Exploring Multistreaming in the Universe

Category: Application

Abstract— Visually striking large scale structures (LSS) such as galaxy filaments and clusters have been observed in cosmological surveys, and are composed of large numbers of galaxies. Even though there are theories for the formation and evolution of these cosmological structures, much remains to be understood, especially in the nonlinear regime of structure formation. This problem is now being attacked with the aid of high accuracy cosmological simulations. Associated with these simulations, there is a new challenge for data analysis. In this paper, we describe how visual analytics can help identify and characterize a particular set of features of interest in the evolution of the universe, the multistreaming events. As the name implies, different velocities (directions and magnitudes) can be observed in locations of such events. It is believed that these events are precursors to the formation of the LSS.

Index Terms—Cosmology, multistreaming, feature detection, velocity field.

1 Introduction

This paper describes the application of visualization techniques to identify and characterize multistreaming events (see Section 3 for details) in the evolution of the distribution of dark matter in the universe. We first give an overview of the cosmology problem including the importance of multistreaming in the formation of LSS in the initially smooth dark matter distribution. We then provide a brief review of previous and related works in cosmological visualization and detection of multistreaming events. Next, we re-examine the different contexts in which multistreaming is described and identify other criteria for detecting these events. Finally, we present our strategies and results for finding these multistreaming events.

An astonishing 99.6% of our Universe is "dark" and not directly observable by emission or absorption of light. Observations indicate that the Universe consists of 70% mysterious dark energy, 25% of a yet unidentified dark matter component referred to as cold dark matter, and only 0.4% of the remaining 5% of ordinary (baryonic) matter is visible. Understanding the physics of this dark sector is the foremost challenge in cosmology today. Although the ultimate nature of the dark sector is unknown, a detailed phenomenology has been constructed which is called the cosmological Standard Model. In this model, tiny density fluctuations in the very early Universe grow under the influence of gravity as the Universe expands. It is this density fluctuations that seed the formation of complex LSS in the matter distribution. Growing modes collapse into dark matter clumps called *halos* which attract and collect baryonic matter (gas) and eventually lights up as galaxies. The galaxies, being visible, can be used to track the dynamics and distribution of dark matter [14].

Understanding the evolution and dynamics of the dark matter can be achieved by following the formation of LSS as observed in the distribution of galaxies from the earliest moments until today. LSS such as galaxy clusters (0-D), filaments (1-D), and surface-like pancakes (2-D) can be considered to correspond to nodes, edges, and faces respectively, in a tessellation of the topology of the universe [2, 12]. These structures have complex geometry and topology as can be seen in Figure 1. Understanding the formation of the LSS is not only interesting in of itself but will also be very important to guide analytical and numerical studies in the quasi-linear regime of structure formation (where one can still apply perturbation theory techniques). Such studies will be essential to interpret results from major up-coming large-scale structure surveys, such as the Large Synoptic Survey Telescope (LSST), the Joint Dark Energy Mission, and the Baryon Oscillation Spectroscopic Survey (BOSS), a key component of the Sloan Digital Sky Survey III.

Precision dark matter simulations are a key foundation of cosmological studies. These simulations track the evolution of the dark matter with very high resolution in time, force, and mass. At the scales of interest to structure formation, a Newtonian approximation in an expanding universe is sufficient to describe gravitational dynamics. The evolution is given by a collisionless Vlasov-Poisson equation

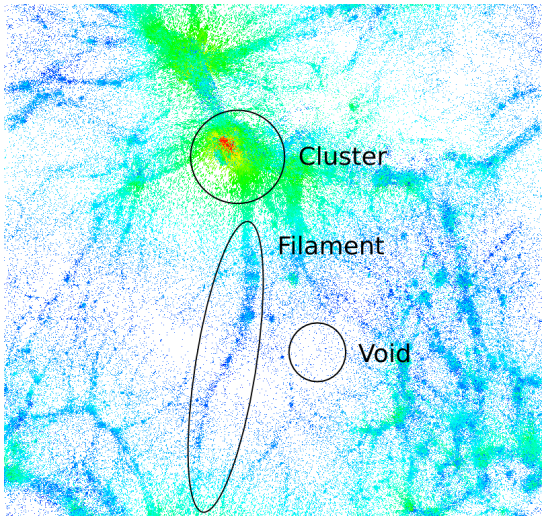

Fig. 1. Large scale cosmological structures of the universe.

[6], a six-dimensional partial differential equation. Thus, on memory grounds alone, a brute force approach is impractical. Having made their first appearance in plasma physics, N-body codes now form a standard approach for dealing with this problem. In the N-body approach, the six-dimensional phase space distribution is sampled by "tracer" particles and these particles are evolved by computing the inter-particle gravitational forces.

The starting point of the simulations is a Gaussian random density field which imprints small perturbations on a uniform density, isotropic universe. The simulations start in the linear regime of the density fluctuations and then evolve under the influence of gravity. At any given length scale, during the early stages the evolution remains linear but as time progresses, the evolution enters the quasi-linear regime before finally reaching the fully nonlinear regime at which point all analytic descriptions break down. There is substantial interest in determining and characterizing the transitions between linear, quasi-linear, and nonlinear dynamics in the simulations by tracking the dynamics of dark matter tracer particles. At the start of the simulation, the velocity dispersion is initially zero, and the phase-space distribution is a three-dimensional sub-manifold of the phase space (only one velocity direction at a given spatial point). As the 3-hypersurface evolves, it folds, leading to the occurrence of singularities in the density field corresponding to the appearance of regions with multistream flow.

