

# **Please Read the Abstract**

By

**Nesar Ramachandra**

Submitted to the Department of People who read Abstracts and the  
Graduate Faculty of the University of Kansas  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy

---

Sergei Shandarin, Chairperson

---

Hume Feldman

Committee members

---

Barbary Anthony-Twarog

---

Mikhail Medvedev

---

Suzanne Shontz

Date defended: October 02, 2012

The Dissertation Committee for Nesar Ramachandra certifies  
that this is the approved version of the following dissertation :

Please Read the Abstract

---

Sergei Shandarin, Chairperson

Date approved: \_\_\_\_\_ October 03, 2012 \_\_\_\_\_

## Abstract

Version 2.1, May 20, 2013. This was a minor update that slightly changed kuthesis.cls. First, the Graduate School no longer exists, so we rephrase the front matter to refer to Graduate Faculty. Second, the page numbering after the front matter begins at arabic number 1. This was an oversight in the previous edition, it was never intended, but users found easy “work arounds” and we were not aware of it.

Version 2, October 2, 2012. This is a minor update to allow insertion of dates in the document preamble.

If you are reading the pdf, this document is produced from “thesis-ku.tex” using the  $\text{\LaTeX}$  class kuthesis.cls and other common packages that are almost surely installed in a working  $\text{\LaTeX}$  system. We suggest you try to compile this document in your system (as a test), and after that use it as a template for your project.

**Look here.** The  $\text{\LaTeX}$  class “kuthesis.cls” extends the base class called “report”. If you have a report document already, migration into this format will be easy, since we inherit everything from “report” and then customize a few things.

**How To Use?** This is a basic  $\text{\LaTeX}$  usage question. You need a “ $\text{\LaTeX}$  distribution” such as TexLive or MikTex. You need to run the file thesis-ku.tex through the  $\text{\LaTeX}$  processing system. At the current time, the preferred output format is pdf from the “pdflatex” program, but it is probably safe to try other programs as well. On most  $\text{\LaTeX}$  systems, your latex processor will look in several places for “kuthesis.cls” while it processes thesis-ku.tex. Usually, you don’t need to install formally that file, since—on most systems—the  $\text{\LaTeX}$  program will look in the current working directory for

style files. If so, then this document will “just work” for you. If you want to use kuthesis.cls for several documents, you need to either copy it into every document directory, or install “kuthesis.cls” in a place where your L<sup>A</sup>T<sub>E</sub>X system will find it, and then run the L<sup>A</sup>T<sub>E</sub>X re-configuration program to make the system notice. In my system, L<sup>A</sup>T<sub>E</sub>X is set up to look in /usr/share/texmf/tex/latex, /usr/local/share/texmf/tex/latex, as well as in my home directory and the current working directory. I copy kuthesis.sty into /usr/local/share/texmf/tex/latex, run the re-configuration program “texhash,” and it is ready to go.

**Does it work with Lyx?** Yes, as a matter of fact, I’m preparing this document in LyX right now. (LyX is a GUI L<sup>A</sup>T<sub>E</sub>X document editor, available for free at <http://www.lyx.org>.) LyX will open the document, but it will warn that the layout file “kuthesis.layout” is unavailable. Unfortunately, the only fix for that is to install the customized layout file, which is provided with this document in the style-files directory. Copy the kuthesis.layout file into the one of the folders where LyX looks for layout configurations. I did not install that in the system-wide LyX layout folder, but rather in my personal user configuration folder. I suggest you do the same, at least while you are testing this. On Linux, that is in the user HOME under .lyx/layouts. (For me, that’s “/home/pauljohn/.lyx/layouts.”) On other OS, here’s how you can find out where to put the layout. Run LyX, in the top menu choose Tools/Preferences. Change something trivial, close. That will force LyX to create a user folder somewhere in your personal HOME directories. In MS Windows 7, I just tested this. I searched for the word “layouts” and found my LyX preference folder is “C:\Users\pauljohn\AppData\Roaming\lyx202\layouts”.

