juhan Frank† Thomas Müller‡ Steffen Frey* Guido Reina* Daniel Weiskopf* Thomas Ertl*

ABSTRACT

The *IEEE SciVis 2019 Contest* targets the visual analysis of structure formation in the cosmic evolution of the universe from when the universe was five million years old up to now. In our submission, we analyze high-dimensional data to get an overview, then investigate the impact of Active Galactic Nuclei (AGNs) using various visualization techniques, for instance, an adapted filament filtering method for detailed analysis and particle flow in the vicinity of filaments. Based on feedback from domain scientists on these initial visualizations, we also analyzed X-ray emissions and star formation areas. The conversion of star-forming gas to stars and the resulting increasing molecular weight of the particles could be observed.

Index Terms: Human-centered computing—Visualization;

1 INTRODUCTION

The *SciVis Contest 2019* is dedicated to the visual analysis of a cosmological simulation that models the evolution of the universe from the age of five million years ($z = 200$) up to now ($z = 0$), where z is the redshift. The simulation, run using the *Hardware/Hybrid Accelerated Cosmology Code (HACC)* [7, 11], comprises 64^3 dark matter particles and 64^3 baryon particles contained in a cubic domain with a side length $64\,{}^{Mpc}/_{h}$ (galactic distance with h reflecting the uncertainty of the Hubble constant). The task of the contest was to summarize the provided data set. No further details were provided. Furthermore, the organizers of the contest recommended including data derived from the ones provided by the simulation, namely the temperature of the baryons and thermodynamic entropy. Moreover, studying the impact of Active Galactic Nuclei (AGNs) on their surroundings is in line with the general goal of the HACC simulation.

Our approach to the analysis adopts a variety of different techniques from all areas of visualization: We use parallel coordinates plots (PCPs) and scatter plot matrices (SPLOMs) to provide an overview of the quantities and to discover patterns in multi-dimensional data. Additionally, these techniques facilitate selection and filtering using brushing and linking with 3D views of the particles or the vector field, thus enabling the user to locate specific points and ranges in space. As additional means to provide an overview and detailed analysis, we employ direct displays of the particles as spheres with different quantities being mapped to size and color, arrow glyphs to detect the velocity of the particles as well as direct volume rendering and iso-surface rendering of reconstructed particle densities. As the majority of matter in the universe resides in galactic filaments and many interesting processes like star formation happen there [1], we developed a filament detection algorithm that can separate these structures from the surrounding medium. For the velocity vectors, we use surface line integral convolution (LIC), which allows us to show the direction of the particles on the surface of dense areas, for instance, on filament structures. Furthermore, we want to obtain a visual representation of the data that can be related to quantities observed in the real universe. To this end, we compute a scalar field showing the X-ray emission in space and time from the given data. All these techniques are smoothly combined in a visual analytics tool based on MegaMol [8], which facilitates high-performance rendering for interacting with large simulation data. Our approach allows researchers to explore, summarize, and quantify the structure formation in cosmic evolution. The presented system is the result of a collaboration between domain scientists and visualization experts.

*Visualization Research Center (VISUS), University of Stuttgart, Germany. e-mail: *firstname.lastname*@visus.uni-stuttgart.de

†Louisiana State University, Baton Rouge, United States. e-mail: {frank,dmarcello}@phys.lsu.edu, pdiehl@cct.lsu.edu

‡Max Planck Institute for Astronomy, Heidelberg, Germany. e-mail: tmueller@mpia.de

2 OVERVIEW OF THE DATA

We first want to report on the different steps of the analysis process for the simulation data using our system. In this section, we start with an overview of the data—following the visual information seeking mantra by Shneiderman [16]. The subsequent sections address various details of the analysis, as well as adapted or novel techniques.

Initially, we tried to inspect the first and last simulation step by depicting particles as solid glyphs. However, this approach was not viable since particles are evenly distributed initially and throughout the simulation high particle densities lead to significant occlusion. To account for the multivariate nature of the data set and to quickly filter basic quantities present in the simulation data, we decided to incorporate a SPLOM and a PCP.

The first time step is shown in Fig. 1(a). Please note the evenly distributed position components, indicating that the particles are initially positioned on a uniform grid. While this circumstance would also have become visible with a simple spatial volume rendering, one might have missed the non-evenly distributed high velocities and gravitational potential. A little less obvious is the equal distribution of the two classes of particles (all baryon particles and dark matter particles have equal mass each). Presumably, these are the initial values of the simulation. The last time step is shown in Fig. 1(b). Contrary to the beginning of the simulation, one can see the fine filament structures emergent in the position as well as the normally distributed velocity. As expected, we also see a strong correlation between temperature and internal energy. Moreover, one can recognize that large smoothing length correlates with slow particles, which is also backed up by correlation with the gravitational potential. Furthermore, entropy is not evenly distributed. The PCP at the top allows for easy detection of correlations between the given magnitudes, as the column order is interchangeable. It also provides brushing and linking based filtering options on the whole data as well as single particles as shown in red in Fig. 1(b).

When inspecting all time steps as an animation (cf. supplemental video), we noticed some leaping values such as molecular weight and density around time step 35 and an overall trend that everything is slowing down, which was to be expected.

3 FILAMENT DETECTION

Especially in later simulation steps of the universe, we observed matter forming filament structures. Based on previous research [1],

we knew those filaments and their surroundings to be of high interest to domain experts as they contain more than half the matter of the known universe. While purely density-based approaches, such as direct iso-surface visualization, usually result in recognizable filament structures, they also tend to cause visual clutter in other regions. Hence, we developed our own filament detection algorithm to filter out the filament structures more reliably and thus facilitate analysis. Our approach is based on the friends-of-friends (FOF) algorithm of Davis et al. [5]. Their algorithm was originally designed to detect dark matter halos. Due to an inherent drawback—the tendency to generate elongated halos—adapted versions were published by other groups [13]. This drawback, however, comes in handy for our needs, as the filament structures we want to detect are elongated by nature. Besides the particle data, the method has only one input value, namely the so-called *linking length* l. Particles belong to the same cluster, if they are closer together than l. Filament structures are typically the densest regions of the data set. Thus, we adapt the original method by only selecting clusters that host at least one particle with a density value larger than 90% of the maximum density; all other clusters are filtered out. Subsequently, small-sized clusters with a particle count below a certain threshold (1000 in our case) are also dropped. We chose a linking length of $l = 450\ {}^{kpc}/_{h}$, which is larger than one would typically use for halo detection. This is due to the fact that the structures we aim to detect now are also larger. Results of our detection method are given in Fig. 2.