The initial appearance of multistreaming is observed during the transition from linear to nonlinear regimes. This transition affects the accuracy of perturbative techniques used to analyze the evolution of

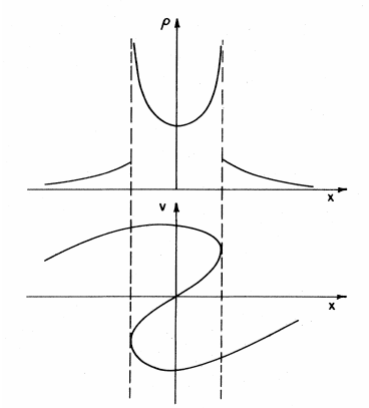

Fig. 2. Phase-space plot of a 1-D flow. A singularity (non-self-intersecting fold) is a necessary condition for multistreaming.

the density field. Multistreaming is therefore of interest both in terms of the underlying phenomenology and for understanding the validity of analytical methods.

Previously, LSS have been identified by thresholding on the density of tracer particles with halo finders (e.g. [10]. In this paper, we study how the velocity information of tracer particles can find and characterize the multistreaming regions which eventually lead to these LSS. This work is carried out in close collaboration between astrophysicists and computer scientists.

2 Previous Work

The visualization of cosmological data sets has received considerable attention recently. Since all cosmological simulations are particle-based, one of the popular tools for visualizing particles directly is Partiview [8]. Within the visualization community, there are also a handful of recent contributions dealing with astrophysics data sets. These include the work of Li et al. [9], focusing on how to display positional and trajectory uncertainties in astrophysical data sets. In the same year, Navratil et al. [13] described their visualization approach for a data set that study the formation and effects of the radiation from the first stars and the impact on subsequent star formation. Last year, Haroz et al. [5] investigated particle-based simulation data sets of the evolution of the universe and studied the relationship of different variables especially in the face of uncertainty arising from the different code settings, e.g. the starting time for the simulation. A similar data set was also investigated by Ahrens et al. [1] where the focus was on comparison rather than on uncertainty visualization.

There have also been several papers on multistreaming events. For example, Yano et al. [18] studied the distribution of caustics (see Section 3) in the expanding universe, while Gouda [4] investigated the relationship between catastrophe theory and gravitational clustering leading to caustics. In these studies, the models describe continuous matter density fields, e.g. density perturbations, singularities of density, etc. A good starting point for reading more about about multistreaming can be found in [3].

3 Multistreaming

Multistreaming is said to occur when there are multiple velocities at one point. A more formal description is illustrated in Figure 2. Here, the phase space plot of a 1-D system shows that within the dashed region, singularity occurs. That is, the mapping from phase space to physical space produces multiple possibilities. Hence, at some location x, there are multiple velocities v. While this illustration is in 1-D, the theory extends to higher spatial dimensions as well.

In cosmology literature, multistreaming is often mentioned in conjunction with caustics. Caustics are high density structures that form in collisionless media. Just as light caustics are formed by the convergence of light rays from multiple directions, multistreaming can also be formed by the convergence of particles. There are differences between the two however. For example, multistreaming can also arise if particles are traveling at different speeds but along the same direction – this can easily be illustrated by additional folds in the phase space plot. Both are due to singularities in the evolution of the density field, and for the purpose of this paper, both terms are used interchangeably.

Looking at the top portion of Figure 2, one can observe a correspondingly higher density of particles, ρ, in the range of x where there are multiple velocities. This is the basis for existing methods of halo finders that look for over-dense regions associated with multistreaming. In our work, one of the questions we seek to answer is how to use the velocity information to find multistreaming, and see if there is any additional information that can be gleaned.

First, we look at the assumptions that go into numerical simulations. In cosmological simulations, initially the universe is close to homogeneous with very low velocity dispersion. At this early stage, the net effect of the gravitational interaction is insignificant, and particle motion is inertial to a good approximation:

$$x(t,q) = q + t \cdot v(q) \tag{1}$$

The particle position in Eulerian space x is a function of time t and particle position in the (initial) Lagrangian space is q. $v(q)$ is the initial velocity field. Over time, the small density perturbations are amplified by the gravitational instability. In an expanding universe, a non-perturbative description of the instability is given by the Zel'dovich approximation describing particle motion as

$$r(t,q) = a(t)[q - b(t)\nabla\Phi(q)], \tag{2}$$

the particle position in Eulerian space r is a function of time t and particle position in Lagrangian space q. $a(t)$ is the cosmological expansion factor, and $b(t)$ is the growth rate of linear density fluctuations. Since $\Phi(q)$ is the gravitational potential field, $-\nabla\Phi(q)$ describes the initial velocity field, which is conservative and irrotational. In Equation 2, the first term is the unperturbed particle position, and the second term is the spatial perturbation. Equation 1 reduces to Equation 2 through the following substitutions:

$$r(t,q) = x(t,q)a(t), \quad b(t) = t, \quad v(q) = -\nabla\Phi(q)$$

This means that particle motion under gravity mimics particle motion under inertia. Furthermore, the Zel'dovich approximation closely matches the evolution of density perturbation to full N-body simulation until multistreaming.

As particles move under gravity, they form structures such as clusters, filaments, pancakes, or voids. Multistreaming happens in these structures: particles from different position $q_1, q_2, q_3, \ldots, q_n$ in Lagrangian space congregate at point x in Eulerian space having different velocities $v_1, v_2, v_3, \ldots, v_n$ [16]. In other words, a multistreaming region consists of heterogeneous particle flows. Although they were not looking explicitly for multistreaming events, this behavior was also observed by Haroz et al. [5] where they found that dense cluster regions are associated with high velocity uncertainty of particles. In such a scenario, any one of three conditions can lead to multistreaming:

- Particle flows have different speed and direction.
- Particles flows have the same speed but different direction.
- Particles flows have different speeds but the same direction.