**Note.** To customize this document, you have to edit the document preamble to put in your name, title, degree name, department, and advisors. If you don’t do that, you won’t get the right title page. To be explicit, in this document, the following values are currently set in the preamble:

```

\title{Please Read the Abstract}
\author{Abstract Writer}
\dept{Department of People who read Abstracts}
\dereetitle{Doctor of Philosophy}
\papertype{Dissertation} %capitalization is important here
\committee{MEMBER 1}{MEMBER 2}{MEMBER 3}{MEMBER 4}{}

```

**Note please.** Replace “Member 1” with a name, of course. You can have at most 5 committee members, but you can have fewer. For the ones you have, put their names inside squiggles; for the ones you don’t want, leave empty squiggles, {}, as I’ve demonstrated on the 5th member in this example.

**Note also.** The chapters included here are just for illustration and testing. I’ve fiddled the format to look like what KU asks for.

**Note also.** Note the file “thesis-ku.tex” (or thesis-ku.lyx, as the case may be), does not have much content. It has the title page, the abstract, and not much else. Everything else is brought in from separate files. I strongly recommend you keep separate chapters in separate folders. If you don’t do that, you will likely go insane.

**Note for the version 2012-10-02.** In the first edition of this setup, I left the dates blank and expected users to write or type them in as necessary. A user requested a fix to allow the dates of the defense and final approval to be added in the document preamble with the other settings. So in the end of the document preamble, insert the following. The first is some L<sup>A</sup>T<sub>E</sub>X code, don’t change it.

```

\@printd@testtrue
\datedefended{October 2, 2012}
\dateapproved{October 3, 2012}

```

**Note finally.** I want you to use L<sup>A</sup>T<sub>E</sub>X to prepare documents. Doing anything else is “finger painting.” But don’t ask me for much help. If you find something wrong with kuthesis.cls, I will be glad to try to fix it if you let me know. I don’t have time/ability to answer a million email questions about L<sup>A</sup>T<sub>E</sub>X or LyX. So please consult the WWW, there is a ton of stuff up there, that’s how I learned. And if you get in a serious hole and ask me for help, it might cost you a t-shirt.

Paul Johnson <pauljohn@ku.edu>

Prof. Political Science & Assoc. Dir., Center for Research Methods and Data Analysis

2012-10-02

And that’s the end of the abstract.

## **Acknowledgements**

I would like to thank mum, dad, Prof. Shandarin, committee members, officemates, friends, exes, uncles, aunts, grandpa, grandma, grandma's dogs, neighbors, neighbors' dogs who made this thesis possible.

# Contents

<b>1 General Introduction</b>	<b>1</b>
1.1 Large Scale structure . . . . .	1
1.2 Observation? . . . . .	1
1.3 Cosmic fields: Characterizing non-linear growth of Cosmic structure . . . . .	1
1.4 Chapter Organization . . . . .	3
<b>2 Background</b>	<b>4</b>
2.1 Zel'dovich Approximation . . . . .	4
2.2 Dynamics: Lagrangian sub-manifold . . . . .	7
<b>3 Multi-stream portrait of the Cosmic web</b>	<b>8</b>
3.1 Introduction . . . . .	8
3.2 The simulation . . . . .	12
3.3 Multi-stream field calculation . . . . .	12
3.4 Global statistics of Cosmic web . . . . .	13
3.4.1 Volume and mass fractions . . . . .	16
3.5 Stream environment around haloes . . . . .	21
3.5.1 Technique . . . . .	22
3.5.2 Voids, filaments and walls around haloes . . . . .	25
3.6 Summary . . . . .	34