4 Active Galactic Nuclei Structure

The particle flow in the vicinity of filaments allows us to understand the impact of AGNs on structure formation. We first compute a density representation from the given particles by splatting density kernels around the positions of the particles, accumulating at a point in space (akin to Westover [18]):

$$\rho(\mathbf{x}) = \sum_{p \in \Omega} k\left(\left|\left|\mathbf{x} - \mathbf{x}_p\right|\right|, h\right) \rho_p\,, \tag{1}$$

where p is a particle at position $\mathbf{x}_p$ in domain Ω, ρ_p is its density, and k is a radial basis function (RBF) with the smoothing length h as kernel size. With this density volume, we can now extract an iso-surface encompassing areas of high density.

We use the reconstructed surface to more closely investigate the movement of particles in its vicinity. Furthermore, we estimate velocity information at arbitrary positions in a manner similar to the density splatting described above. However, contrary to Eq. 1, we have to normalize the resulting velocity by dividing by the sum of the RBFs since we want to interpolate velocity:

$$\mathbf{u}(\mathbf{x}) = \frac{\sum_{p \in \Omega} k\left(\left|\left|\mathbf{x} - \mathbf{x}_p\right|\right|, h\right) \mathbf{u}_p}{\sum_{p \in \Omega} k\left(\left|\left|\mathbf{x} - \mathbf{x}_p\right|\right|, h\right)}\,, \tag{2}$$

with velocity $\mathbf{u}_p$ at position $\mathbf{x}_p$.

The resulting 3D vector field allows us to depict the flow of particles on the extracted surfaces. For this, we use a hybrid screen-space variant of line integral convolution (LIC) [4, 17]. To capture the flow on the surface, we project the velocity onto the surface and use this projection for the streamline computation. We convolve the values from a noise texture along the streamline using a radial basis function to yield a texture depicting flow direction. Velocity magnitude is encoded by color. To distinguish between near and far objects, we add visual highlights depending on the surface normal and a fixed light source. The result is shown in Fig. 3. The small differences in magnitudes (c) indicate that the velocities are mostly tangential to the surface. However, additionally showing arrows in the vicinity of the cluster in (d) allows us to infer that the particle flow is directed toward the surface before reaching it. Fig. 6 further illustrates this behavior.

5 X-ray emission

The particles in the data set depict matter that emits electromagnetic waves. Some of these particles emit in the X-ray spectrum, which is of particular interest. For example, X-ray emission of astronomical objects [10] is related to these objects containing hot gases (around one million Kelvin or higher). X-ray emission from the hot gas is expected where structures such as cluster dominant galaxies and AGNs are growing at the expense of the surrounding material. Thus, the reconstruction of the X-ray emission enables to validate the assumption that AGNs should emit X-rays. To show the distribution of X-ray emission in the data set, we implemented a combination of raytracing and raycasting to generate and render a volume that captures the X-ray emission. Our approach is inspired by the visualization model of Magnor et al. [14] and the irradiance volume technique of Greger et al. [9].

We accumulate the X-ray emission for each voxel and attenuate it by the absorption, but neglect other effects, such as scattering or other models like radiance distribution functions. The overall process comprises two steps. In the first step, we determine the X-ray emissivity of each baryon particle with Eq. 3, derived from discussions with our domain scientists and in line with Dolag et al. [6], approximating emissivity from Bremsstrahlung

$$\varepsilon_X = \rho^2 \sqrt{T}, \tag{3}$$

where ρ is density and T is temperature of a particle.

In the second step, we shoot rays from each particle position into the volume to capture the emission. The ray origins are uniformly distributed over the full sphere and offset by the smoothing length to avoid "self-illumination". At each sample position, we spawn additional rays with directions sampled within a cone around the original direction, which is inspired by cone tracing [2, 12]. We add the emission attenuated by the absorption coefficient shown in Eq. 4 to each voxel hit by the respective ray with a fixed step size:

$$\alpha = 0.018 \sqrt{T^{-3}}\ 1.4 \left(\frac{\rho}{m}\right)^2 g\,. \tag{4}$$

Eq. 4 is adapted from equation 5.19b from Rybicki and Lightman [15]. Since we have no information about the number density of electrons or ions, we used density ρ and mass m via the term $1.4\,{}^{\rho^2}/_{m^2}$, a replacement suggested by our domain scientists for the term $Z^2 n_e n_i$ in the original equation. Moreover, we did not include the frequency in the equation and took 1.2 as the Gaunt factor g (cf. Sect. 5.2 [15]). For both equations, we normalized the resulting quantities prior to the calculation of the final volume and rendering, because the rendering is intended as a qualitative illustration.

Note that the quantities in Eq. 3 are based on particles. The quantities in Eq. 4 are splatted into volumes (cf. Sect. 4) prior to ray traversal in the pre-processing steps. Once the X-ray emission volume is ready (and the voxel values normalized), we render it with a classic direct volume raycasting approach. As can be seen in Fig. 4, this approach highlights possible AGN candidates and AGN particles, which helps guide the user attention to these areas independent of the currently depicted quantities.

6 Star Formation

Based on the feedback of our domain scientists, we also analyzed the formation of stars, which can happen in cold areas with high density. Fig. 5 shows the top 10 % of the densest particles as well as the bottom 10 % of the coldest particles.

Especially in the video, one can observe star-forming gas being converted to stars, as expected. Preceding to the particles being detected as star-forming gas, the molecular weight of the particles is increasing. However, this only seems to be a necessary, but not a sufficient condition for star formation.

Furthermore, one can observe the formation of AGNs in the vicinity of star clusters as more and more star-forming gas starts to accumulate in those areas. However, the AGN classification is being revoked as the star-forming gas starts to disappear. Retrospectively, this behavior makes sense since AGNs seem to suspend their activity without further mass such as star-forming gas nearby. The presumption is reinforced by the observation that some AGNs keep waking up when fueled by new gas to then suspend again.