From these conditions, one can hypothesize possible metrics for extracting multistreaming events. For example, regions with high velocity variance, or regions with high shear can account for any of the three situations above.

In addition, we found recurring descriptions associated with multistreaming that suggests additional alternatives for finding them. For example, Sahni and Coles [15] described that "... as evolution proceeds, the map connecting initial to final positions develops singularities (caustics) corresponding to multiple flow directions at a given

spatial point. Regions of multistream flow form, and even though each stream is irrotational (curl-free), the velocity field is no longer a potential flow." This suggests that flows where imaginary components in the eigenvalues may now appear. Another example, is the frequent mention of multistreaming happening around the time when the system transition from a linear to a nonlinear regime. If we extrapolate this to linear versus nonlinear flows, one could possibly test for nonlinearity in flows as an indication of multistreaming.

In summary, there are a number of potential identifiers for caustics or multistreaming events e.g. high density regions, exploiting singularities in phase-space, high velocity variance, and high shear regions. Other possibilities include looking for changes in the velocity field — e.g. see if the flow remains curl-free as well as linear. In the next section, we describe and investigate velocity-based indicators of multistreaming events.

4 Data Sets

We use two cosmological simulation data sets in this study — (i) the Zel'dovich pancake where multistreaming regions form flat pancake-like sheets and (ii) a simulation from the Mesh-based Cosmology Code (MC^2) [6]. Both are particle-based simulations of cosmological evolution. Each particle has a unique tag, position, and velocity. More importantly, particles are collisionless. They can occupy the same space without physically colliding into each other.

4.1 The Zel'dovich Pancake

This data set simulates particle behavior within a $(10\sqrt{3}\,\text{Mpc})^3$ box using the MC^2 code. Each megaparsec (Mpc) is approximately 3.26 million light years. The simulation has 64^3 particles and 250 time steps. It models the 1D flow of particles along the main diagonal of the data cube, where it is known that multistreaming starts simultaneously in three orthogonal pancake regions at time step 37. It is an important test data for validating the hypotheses behind our velocity-based feature extraction. We describe our feature extraction methods with results using this data set in Section 5.

4.2 MC^2

This data simulates particle behavior within a $(90\ \text{Mpc})^3$ box using the MC^2 code. It has 256^3 particles and 251 time snapshots. Unlike the Zel'dovich Pancake, the locations and onset of multistreaming events in this data set are unknown. We describe the results of our proposed methods using this data set in Section 6.

4.3 Regridding

To facilitate processing of the particle-based data, we resample the data onto a regular grid. The choice of grid resolution is quite important since if the grid is too coarse we may miss the multistreaming event, and if the grid is too fine it would result in a low particle count and confidence in each cell. We choose the grid size for the density calculation in such a way that each cell on average has eight particles. For the MC^2 data set, this leads to a 128^3 grid, and for the Zel'dovich Pancake data set, a 64^3 grid. At the start of the simulation, each cell contains eight particles on average. As time progresses, some regions become more dense while others become sparse or even empty. Empty cells as well as those in their immediate vicinity must be treated with care so that they do not produce erroneous results in the analysis.

5 Identifying and Characterizing Multistreaming Events

Incorporating velocity from particle simulations opens up several ideas for finding multistreaming events. In this section, we describe how those ideas can be formulated into feature detectors. We also test each method against the pancake data to see how well it detects the onset and identify the multistreaming regions. We use ParaView [17] to visualize our results.

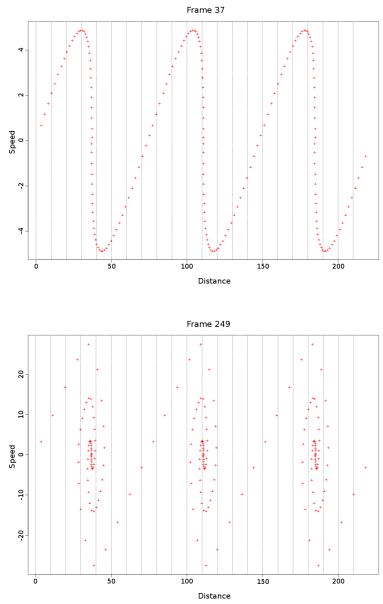

Fig. 3. Phase space plot. Top image is from frame 37 of the simulation, which coincides with the onset of multistreaming. Note the three vertical sections of the curve where multiple velocities are associated with each position. Bottom image is from frame 249, the last frame of the simulation. One can see that the curve has evolved into multiple, but non-self-intersecting curves. The range within each of the three multistreaming regions have also grown wider.

5.1 Phase Space Plots

First, we verify the onset of multistreaming events in the Zel'dovich pancake by looking at the phase space plots (PSP) of the particles at each time step. Because the flow in the Zel'dovich pancake data is along the main diagonal d of the data cube, we find the PSP by defining a plane going through the origin o and perpendicular to d. Given a particle with position p and velocity v, we calculate its distance p' and velocity v' along d with respect to p.

$$\begin{aligned} p' &= |p-o|\cos\alpha \\ v' &= |\overrightarrow{v}|\cos\beta \end{aligned}$$

where α is the angle between $p-o$ and d, and β is the angle between v and d. The PSP shows the relationship between p' and v' of all the particles in a time frame. Figure 3 confirms that at the onset of multistreaming, frame 37, the PSP is not a functional anymore giving rise to multiple velocities at a given position. The same information displayed in physical space as a volume rendering of the particle density field is shown in Figure 4. Here, we can clearly see the three pancake regions in green which intensifies and gets thicker over time. The three sections of the PSP in Figure 3 correspond to the three pancake regions in Figure 4. The thicker pancakes correspond to the expanding region of overlapping folds in the PSP.