<b>4 Topology and geometry of the dark matter web: a multistream view</b>	<b>39</b>
4.1 Introduction . . . . .	39
4.2 The simulation . . . . .	46
4.2.1 Multi-stream field at $z = 0$ . . . . .	46
4.3 Voids in the multistream field . . . . .	48
4.3.1 Connectivity of the voids . . . . .	50
4.3.2 Halo boundaries within the void . . . . .	51
4.4 Percolation in the multistream web . . . . .	53
4.5 Local geometry of the multistream field . . . . .	56
4.5.1 Softening of the multistream field . . . . .	61
4.5.2 Resolution dependence . . . . .	62
4.6 Discussion . . . . .	66
4.7 Summary . . . . .	70
<b>5 Dark matter haloes: a multistream view</b>	<b>73</b>
5.1 Introduction . . . . .	73
5.2 Phase-space representation of gravitational clustering . . . . .	78
5.2.1 Collapse in higher dimensions . . . . .	80
5.3 Simulations and tools . . . . .	82
5.4 Haloes in the multistream field . . . . .	83
5.4.1 Hessian of multistream fields . . . . .	84
5.4.2 Halo finder algorithm . . . . .	85
5.4.3 Effect of smoothing . . . . .	89
5.4.4 Effect of multistream thresholds . . . . .	90
5.5 Halo properties . . . . .	92
5.6 Correspondence with other halo finders . . . . .	96
5.6.1 Haloes in the percolating filament . . . . .	102
5.7 Discussion . . . . .	105

5.8	Summary	108
<b>A</b>	<b>My Appendix, Next to my Spleen</b>	<b>127</b>
<b>B</b>	<b>The multistream flow field in one-dimension</b>	<b>128</b>
B.1	Variations in the multistream field	130
<b>C</b>	<b>Hessian signatures of the multistream field</b>	<b>133</b>

# List of Figures

1.1	Classification of some of the fields/parameters used in cosmological analyses. Some fields utilise position co-ordinates only whereas others use the full phase-space information. In addition, the fields may be defined on a regular grid, or may be defined on an unstructured grid (For instance, flip-flop is a number-valued field defined on each dark matter particle). Fields like mass density and multi-streams can be defined on either grids, depending on the numerical technique. List is obviously not exhaustive- velocity and potential fields are not included, and discussions of correlation functions are excluded as well. . . . .	2
2.1	This plot shows the idea of Lagrangian tessellations. Left: The distribution of Dark matter particles in Lagrangian space is on the regular grid. The tetrahedra surfaces in this case are mostly regular (except at the edges). Right panel shows the Eulerian positions of the same particles at $z = 0$ . The particles clearly have undergone multiple flip-flops, and the intersections of Lagrangian tetrahedra signifies locations of caustic surfaces. Specific tessellation schemes can be utilized separately to identify these surfaces. . . . .	6
3.1	The cosmic web structure in a slice of $30h^{-1}\text{Mpc} \times 100h^{-1}\text{Mpc} \times 100h^{-1}\text{Mpc}$ in a simulation box of size $100h^{-1}\text{ Mpc}$ and $128^3$ particles. Regions with 17 or more streams (blue) form a filamentary structure. The haloes determined by FOF (red) are predominantly embedded in the intersections of the filaments. . . . .	15

3.2	Volume and mass fraction of each stream, mean density of each stream in a box of size $100 h^{-1}$ Mpc and $128^3$ particles. Exact values of fractions, density and their curve fit for the range $n_{str} \geq 5$ from Eq. 3.1, Eq. 3.2 and Eq. 3.3 are shown. Multi-streams are calculated with refinement factor of 8. The void ( $n_{str} = 1$ ) occupies 93% of the volume and 55% of the mass. . . . .	17
3.3	Top: Volume fraction distribution of streams in 6 simulation boxes of size $100 h^{-1}$ Mpc, $200 h^{-1}$ Mpc and $128^3, 256^3, 512^3$ grids (with refinement factor of 1). Volume fractions are similar for simulation boxes with same inter-particle resolution. Slopes of the curve fits are shown in Table 3.1. Bottom: Volume fraction distribution for different refinement factors for $100 h^{-1}$ Mpc, $128^3$ box. A considerably smoother volume fraction distribution is obtained at high number of streams in multi-stream fields with high refinement factor. . . . .	19
3.4	Top: Mass fraction distribution of streams in 6 simulation boxes of size $100 h^{-1}$ Mpc, $200 h^{-1}$ Mpc and $128^3, 256^3, 512^3$ grids ( with refinement factor of 1). Mass fractions are similar for simulation boxes with same inter-particle resolution. Slopes of the curve fits are shown in Table 3.1. Bottom: Mass fraction distribution for different refinement factors for $100 h^{-1}$ Mpc, $128^3$ box. Single-streaming void regions have more fraction of mass particles in multi-stream fields calculated at low refinement factors of 1 and 2. This effect is minimized when calculation is done at better refinement. In addition, mass fraction distribution is less noisy. . . .	20