7 Conclusion

We have analyzed a cosmic evolution using a visual analytics system based on the software framework MegaMol. MegaMol was instrumental to our analysis, because of its flexible configuration and easy extensibility for fast prototyping and scalable data analysis. From our experience, it was crucial to include domain scientists using pair analytics [3]. In this way, we could explore questions that went beyond the original tasks of the *SciVis 2019 Contest*.

Acknowledgments

This work was partially funded by the Center of Computation & Technology at Louisiana State University; by Deutsche Forschungsgemeinschaft (DFG) as part of the Cluster of Excellence EXC 2075 "SimTech" (390740016), Transregional Collaborative Research Centre SFB/Transregio 75 (84292822) and SFB/Transregio 161 (251654672), Collaborative Research Centre SFB 1244 (279064222), the International Research Training Group GRK 2160 "DROPIT" (270852890), and "Research Software Sustainability for the Open-Source Particle Visualization Framework MegaMol" (391302154); by the German Bundesministerium für Bildung und Forschung (BMBF) as part of project "TaLPas" (Task-based Load Balancing and Auto-tuning in Particle Simulations); by Baden-Württemberg Stiftung as part of project "DiHu" (High Performance Computing II grant); and by Intel Corporation as part of the Intel Graphics and Visualization Institutes of XeLLENCE program.

References

[1] P. Ade, N. Aghanim, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. Balbi, A. Banday, R. Barreiro, et al. Planck intermediate results–viii. filaments between interacting clusters. *Astronomy & Astrophysics*, 550:A134, 2013.

[2] J. Amanatides. Ray tracing with cones. *ACM SIGGRAPH Computer Graphics*, 18(3):129–135, 1984.

[3] R. Arias-Hernández, L. T. Kaastra, T. M. Green, and B. D. Fisher. Pair analytics: Capturing reasoning processes in collaborative visual analytics. In *Proceedings of the 44th Hawaii International Conference on Systems Science (HICSS-44 2011)*, pages 1–10, 2011.

[4] B. Cabral and L. C. Leedom. Imaging vector fields using line integral convolution. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH*, pages 263–270, 1993.

[5] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *The Astrophysical Journal*, 292:371–394, 1985.

[6] K. Dolag, S. Borgani, S. Schindler, A. Diaferio, and A. M. Bykov. Simulation techniques for cosmological simulations. *Space Science Reviews*, 134(1):229–268, 2008.

[7] N. Frontiere, C. D. Raskin, and J. M. Owen. CRKSPH–a conservative reproducing kernel smoothed particle hydrodynamics scheme. *Journal of Computational Physics*, 332:160–209, 2017.

[8] P. Gralka, M. Becher, M. Braun, F. Frieß, C. Müller, T. Rau, K. Schatz, C. Schulz, M. Krone, G. Reina, and T. Ertl. MegaMol – a comprehensive prototyping framework for visualizations. *The European Physical Journal Special Topics*, 227(14):1817–1829, 2019.

[9] G. Greger, P. Shirley, P. M. Hubbard, and D. P. Greenberg. The irradiance volume. *IEEE Computer Graphics and Applications*, 18(2):32–43, 1998.

[10] H. Gursky. The x-ray emission from rich clusters of galaxies. *Publications of the Astronomical Society of the Pacific*, 85:493, 1973.

[11] S. Habib, A. Pope, H. Finkel, N. Frontiere, K. Heitmann, D. Daniel, P. Fasel, V. Morozov, G. Zagaris, T. Peterka, et al. HACC: Simulating sky surveys on state-of-the-art supercomputing architectures. *New Astronomy*, 42:49–65, 2016.

[12] D. B. Kirk. The simulation of natural features using cone tracing. *The Visual Computer*, 3(2):63–71, 1987.

[13] A. Knebe, S. R. Knollmann, S. I. Muldrew, F. R. Pearce, M. A. Aragon-Calvo, Y. Ascasibar, P. S. Behroozi, D. Ceverino, S. Colombi, J. Diemand, et al. Haloes gone MAD: the halo-finder comparison project. *Monthly Notices of the Royal Astronomical Society*, 415(3):2293–2318, 2011.

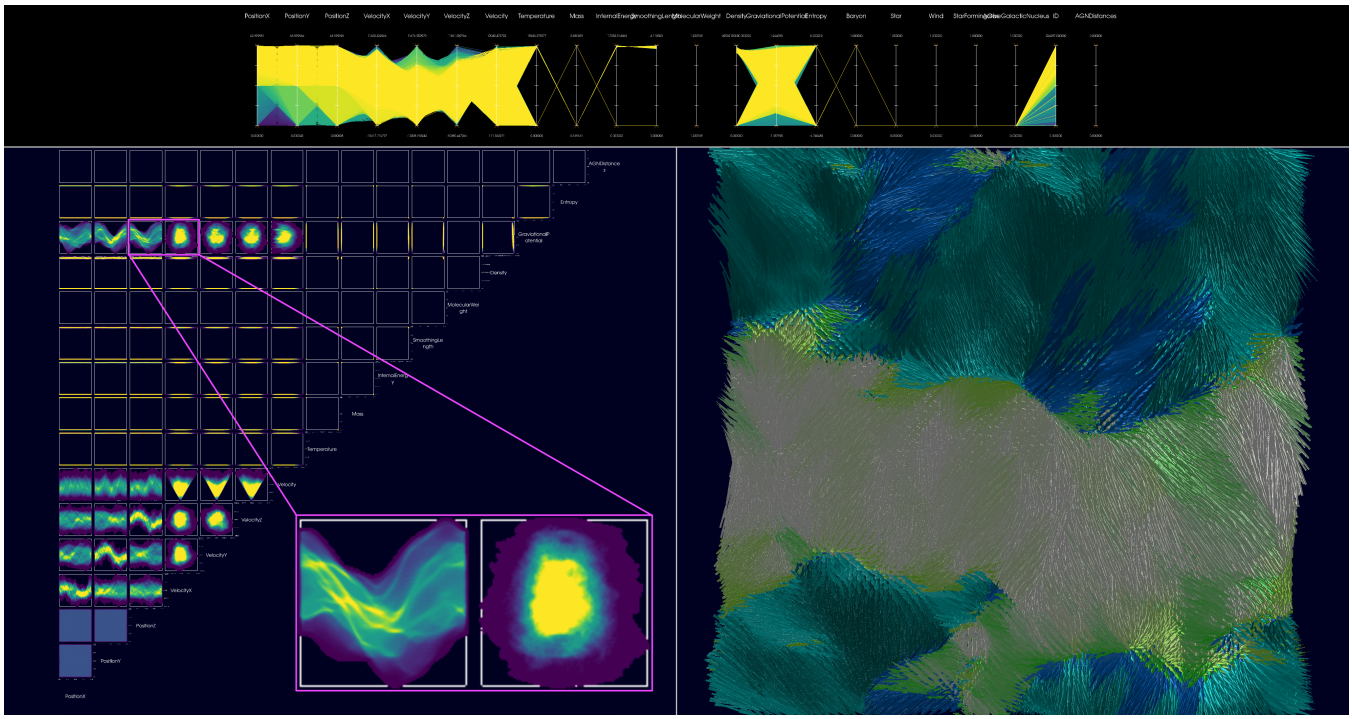
[14] M. A. Magnor, K. Hildebrand, A. Lintu, and A. J. Hanson. Reflection nebula visualization. In *Proceedings of the IEEE Visualization conference*, pages 255–262, 2005.