Unfortunately, the PSP plot is useful only if the velocities are predominantly 1D. In general 3D flows, the PSP combines velocities that are not orthogonal to the reference plane, and hence cannot be used directly in the analysis.

5.2 Maximum Shear Stress

Shear in the velocity field can be one of the mechanisms for multistreaming. Particles going in the same or opposite directions but at different speeds lead to shear in the velocity field. To find the maximum shear stress, we first calculate the velocity gradient tensor of the velocity field, then find its symmetric tensor and associated eigenvalues λ_1, λ_2, and λ_3. We use the von Mises criterion for maximum shear

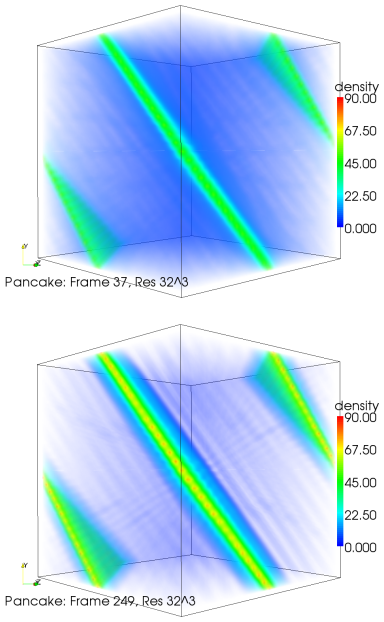

Fig. 4. Volume rendering of the particle density field. Top image is from frame 37. The three green pancake regions correspond to the vertical sections of the PSP in Figure 3. Bottom image is from frame 249, and shows how the pancakes have intensified and thickened.

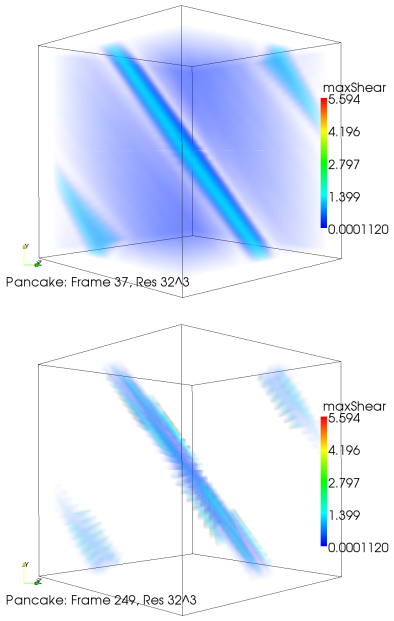

Fig. 5. Maximum shear stress. Top image is from frame 37 where we can see well defined three regions of high shear. Over time, the maximal shear has increased and become more pronounced as shown in frame 249 on the bottom. On the other hand, the size of the high shear region has tapered off and become more concentrated in the pancake region.

Red	Repelling-Spiral-Saddle
Green	Repelling-Node-Saddle
Blue	Repelling-Spiral
Yellow	Repelling-Node
Purple	Attracting-Spiral-Saddle
Cyan	Attracting-Node-Saddle
White	Attracting-Spiral
Gray	Attracting-Node

stress which is defined as:

$$MS = \sqrt{\frac{(\lambda_1-\lambda_2)^2+(\lambda_1-\lambda_3)^2+(\lambda_2-\lambda_3)^2}{2}} \quad (3)$$

Figure 5 shows the results of this metric.

5.3 Critical Points

Because flows in the vicinity of critical points usually involve different velocities (directions and magnitudes), we do a quick study on the location of critical points. In addition, because flows in multistreaming regions are not potential flows anymore, we also look at the type of critical points. In particular, we try to identify those that have rotational components. Figure 6 shows the location and color code the different types of critical points according to the following scheme:

5.4 Divergence

Divergence is a scalar quantity that measures the degree to which a vector field is a source or a sink at a given location. Positive values indicate a source-like behavior, while negative values indicate a sink-like behavior. Divergence can potentially be used to find multistreaming because it finds regions where light converges (sink) as in caustics. Looking at the evolution of the divergence field in Figure 7, one can see that the pancake regions, corresponding to negative divergence, are starting to pull in particles most noticeably when multistreaming started. At the same time, the particles in-between the pancake regions, with positive divergence, are being forced away and towards

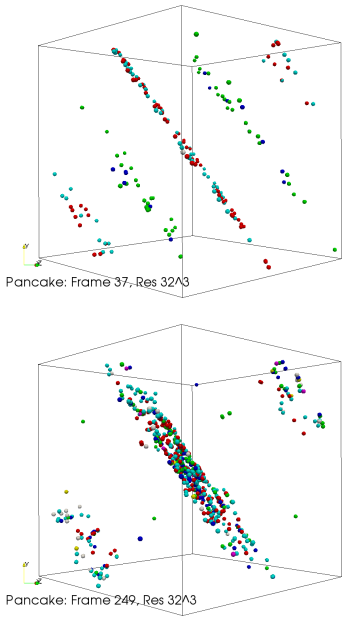

Fig. 6. Type and locations of critical points at time frames 37 and 249. One can observe that the number of critical points grow over time and congregate at the pancake regions. One can also observe that critical points with rotational components (colored red, white, blue, and purple) can also be found in these regions. On the top image, one can also see two planes of predominantly green and blue critical points, both of which are repelling types. We surmise that these critical points "push" the particles in their towards the three pancake regions towards the end of the simulation. On the other hand, the high concentration of different types of critical points in the pancake regions indicate the highly complex behavior even for this simple 1D model.