3.5 Multi-stream flow regions in a small box of the simulation. Top left: regions with more than 3-stream flow are identified as walls (brown). Intersection of multiple walls have higher $n_{str}$ regions (green, 17+ streams). Single-streaming voids (white) occupy large volume and are very close to the filaments in some directions. Top right: 17+ streams (green) form filamentary structures with nodes at the intersections (red, 90+ streams) Bottom left and right: Closer look at the highly non-linear region reveals that a filament is sandwiched between the walls (brown). The 90+ stream region (red) forms a compact structure and is entirely contained within the filament. The black dots show the particles around the FOF halo within linking length of 0.2.	23
3.6 Modelling a wall and a filament. A diagnostic spherical surface is intersected by a cylinder and plane.	24
3.7 Large halo of mass $3.7 \times 10^{14} M_\odot$ and FOF radius $1.2 h^{-1}$ Mpc (dotted-violet line in the top figure). Top: Fractional distribution of streams on the surface of spheres of increasing radii. Dashed-black lines are for $r^{-1}$ and $r^{-2}$ scaling. 3+ streams are closest to $r^{-1}$ and 17+ scales close to $r^{-2}$ . for higher thresholds, the fractional distribution departs smoothly from $r^{-2}$ . Bottom: Mollweide projection of multi-streams on the surface of the sphere.	27
3.8 Halo of mass $5.0 \times 10^{13} M_\odot$ and FOF radius $0.7 h^{-1}$ Mpc. Top: Fractional distribution of streams deviates from $r^{-1}$ and $r^{-2}$ scales since the high stream filament passes along the surface of the sphere. Bottom: Filament passing through the surface is seen from 2 to 5 times the halo radius.	28
3.9 Halo of mass $7.0 \times 10^{12} M_\odot$ and FOF radius $0.4 h^{-1}$ Mpc. Top: All lines clearly scale between $r^{-1}$ and $r^{-2}$ . Bottom: The filament is passing through the centre. It persists from radius of halo to $4 h^{-1}$ Mpc. It is also surrounded by a single wall appearing as a line in the middle.	29

3.10 Halo of mass $1.1 \times 10^{12} M_{\odot}$ and FOF radius $0.2 h^{-1}$ Mpc (dotted-violet line). This halo just has 26 particles, hence the resolution is lesser than previous haloes for surfaces with low radii. Top: There is a bump in fraction of each of 3+ streams a little over $1 h^{-1}$ Mpc. This is due to the presence of an additional halo nearby, as seen in the projections. The Mollweide projections from $0.4 h^{-1}$ Mpc to $2 h^{-1}$ Mpc have high stream flow regions near the lower surfaces of diagnostic spheres. Bottom: Corresponding FOF halo (red, at the centre) has a more massive neighbouring FOF halo (blue) within distance of $2 h^{-1}$ Mpc. The 17+ stream regions (green) are increased around the neighbouring halo. . . . .	30
3.11 Scatter plot of minimum and maximum number of streams on the surface of FOF radius around all haloes in FOF catalogue. For the analysis, we use a total of 5521 haloes that are identified using the FOF technique with linking length, $b = 0.2$ . Since several of the haloes coincide, distributions of number of haloes corresponding to minimum and maximum $n_{str}$ on their FOF radii are shown above and beside the scatter plot respectively. Top: Full box of $L = 100 h^{-1}$ Mpc and $N = 128^3$ (i.e. $L/N = 0.78 h^{-1}$ ), with a low refinement factor of 1 is utilized for multi-stream field calculation. Bottom: Same simulation box; multi-stream field calculated with higher refinement factor of 8. . . . .	32
3.12 Scatter plot of minimum and maximum number of streams on the surface of FOF radius around all haloes in FOF catalogue. Distributions of number of haloes corresponding to minimum and maximum $n_{str}$ on their FOF radii are shown above and beside the scatter plot respectively. Top: In the simulation box of $L = 100 h^{-1}$ Mpc and $N = 512^3$ (i.e. $L/N = 0.19 h^{-1}$ ), multi-stream field is calculated on a smaller slice of $25 h^{-1}$ Mpc with a refinement factor of 1. Multi-stream field is projected onto surfaces of 3448 FOF haloes within this small box. Bottom: Same simulation box; the multi-stream field calculated on the same small box, but with a refinement factor of 8. . . . .	33