[15] G. B. Rybicki and A. P. Lightman. *Radiative Processes in Astrophysics*. John Wiley & Sons, 2004.

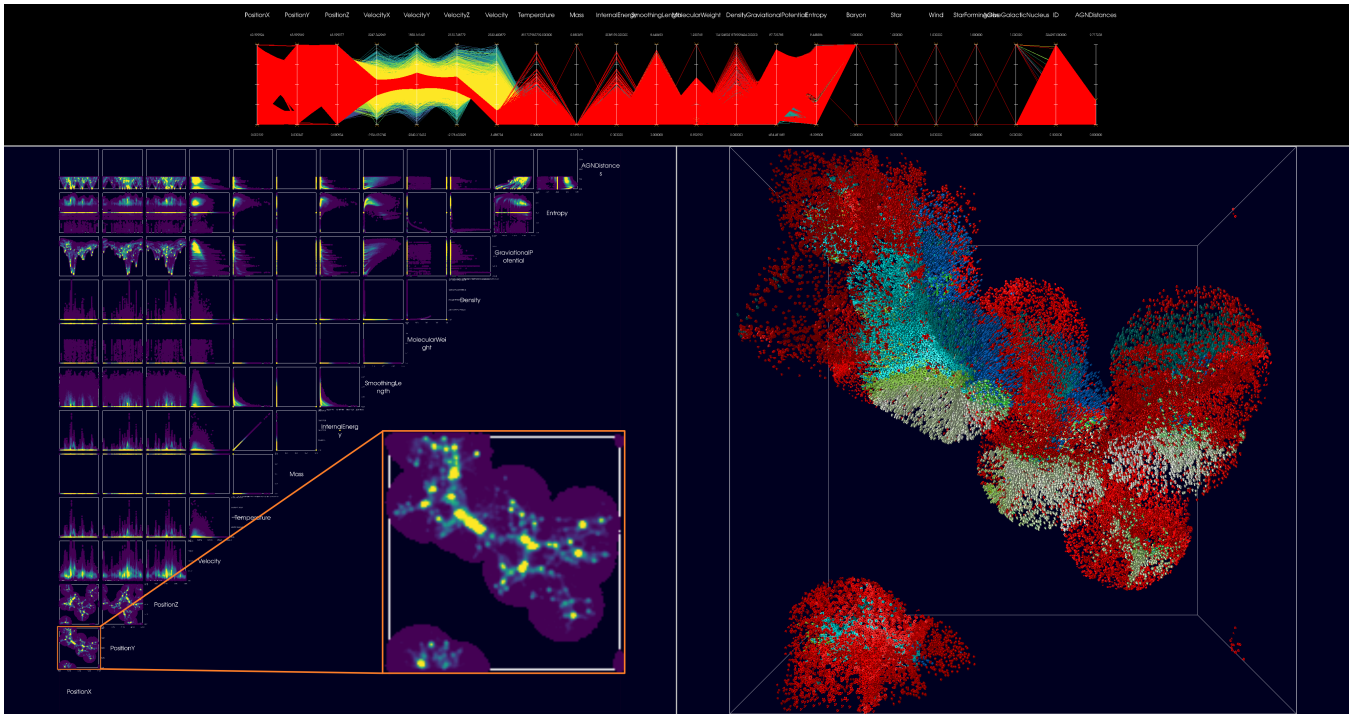
[16] B. Shneiderman. Eyes have it: A task by data type taxonomy for information visualizations. *Proceedings of IEEE Symposium on Visual Languages*, pages 336–343, 1996.

[17] D. Weiskopf and T. Ertl. A hybrid physical/device-space approach for spatio-temporally coherent interactive texture advection on curved surfaces. In *Proceedings of the Graphics Interface Conference*, pages 263–270, 2004.

[18] L. Westover. Footprint evaluation for volume rendering. *ACM Siggraph Computer Graphics*, 24(4):367–376, 1990.