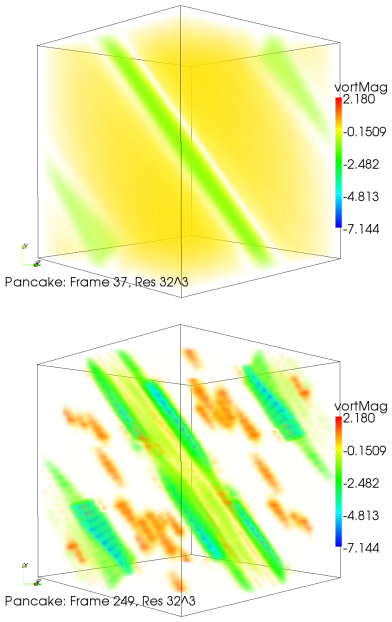

Fig. 7. Divergence field at time frames 37 and 249. The negative divergence in the three green pancake regions tend to draw in the particles from their surrounding neighborhood. These regions seem to intensify and grow larger over time.

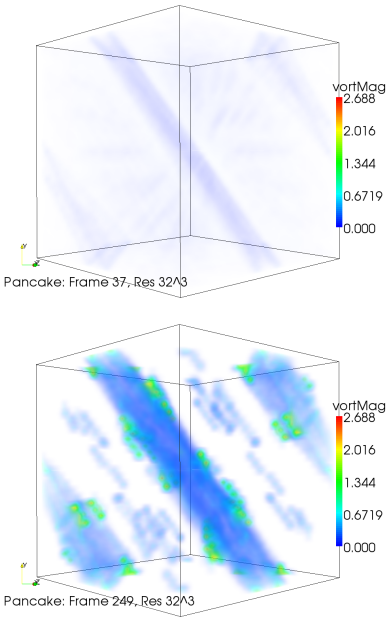

Fig. 8. Vorticity magnitude can be seen as faint bluish regions in the three pancake regions in frame 37, that becomes more pronounced in frame 249. We also note that vorticity around the boundaries of the pancake regions seem to be stronger.

the pancake regions. This corroborates the observations using critical points.

5.5 Vorticity

Vorticity measures the tendency of vector field elements to spin. In cosmological simulations the velocity field is irrotational (zero curl) prior to multistreaming. We hypothesize that vorticity may be used as an indicator for multistreaming.

The vorticity at a point is a vector and is defined as the curl of the velocity. That is, vorticity is $\nabla \times \vec{V}$ where $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$. Since we are primarily interested in detecting the presence of regions with rotational motions and not their particular orientations, we look at the vorticity magnitudes in the simulations. Figure 8 shows that vorticity magnitude increasing over time with higher vorticity in the pancake regions.

5.6 Linearity Test

Another test for multistreaming is to check if the velocity field is still linear. This is motivated by the description that the simulations start out being linear, then transition through a quasi-linear to a nonlinear behavior. Detecting changes in the linearity of the velocity field may be an indicator of multistreaming.

Given a velocity field $\vec{V}$, position p, and velocity gradient J, we can obtain the velocity of a nearby point dp using first order approximations if the field is linear.

$$\vec{V}(p+dp) = \vec{V}(p) + J \cdot dp \tag{4}$$

where dp is one of $[\pm 1,0,0]$, $[0,\pm 1,0]$, or $[0,0,\pm 1]$ depending on which neighboring velocity we want to get. To check whether the velocities around p are linear, we compare the first order approximation of $\vec{V}(p+dp)$ against the original velocity at $p+dp$ by taking their pairwise dot products. Velocities are normalized before dot products are taken. Each direction that has a dot product greater than 0.9 counts as a vote for linearity. If any neighbor of p is empty, we skip the calculation of $J(p)$ and do not apply the linearity test at p. Figure 9 shows results of running the linearity test on the pancake data. Cells with vote counts of 3 or more are considered linear. Regions with linear flows are not of interest and are made transparent. Regions with nonlinear flows are rendered opaque and bluish. Their opacity drops off as they tend towards linear flows. While we can see that the nonlinear regions correspond to the pancake regions found using other methods at the last time frame, the linearity test did not show the pancake region at frame 37. In fact, the test did not detect any nonlinear flows until 14 frames after frame 37. This might be due to the strict cutoff of having not more than 3 votes, or the test, using first order approximations, was simply late in predicting multistreaming.

5.7 General Observations

We know that multistreaming happens at time frame 37 for the pancake data. This is confirmed in the PSP images. There is strong agreement with the three pancake regions among the density field, the critical points, maximal shear, divergence, vorticity, and nonlinearity tests. In addition to consistently identifying the multistreaming regions, the velocity-based methods were also able to provide valuable new information not available with density-based methods alone. This can be observed in the form of complex behaviors within multistreaming regions as well as the behavior of nearby regions that contribute to the formation of multistreaming regions. One drawback to the current set of methods, however, is that they are not able to accurately predict the onset of multistreaming. The main limiting factor is that one needs to set the appropriate threshold parameter for each test e.g. what is the minimum value of maximum shear before one can consider a region to be multistreaming, or what is the maximum number of affirmative votes for linearity before a cell is considered to have nonlinear flow? Even for the existing density-based methods of halo finding, one also needs to set a threshold for particle density [7].