4.1	3D rendering of the multistream field: the cosmic web structure of a $50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc}$ slice in a simulation box of side length $100h^{-1}\text{ Mpc}$ and $128^3$ particles. The multistream field is calculated at 8 times the native resolution. void(black) is a percolating structure with $n_{str} = 1$ . Regions $n_{str} \geq 17$ show a filamentary structure (gray) and the bright spots at the intersections of the filaments are regions with $n_{str} \geq 100$ . . . . .	47
4.2	Opposite faces of the multistream field for the simulation box with $N_p = 128^3$ . Non-void regions (gray) have $n_{str} > 1$ . The largest void (white) in the entire field spans over the entire box. Rest of the smaller isolated voids (red) occupy very small volume fraction. . . . .	49
4.3	Single-streaming void distribution on diagnostic spheres around FOF-haloes are considered. At radius $r_{10}$ , each diagnostic sphere has $n_{str} = 1$ on 10 per cent of its spherical surface. Distribution function of $r_{10}$ (blue) and FOF-radii $r_{vir}$ (red) are shown. Inner plot shows the distribution function of $r_{10}/r_{vir}$ . The haloes within the dashed line have at least 10 per cent of their virial-surfaces in contact with $n_{str} = 1$ regions. . . . .	52
4.4	Percolation plot in the multistream field and mass density. Two density estimators - CIC and DTFE are shown. Percolation transition (at $f_1/f_{ES} = 0.5$ shown by the horizontal red line) occurs at smaller excursion set volumes for the multistream field, as seen by the dashed lines for both the curves. It is worth stressing that the percolation curves for $n_{str}$ field are bounded by conditions $f_{ES} < 0.1$ . . . . .	53

4.5	Percolation threshold in the multistream (left panel) and matter density fields. Matter density is calculated using DTFE (middle panel) and CIC (right panel) with a refinement factor of 2 in the simulation with $256^3$ particles. The volume fraction of excursion set and the filling fraction of the largest structure is shown. Percolation transition in multistream field at $n_{str} = 17$ is shows by the dashed vertical line. Percolation at $\rho_{DTFE}/\rho_b = 5.16$ and $\rho_{CIC}/\rho_b = 5.49$ are shown by the dashed vertical line. . . . .	54
4.6	Probability distribution function of the sorted eigenvalues of the Hessian $\mathbf{H}(-n_{str})$ in the simulation box with $N_p = 128^3$ . Top panel: Distribution in the entire simulation box. The multistream field is calculated at refinement factor $l_l/l_d = 2$ and smoothing scale of equal to $l_d$ . All the three eigenvalue data fields have a highest number of points where their value is 0. Bottom panel: Hessian eigenvalues for the non-void region ( $n_{str} > 1$ ) is shown. Total number of eigenvalue triplets are less than 10 per cent of that of the full simulation box. Eigenvalues close to zero in non-void regions are notably fewer than in the entire simulation box. . . . .	57
4.7	Comparison between small eigenvalues of the multistream Hessian $\mathbf{H}(-n_{str})$ . Percentage of eigenvalues with absolute values less than a cut-off, $\lambda_{th}$ are shown for full simulation box (dashed lines) and the multistream web structure (solid lines). The multistream web has fewer eigenvalues below $\lambda_{th} = 0.1$ . The void seems to have most of the small eigenvalues. . . . .	58
4.8	Eigenvalues of the Hessian matrix $\mathbf{H}(-n_{str})$ in a slice of $50 h^{-1} \text{ Mpc} \times 50 h^{-1} \text{ Mpc}$ slice of the simulation box of $128^3$ particles. Variation in the eigenvalues in the multistreaming web structure is shown. The largest eigenvalue $\lambda_1$ (top left panel) has positive values throughout the structure. The smallest eigenvalue $\lambda_3$ (bottom left) has negative values surrounding positive definite regions of the $n_{str}$ field. Corresponding multistream field is shown in the bottom right panel for single, three and more than five streams. . . . .	60