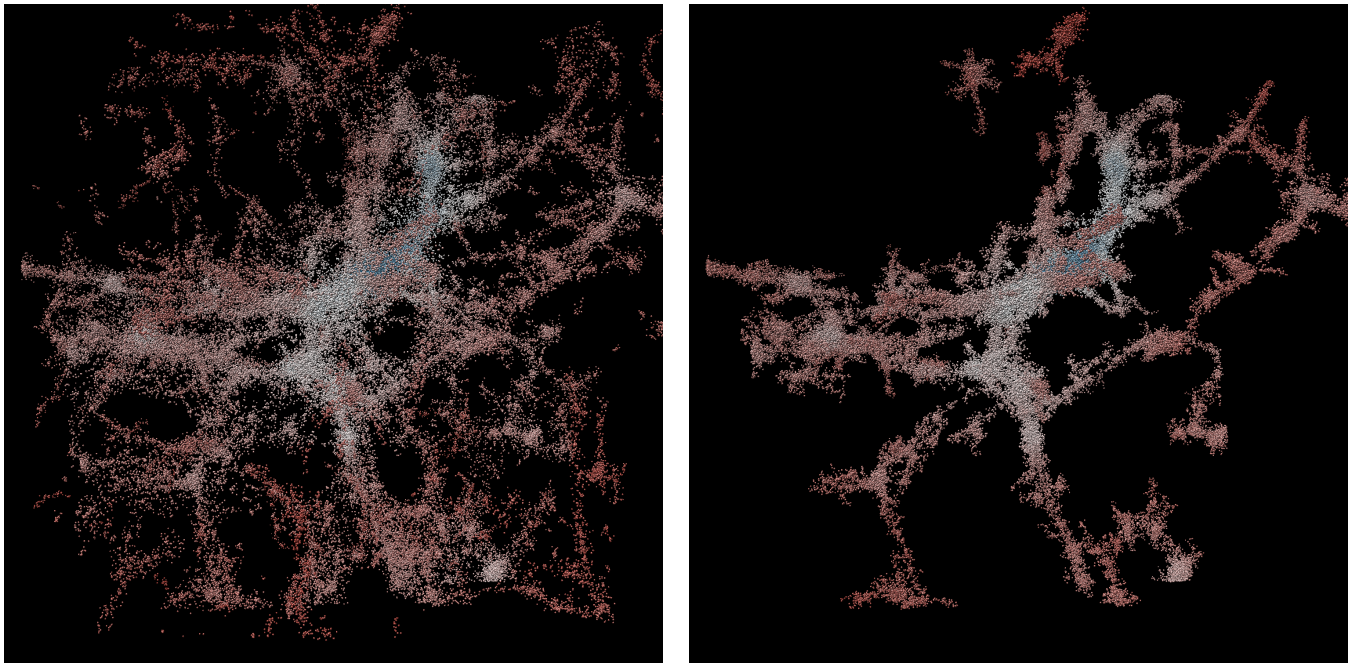


(a) First time step



(b) Last time step

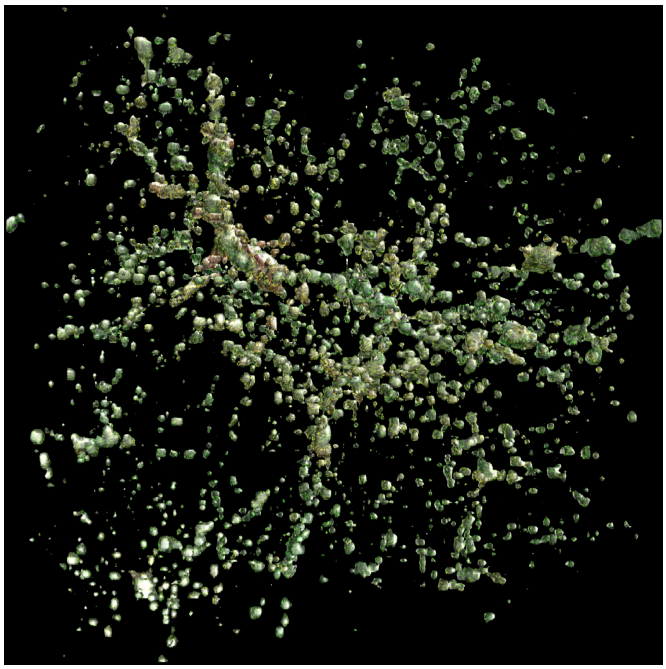
Figure 1: Linked visualizations providing an overview of the data. A PCP (top) and a SPLOM (left) with high-dimensional points aggregated as density and then mapped to color (kernel density estimation, low-to-high is blue-to-yellow). In addition, position and velocity are encoded into arrow glyphs (right). In the last time step, the PCP is used to filter for particles close to an AGN and to brush all particles with a low velocity. These distance values were not given in the original data and were precomputed by us during data loading. The highlighted and zoomed in areas depict: (a) uneven gravitational potential distribution and uneven velocity distribution, (b) particle distribution forming filament structures.



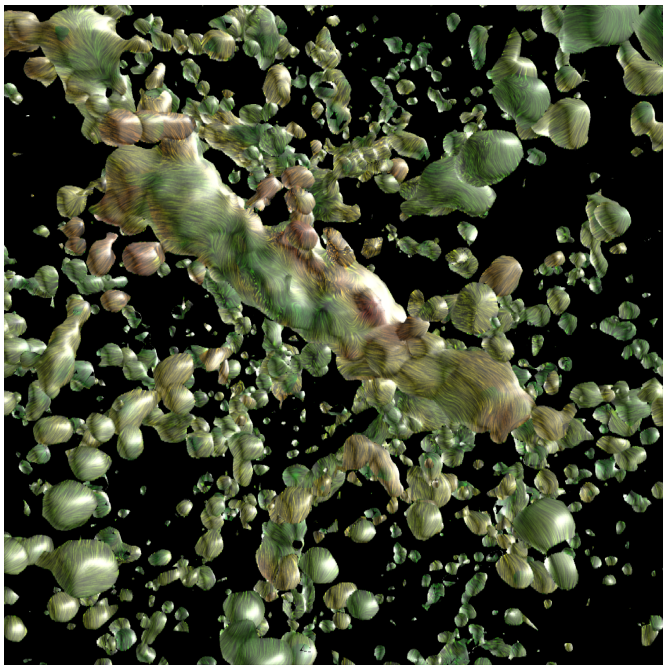
(a) Purely density-based filtering

(b) Our filament detection

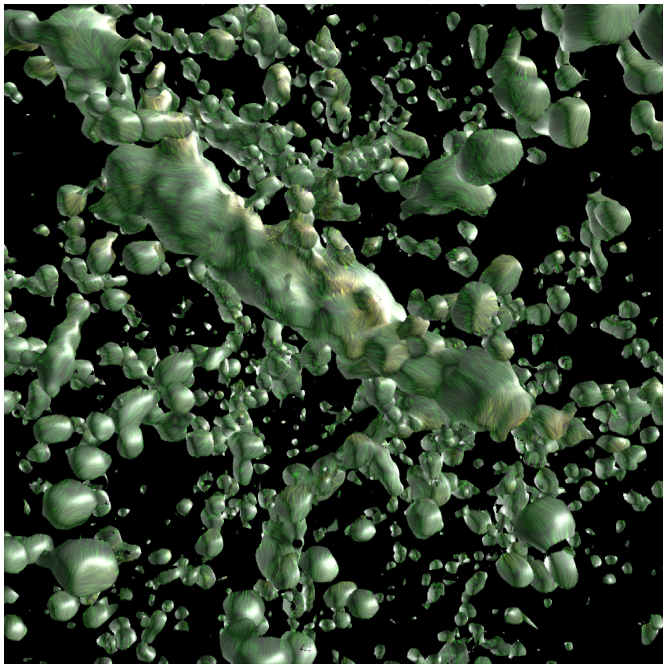
Figure 2: Filament detection examples with particles colored by gravitational potential. With the applied settings, smaller filaments are filtered out alongside unrelated particles. It is possible to include the smaller filaments filtered out in (b) by tweaking the parameters of the method, but this would also add some of the clutter that can be seen in (a).