In our experiments, we looked at the histograms of the relevant derived fields and selected threshold values that made physical sense and produced results corresponding to the expected results for the pancake data set. There are several problems with having to manually find and set thresholds: (i) the thresholds may vary from one data set to another, (ii) even worse, the threshold may vary one time frame to another. Absent an automated method for finding the proper threshold value, we

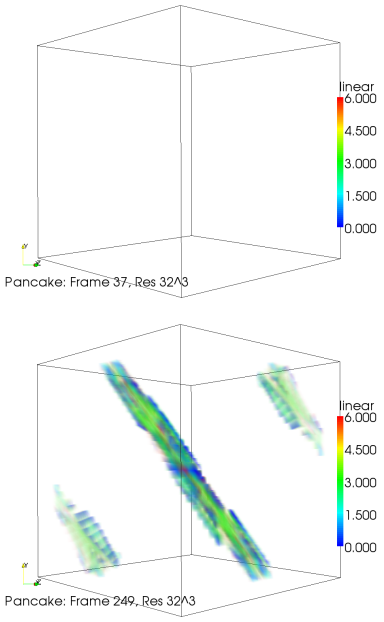

Fig. 9. Nonlinear regions in frames 37 and 249. Linear regions are reddish but has opacity of zero since we are interested in the nonlinear regions. Nonlinear regions are bluish and more opaque. In frame 37, this test does not detect any nonlinear flows. In frame 249, nonlinearity is detected in the three pancake regions.

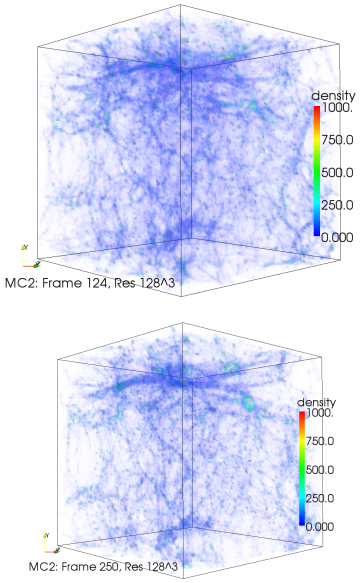

Fig. 10. Particle density field of the MC^2 data set for frame 124 on the top and 250 on the bottom. Clusters and filaments can already be observed in the top image. These complex structures continue to evolve with some structures disappearing while others becoming more prominent towards the end of the simulation.

are unable to predict the actual onset of multistreaming. However, we see this as avenues of new research within the cosmology field to identify relevant thresholds, just as they have determined the appropriate density threshold for halo finders.

6 Blind Test

While the different velocity-based methods can identify the multistreaming regions, none of the methods, in their current form, can accurately predict the onset of multistreaming. Therefore, in the next set of tests with the MC^2 data set, we are primarily interested in the evolution and structure of the multistreaming regions. As with the pancake data set, the MC^2 data set starts out as fairly homogeneous at the start of the simulation. Therefore, we are showing only the middle (frame 124) and last frames (250) of the simulation.

As a point of reference, we first show the results from density-based halo finder in Figure 10. We note that the complex structure is already quite prominent at the halfway mark in the simulation and continue to change over time.

Figure 11 shows the results of maximum shear stress on the MC^2 data set. The top image shows high correlation with the density method in Figure 10. The bottom image shows that regions of high shear are diminishing which may suggest that multistreaming due to the shearing mechanism may be weakening even though the density image shows very high concentration. That is, while the particles have aggregated to form LSS, there is actually less shearing action within the aggregate. If so, this could be interpreted as leading to a process where there is more structural stability and less change in the system. Of course, this kind of observation need to be verified with additional research and analysis by the scientists.

The interpretation of the critical points in this data set is less conclusive. Figure 12 shows that the number of critical points are decreasing over time. On the other hand, the locations of the critical points roughly correspond to the locations of the LSS. Also, we can still see red, white, and blue critical points (purple is harder to see) in those regions. So, the local flow behavior in the vicinity of those critical points still exhibit swirling motion. Perhaps, a plausible explanation

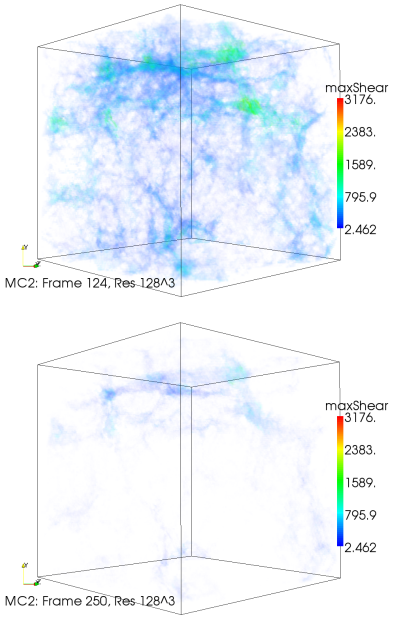

Fig. 11. Maximum shear stress of MC^2 data set at frames 124 and 250. The LSS on the top correlates very well with those in Figure 10. The bottom image shows that regions of high shear have diminished. A possible explanation is that even though particle densities are high in the LSS, the structures are becoming more stable. This also suggests that regions of high shear stress may be better characterized as places where one can have a higher chance of finding multistreaming activity. In contrast, traditional density-based methods find the LSS but does not say much about their dynamics.