4.9	Surfaces identified in the multistream field. Blue regions are closed regions with $\lambda_3 > 0$ , which we identify as two haloes. Other surface has an open curvature along one direction, with $\lambda_1 > \lambda_2 > 0$ and $\lambda_3 < 0$ . . . . .	61
4.10	Probability distribution function of the multistream $n_{str}$ values in the simulation box with $N_p = 256^3$ . The multistream field is calculated at refinement factor $l_l/l_d = 2$ . Unsmoothed multistream field is compared with different Gaussian filtering scales. Softening scales of equal to 0.5, 1, and 2 times the side length of diagnostic grid $l_d$ correspond to $0.10h^{-1}\text{Mpc}$ , $0.20h^{-1}\text{Mpc}$ , and $0.39h^{-1}\text{Mpc}$ respectively. . .	63
4.11	Multi-stream field at various softening scales in the simulation box with $N_p = 256^3$ . The multistream field is calculated at refinement factor $l_l/l_d = 2$ . Unsmoothed multistream field is compared with different Gaussian filtering scales equal to $0.10h^{-1}\text{Mpc}$ , $0.20h^{-1}\text{Mpc}$ , and $0.39h^{-1}\text{Mpc}$ respectively. . . . .	64
4.12	Two top panels show three contours ( $n_{str} = 3, 11, 17$ ) in a slice $100h^{-1}\text{Mpc} \times 100h^{-1}\text{Mpc} \times 10h^{-1}\text{Mpc}$ in the simulation with $128^3$ particles, computed at two refinement factors: 2(upper) and 8 (lower). The bottom panel shows the caustic surfaces in the same slice. . . . .	68
5.1	Dynamical collapse of dark matter in one-dimensional universe: Top panels show the $(\mathbf{p}, \mathbf{x})$ phase-space manifold of the dark matter sheet at redshifts $z_1, z_2, z_3$ and $z = 0$ . Dots represent the dark matter particles. The momentum values are chosen at arbitrary scales. Bottom panels show the corresponding multistream field multistream field $n_{str}(\mathbf{x}, z)$ (red) and density field $\rho(\mathbf{x}, z)$ (gray). . . . .	78
5.2	Multistream field contours: The multistream field is calculated at 8 times the native resolution. Each inner convex blobs (red) surround local multistream maxima inside. Surrounding outer shell(blue) is not convex throughout the surface, and the outermost gray multistream surface displays a filamental geometry. . . . .	82

5.3	Detection of potential halo candidates in the multistream field: algorithm of segmentation and filtering are illustrated in a smaller slice of $40h^{-1}Mpc \times 40h^{-1}Mpc$ slice of the simulation box. Top left figure shows the multistream field of the slice. Voids (white) are the regions with $n_{str} = 1$ , rest are non-void structures. Blue patches within the structure (gray) are the regions with gravitational collapses in more than one direction, i.e., $n_{str} \geq 7$ . Top Right figure shows the smallest eigenvalue $\lambda_3$ field. The value of $\lambda_3$ is close to 0 in most of the regions (yellow), including the voids. Regions with $\lambda_3 > 0$ and $n_{str} > 1$ , are isolated (black spots) using image segmentation techniques. Bottom left panel shows the filtering scheme: the red patches do not have maxima of $n_{str} \geq 7$ in the regions, hence are filtered out. The remaining potential halo regions with more than 20 particles are shown in the bottom right panel.	87
5.4	PDF of highest $\lambda_1$ , $\lambda_2$ and $\lambda_3$ values in each of 4492 haloes detected by our algorithm. The peaks of the PDF are in the range 1-10. Shaded regions represent $1\sigma$ error.	88
5.5	The distribution of $\lambda_3$ in the simulation box of $128^3$ particles and multistream field of refinement factor $l_l/l_d = 2$ . Three smoothing scales are shown.	89
5.6	Top panel shows Halo Mass functions of the potential $\lambda_3$ -haloes in the multistream field refinement factor $l_l/l_d = 2$ with various smoothing scales. Simulation box has $256^3$ particles. Lower panel shows the deviation of the each halo mass function with respect to their average.	91
5.7	Halo Mass functions of the potential $\lambda_3$ -haloes in the multistream field refinement factor $l_l/l_d = 2$ (in Simulation box with $N_p = 256^3$ ) with various thresholds on local maxima of $n_{str}$ within the halo. Lower panel shows the deviation of the each halo mass function with respect to their average.	93