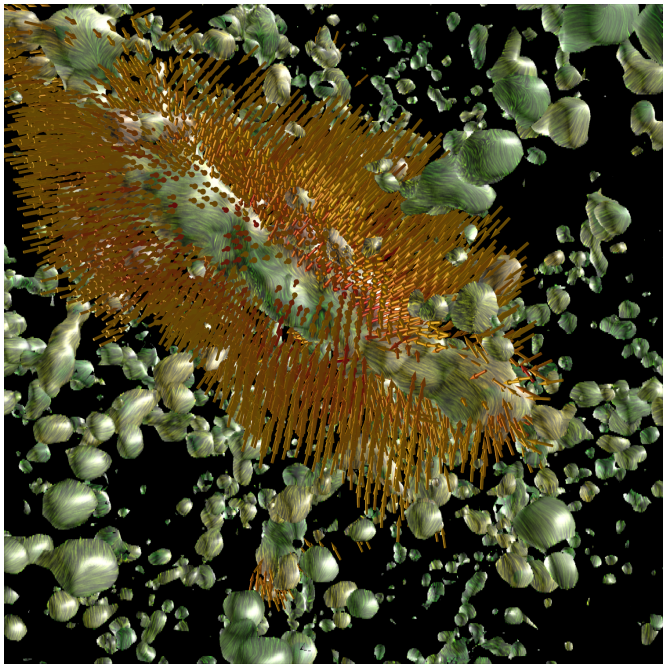
(a) Original velocities



(b) Closeup of original velocities



(c) Difference in magnitude between original and projected velocities



(d) Difference in magnitude with arrows

Figure 3: Surface LIC rendered on the reconstructed iso-surface. Colors indicate the velocity magnitude, ranging from low (green) to high (red). Orignal velocities are shown in (a) and (b). Difference between original and projected velocities are shown in (c). Arrow glyphs in (d) point toward dense structures, indicating continued growth as expected. The base of the arrow is located at the corresponding particle position.

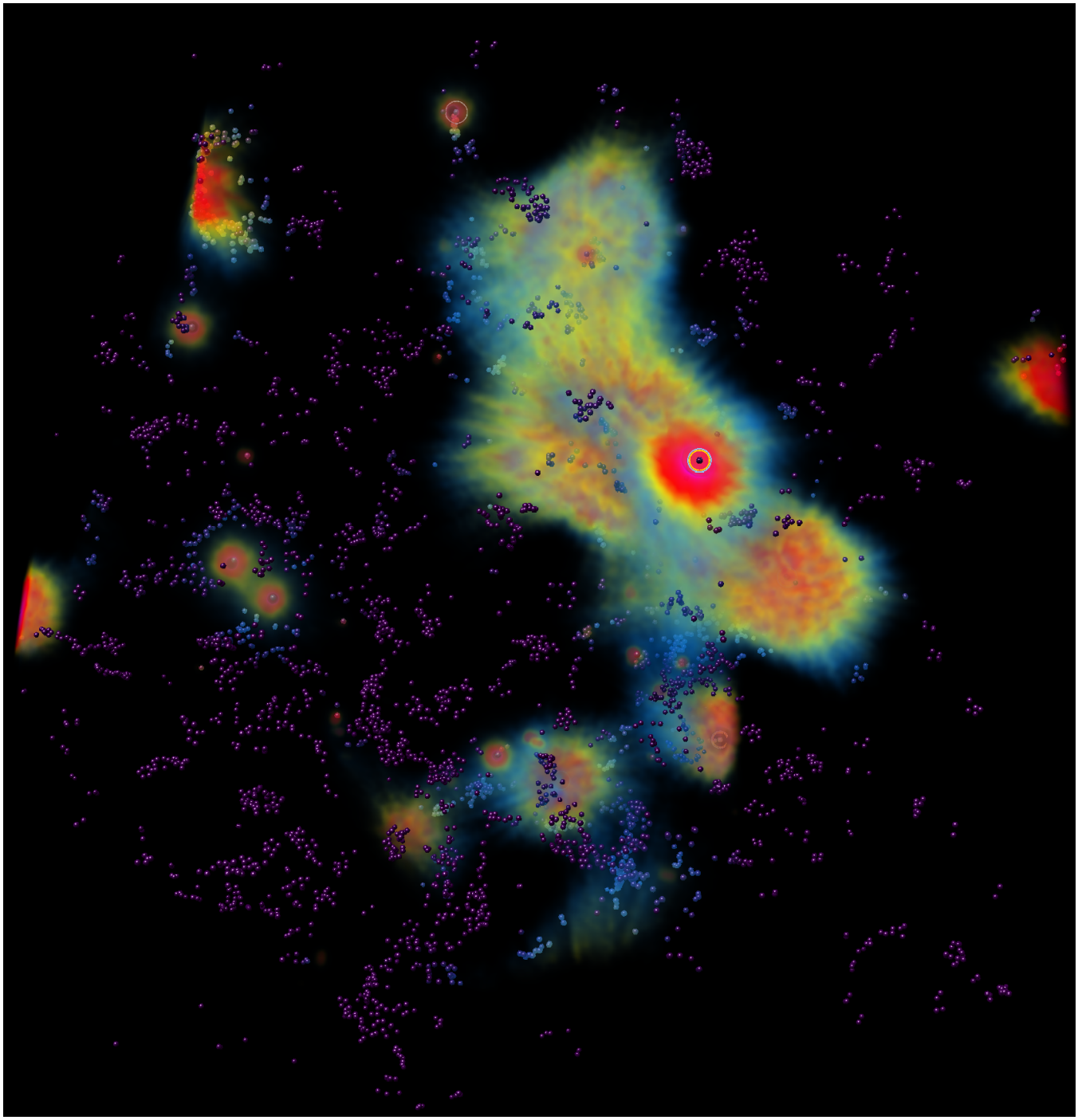
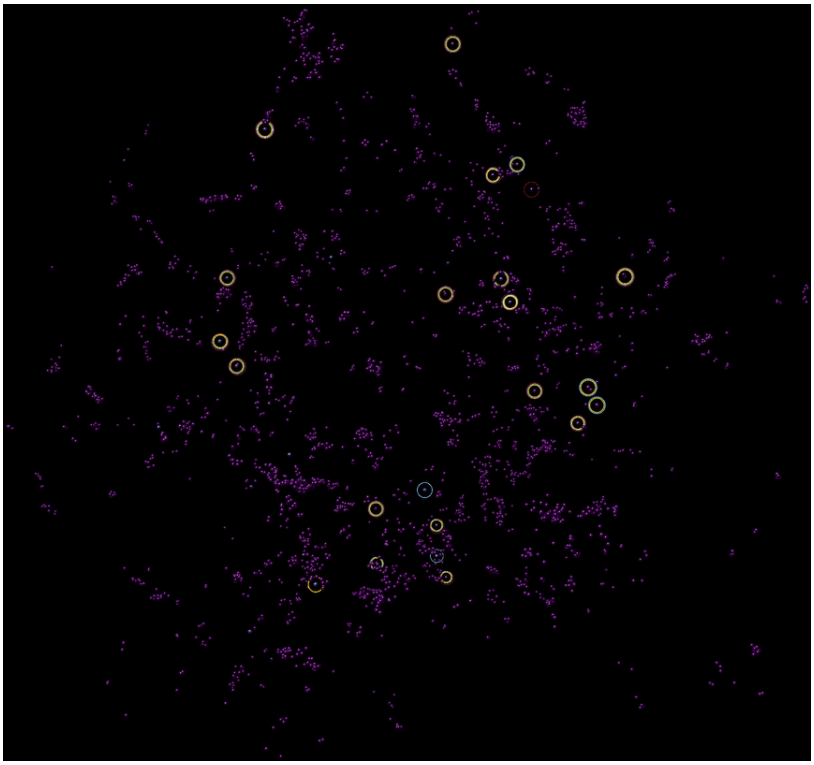
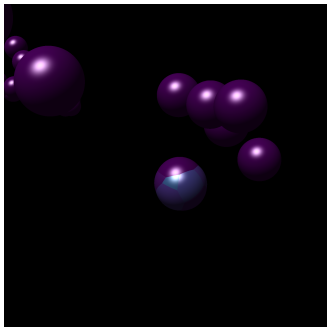