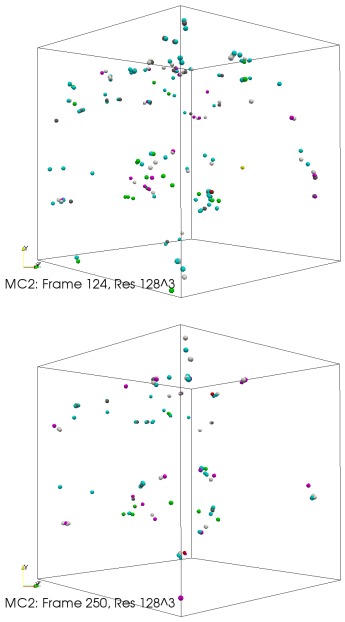

Fig. 12. Critical points in time frame 124 and 250 of the MC^2 data set. Refer to Table 5.3 for the color legend of different types of critical points. The number of critical points diminish over time suggesting that the flow patterns are becoming less complex.

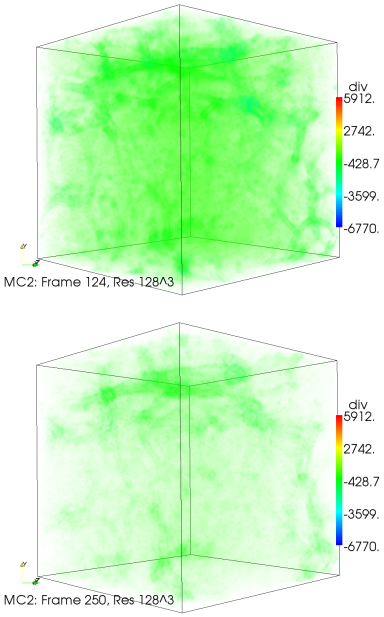

Fig. 13. Divergence field at frame 124 and 250 of the MC^2 data set. This pair of images show that competition for particles from neighboring regions have subsided over time and are concentrated at the already high density regions at the end of the simulation.

might be that over time, this particular simulation produced a system that is becoming more stable over time. Such that at the end of the simulation, there is actually less critical points, and a correspondingly less complex behavior of particle motions. This explanation coincides with the observation using the maximum shear stress criterion.

Figure 13 shows the divergence field. Recall that regions with negative divergence are places that have a sink-like behavior and tend to draw particles in from neighboring regions. We can observe that over time, the divergent region becomes concentrated and co-located with regions of high density. Therefore, competition for pulling particles from neighboring areas are now localized to these high density regions. This suggests that further drastic movements away from the established high density LSS will be less likely, and the structure at the end of the simulation is becoming more stable.

Two frames of the vorticity magnitude of this data set are shown in Figure 14. Here, we also notice that the structure of the vorticity magnitude field seem to subside over time, and that the similarity with the structure from the density field is easily identifiable. Compared to the maximum shear stress, the LSS are better preserved. This suggests that local spins and rotations remain the predominant mechanism relative to shearing at the end of the simulation. Also, one can notice that the high vorticity regions seem to be more pronounced in the cluster region than in the filaments.

Results using the nonlinearity test are shown in Figure 15. While we can see that there is also good correspondence between the nonlinear regions and the high density regions, it is difficult to draw conclusions with this criterion. The reason is that nonlinear regions were detected very early in the simulation. It is possible that we are detecting false positives due to the first order approximation test for nonlinearity, and therefore this method needs to be refined further.

Overall, the suite of methods to find the multistreaming regions produce results that agree very well with those found by density-based methods. The main advantage, however, is that they provide additional information about the behavior within the multistreaming regions. Some of the observations made as a result of these methods have raised some questions that require additional research in cosmology.

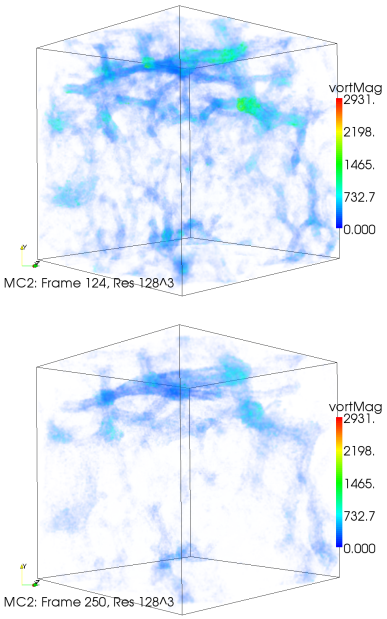

Fig. 14. Vorticity magnitude of the MC^2 data set for frames 124 and 250. High vorticity regions can be observed to be located in high density regions. These regions seem to subside over time and concentrate around clusters more so than in the filament regions.

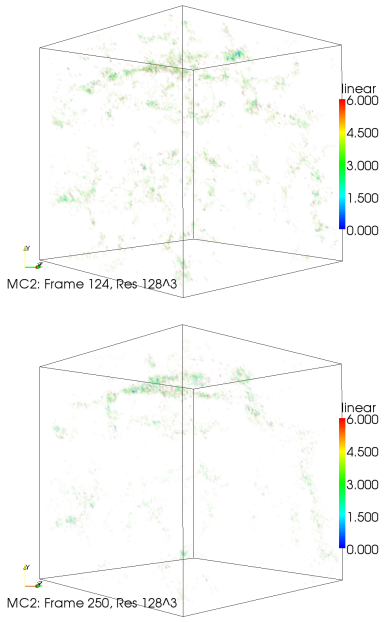

Fig. 15. Nonlinear regions in frames 124 and 250 of the MC^2 data set. The locations of the nonlinear regions correspond to the high density regions in Figure 10.

7 Conclusions and Future Work

We started this investigation with a general question of whether we can use the velocity information in the simulation data to detect and characterize multistreaming events. We hypothesized how the flow field should behave given the various descriptions of multistreaming in the cosmology literature and formulated ways to extract regions with those behaviors. Not all of our hypotheses were verified. Compared to density-based techniques, the velocity-based techniques produce qualitatively very similar results. More importantly, they not only identified the multistreaming regions but also provided information about the particle flow behavior within them and their vicinity. On the other hand, they were not able to accurately detect the onset of multistreaming. It is possible though that this inadequacy is due to other variables, e.g. gridding strategy, and not the formulations themselves.