5.8	Maximum, minimum and median of $n_{str}$ in each of 4492 halo candidates. The closed contours of haloes encompass a wide range of multistream values. None of the haloes are in contact with the void region, since lowest value of $\min(n_{str})$ is 3. Shaded regions are the $1\sigma$ absolute errors in the number of $\lambda_3$ -haloes. . . . .	94
5.9	Multistream environment of a $\lambda_3$ -halo. The contours represent regions with 3 different multistream values: Outermost $n_{str} \geq 25$ (gray) is tubular, The blue region has $n_{str} \geq 75$ . The inner region is highly non-linear with $n_{str} \geq 200$ . The black dots represent the mass particles belonging to a $\lambda_3$ -halo, as detected by our algorithm. .	95
5.10	Halo Mass functions from AHF-, FOF- and $\lambda_3$ -haloes. The AHF-haloes are fewer than FOF- or $\lambda_3$ -haloes. The number of haloes above a mass threshold are binned and taken along vertical axis, normalized to simulation box volume. Error of $1\sigma$ is shown in shaded region. . . . .	97
5.11	A large halo that was detected by AHF (top, red), our geometric analysis in the multistream field (centre, green) and FOF (bottom, blue). Halo boundary differs for each halo finder method. AHF detects particles within a sphere of virial radius. FOF-halo is irregularly shaped. $\lambda_3$ -halo particles are in a non-spherical, yet compact structure. . . . .	98
5.12	Potential haloes detected by AHF (top), our analysis (centre) and FOF (bottom). Most of the haloes are embedded in a percolating filament with $n_{str} \geq 9$ . AHF-haloes are spherical by definition. FOF-haloes without any post processing are elongated along the filament. $\lambda_3$ -halo candidates are neither spherical, nor elongated. Boundaries of $\lambda_3$ -haloes are well resolved. . . . .	100
5.13	Mass fraction of haloes $f_h$ (in per cent) as a detected by the three finders. Each circle represents fractions of mass of all halo particles (AHF, FOF or $\lambda_3$ ) in the total mass of the simulation box. The particles concurrently detected as belonging to haloes by different frameworks are shown in the intersections. . . . .	103

5.14 Percentage of haloes detected (in the simulation with $N_p = 256^3$ ) that are embedded in the non-excursion set (top panel), largest excursion set segment (center panel) and the rest of excursion set regions (bottom panel). Multistream haloes, AHF-haloes and FOF-haloes are shown in green, red and blue respectively. Vertical dashed line at $n_{str} = 17$ is where percolation transition occurs. . . . .	104
B.1 Multi-streaming in one-dimension gravitational collapse. Top panel: $(\mathbf{p}, \mathbf{x})$ phase-space representation redshift $z_{ini}$ and $z = 0$ . Dots represent the dark matter particles. Initially the mass particles are in the linear stage of evolution. At $z = 0$ , multiple values of $\mathbf{p}(\mathbf{x})$ is seen in the collapsed regions. Middle panel: Equivalent Lagrangian sub-manifold $\mathbf{q}(\mathbf{x})$ . At $z_{ini}$ , the dashed line represents $\mathbf{q} = \mathbf{x}$ . Number of streams are parametrized from this sub-manifold. Bottom panel: The multistream field $n_{str}$ and the number-density using CIC algorithm, $n_{CIC}$ at $z = 0$ . . . . .	129
C.1 Projections of regions of $f(x)$ from (1+1)-dimensional function space onto one-dimensional coordinate space. Convex regions and regions above a threshold of an arbitrary function $f(x)$ are shown. Both the regions intersect around a few maxima, but not universally. . . . .	134