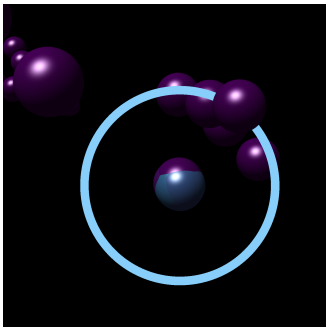
Figure 4: Raycasted X-ray emission volume of time step 490 rendered alongside with AGN or star candidate particles. The highest visible emission is radiated from the only AGN particle in the shown time step.



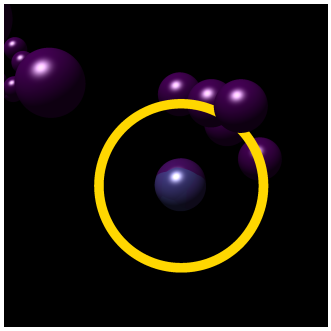
(a) Zoomed-out view



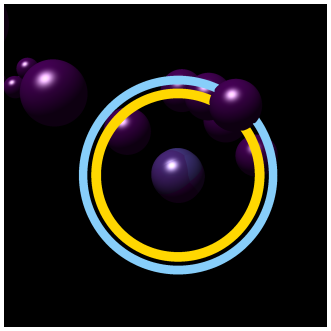
(b) Increased molecular weight



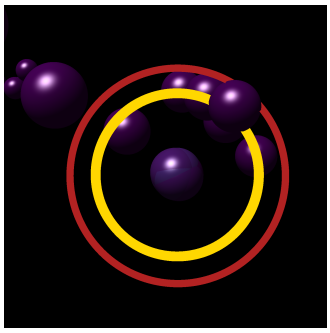
(c) Gas cloud forming



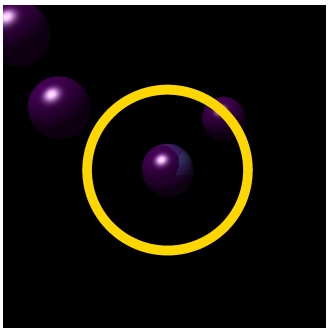
(d) Star creation



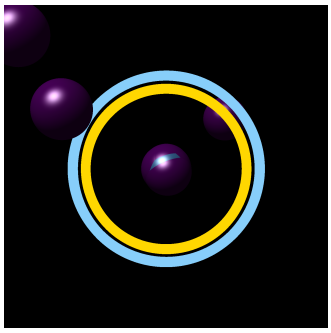
(e) Additional gas



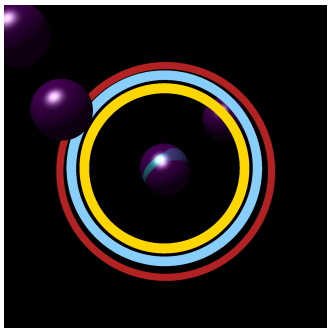
(f) AGN emerges



(g) AGN gone due to lack of gas

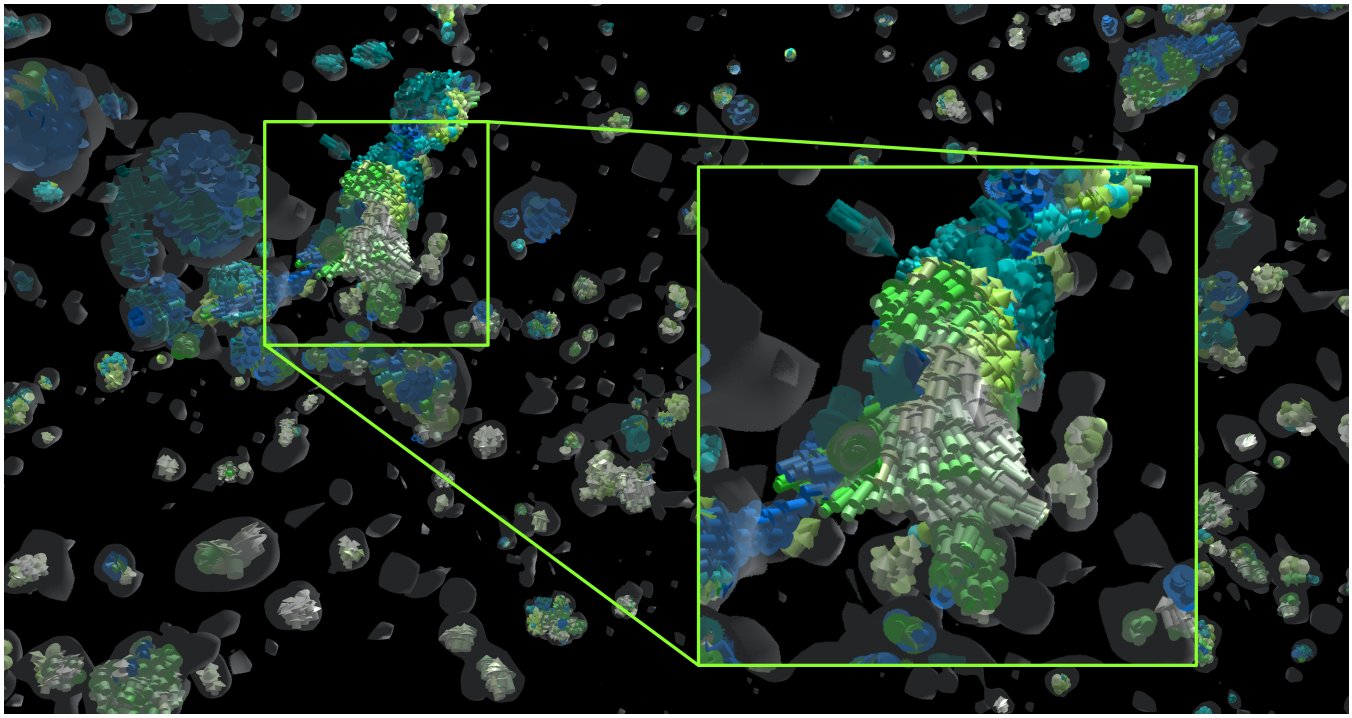


(h) Additional gas appearing

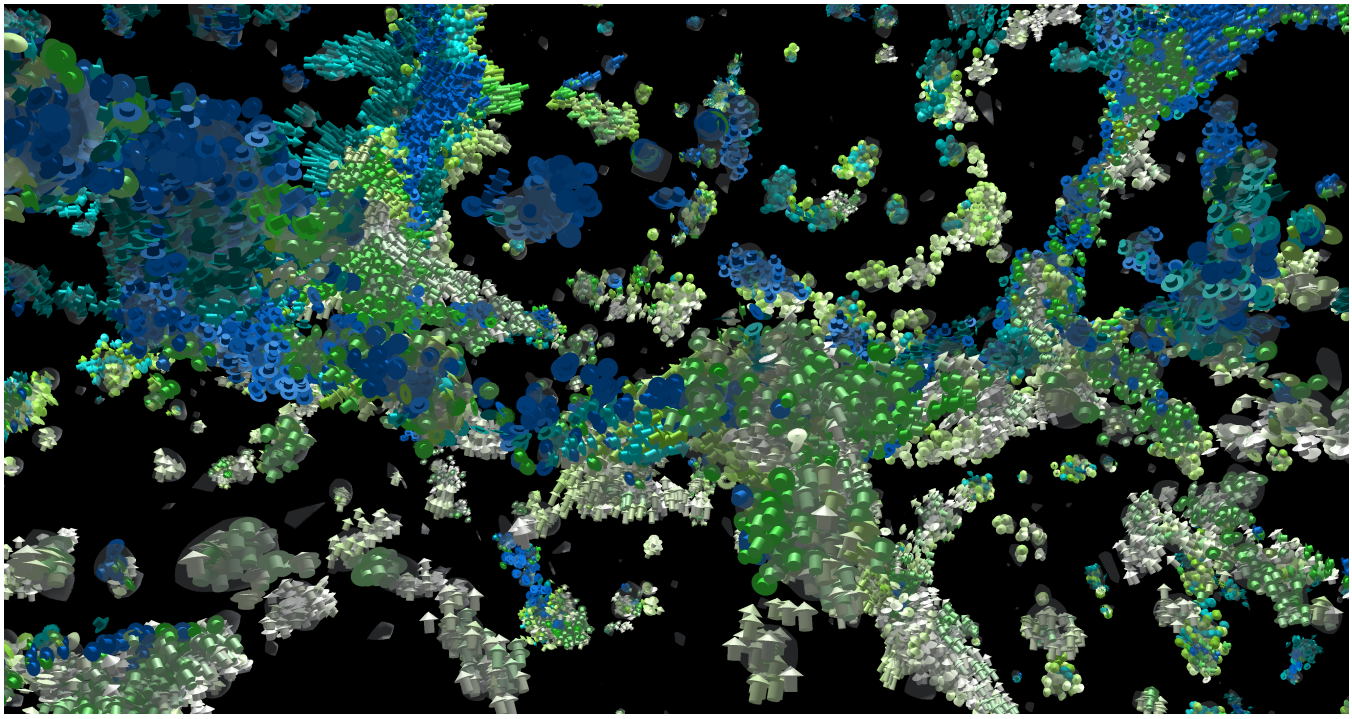


(i) AGN classification restored

Figure 5: Detailed illustration of particles filtered by density and temperature, focused on one specific AGN. The particles are colored by molecular weight (purple low, any other color higher). Additionally, circles highlight star-forming gas (blue), stars (yellow), and AGN particles (red). From (b) to (i), we depict the whole process of AGN creation and one period of the described fuel cycle ((f) to (i)).



(a) Internal particle movement in filaments



(b) External movement to the filaments

Figure 6: Density iso-surfaces depicted alongside with density-filtered velocity arrows for particles (colored by velocity direction). (a) Internal particle movement revealed by a higher density threshold. The inset shows how particles flow along the filament structure. (b) Outer particle movement by using a lower density threshold for filtering. As opposed to the internal ones, the outer particles move toward the filament structures. Their movement seems to be coordinated as many particles have a similar velocity direction.