From our investigations, particularly through animations, we have learned that (i) once multistreaming starts, it is not necessarily monotonous but can grow and shrink over time, (ii) multistreaming regions are not static and can move around, and (iii) there are very interesting dynamic behaviors within multistreaming regions.

We have met the original goal of this work and can affirmatively conclude that velocity-based methods are viable tools for finding multistreaming regions. The different methods not only find the multistreaming regions, but also provide additional information about the regions and the behavior within the regions. While one may wonder which is the best method, we are not ready to provide an answer yet as the methods have different strengths and weaknesses with respect to the different properties of multistreaming.

The results we have to date are very promising, but they are also far from being definitive. During the course of this work, we have identified a few areas for improvement and new research questions in visualization and astronomy alike. For example, the main weakness of the current set of techniques is the reliance on a user-specified threshold. There are at least two possible ways to address this: (i) find an automated way for determining threshold levels based on data distribution e.g. we have looked at the histogram but have not found a generalizable method so far; and (ii) correlation between simulations and empirical data sets to arrive at the threshold e.g. what has been done for the density-based method by the astronomers. The current implementation relied on very simplistic regular gridding. The effects of different gridding strategy (such as adaptive meshing, application of sparse data interpolation, and cloud-in-cell algorithms, etc.) on the resulting multistreaming regions need to be studied further. The PSP was the benchmark for 1D flows, but unfortunately not directly applicable to 3D flows. Future research can look at extending PSP to support multi-directional flows. While we do not have the theory developed yet, a close analogy would be the projection slice theorem commonly found in 3D reconstruction of 3D scalar fields. Another possible area of investigation is the application of multivalued feature extraction techniques [11] to this problem. Cosmology simulations tend to be very large. Hence, scalability of algorithms is also another area of concern and further research.

With respect to cosmology, our work have identified some possible directions of research. For example, what types of behavior can one find in a multistreaming region? How do these behaviors evolve over time – does one type of behavior become more dominant? Or are certain types of behavior associated to different types of large scale structures e.g. filaments versus clusters? Can filaments (or cluster) happen in isolation, or do they have to be topologically connected? Is there a density (or velocity) threshold below which multistreaming cannot occur?

References

[1] J. Ahrens, K. Heitmann, S. Habib, L. Ankeny, P. McCormick, J. Inman, R. Armstrong, and K.-L. Ma. Quantitative and comparative visualization applied to cosmological simulations. *Journal of Physics*, Conference Series 46, SciDAC:526–534, 2006.

[2] P. Coles. Gravitational instability and the formation of the supercluster-void network in the universe. *Chaos, Solitons and Fractals*, 16:513–525, 2003.

[3] U. Frisch and R. Triay. Caustics and cosmological structures, 2005. `http://www.cpt.univ-mrs.fr/~cosmo/NLCP_2005/NLCP2005_Scope.html`.

[4] N. Gouda. Morphology in cosmological gravitational clustering and catastrophe theory. *Progress of Theoretical Physics*, 99(1):55–68, 1998.

[5] S. Haroz, K.-L. Ma, and K. Heitmann. Multiple uncertainties in time-variant cosmological particle data. In *IEEE Pacific Visualization Symposium*, pages 207–214, 2008.

[6] K. Heitmann, P. Ricker, M. Warren, and S. Habib. Robustness of cosmological simulations I: Large scale structure. *Astrophysics Journal Supplement*, 160(28), 2005.

[7] S. R. Knollmann and A. Knebe. Ahf: Amiga's halo finder. *The Astrophysical Journal Supplement Series*, 182(2):608–624, 2009.

[8] S. Levy. Partiview@ncsa. http://dart.ncsa.uiuc.edu/partiview/.

[9] H. Li, C.-W. Fu, Y. Li, and A. Hanson. Visualizing large-scale uncertainty in astrophysical data. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1640–1647, 2007.

[10] Z. Lukic, D. Reed, S. Habib, , and K. Heitmann. The structure of halos: Implications for group and cluster cosmology. *The Astrophysical Journal*, 692(1):217–228, 2009.

[11] A. Luo, D. Kao, and A. Pang. Visualizing spatial distribution data sets. In *VisSym'03*, pages 29–38, 238, May 2003.

[12] R. Mohayaee, S. Colombi, B. Fort, R. Gavazzi, S. Shandarin, and J. Touma. Caustics in dark matter haloes. *EAS Publications*, Series 20:19–24, 2006.

[13] P. Navratil, J. Johnson, and V. Bromm. Visualization of cosmological particle-based datasets. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1712–1718, 2007.

[14] P. Peebles. The standard cosmological model, 1998. http://nedwww.ipac.caltech.edu/level5/Peebles1/paper.pdf.

[15] V. Sahni and P. Coles. Approximation methods for non-linear gravitational clustering. Phys. Rep., 262:1–135, Nov. 1995.

[16] S. F. Shandarin and Y. B. Zeldovich. The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium. *Rev. Modern Physics*, 61(2):185–220, 1989.

[17] A. H. Squillacote. *The ParaView Guide: A Parallel Visualization Application*. Kitware, Inc., 2007.

[18] T. Yano, H. Koyama, T. Buchert, and N. Gouda. Universality in the distribution of caustics in the expanding universe. *The Astrophysical Journal Supplement Series*, 151:185–192, 2004.