# List of Tables

3.1	Comparison of the approximate power law dependences of curve fits in Fig. 4.3 and Fig. 3.4. Power law relations for volume fraction $f_{vol}(n_{str})$ and mass fractions $f_{mass}(n_{str})$ as a function of number of streams at $n_{str} \geq 5$ are shown (amplitudes are not shown). The boxes of size $100 h^{-1}$ Mpc on $128^3$ grids and, $200 h^{-1}$ Mpc on $256^3$ grids have same $L/N = 0.78h^{-1}$ Mpc. Similarly, $L/N = 0.39h^{-1}$ Mpc for boxes of size $100 h^{-1}$ Mpc on $256^3$ grids and, $200 h^{-1}$ Mpc on $512^3$ grids. . . . .	18
3.2	Comparison of the volume and mass fractions of the void ( $n_{str} = 1$ ) regions of the cosmic web for various simulation boxes at refinement factor of 1. Mean density is the ratio of mass fraction to the volume fraction. It is given in units of the mean density of the universe. The boxes of size $100 h^{-1}$ Mpc on $128^3$ grids and, $200 h^{-1}$ Mpc on $256^3$ grids have same $L/N = 0.78h^{-1}$ Mpc. Similarly, $L/N = 0.39h^{-1}$ Mpc for boxes of size $100 h^{-1}$ Mpc on $256^3$ grids and, $200 h^{-1}$ Mpc on $512^3$ grids.	21
3.3	Comparison of the volume and mass fractions of the void ( $n_{str} = 1$ ) regions of the cosmic web for a simulation box at different refinement factors. Mean density is the ratio of mass fraction to the volume fraction. All the multi-streams for simulation box of length $100 h^{-1}$ Mpc on raw resolution of $128^3$ grids ( $L/N = 0.78h^{-1}$ )	
. . . . .	. . . . .	21

3.4	Comparison of the volume and mass fraction of the elements of the cosmic web between our analysis and Falck & Neyrinck (2015). $L/N = 0.78h^{-1}$ Mpc for simulations used in both techniques. We use a refinement factor of 8 for the multi-stream grid. The mean density is given in units of the mean density of the universe. . . . .	37
4.1	Parameters for the simulation boxes: Side length $L$ , number of particles $N_p$ , mass of each particle $m_p$ , and the gravitational softening length $\varepsilon$ for the GADGET simulations are shown. . . . .	46
4.2	Volume fraction $VF_V$ of the voids, total number of isolated voids $N_V$ and the filling fraction of the largest void $FF_1/VF_V$ at different refinement factors $l_l/l_d$ . The filling fractions of the largest void at each refinement factor show that most of the $n_{str} = 1$ region is almost entirely a single percolating structure. . . . .	50
4.3	Volume fraction (in per cent) of $n_{str}$ thresholds for cosmic web structures as defined by Ramachandra & Shandarin (2015). Multi-stream field is calculated at 1, 2, 4, and 8 times the native simulation resolution of $64^3$ grids. Small slice of $50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc}$ is chosen for the analysis. . . . .	65
4.4	Volume fraction of criteria based on $n_{str}$ and $\lambda$ s of $\mathbf{H}(-n_{str})$ calculated at various resolutions. We chose a smaller slice of $50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc}$ i.e., half the volume of the original GADGET simulation box. The refinement factors are the multiplication factors of 1, 2, 4 and 8 times of the native resolution ( $64^3$ ) of the simulation grid along each axis. Eigenvalues of the Hessian of the field are local geometric parameters. The void is globally defined as $n_{str} = 1$ and the multistream web structure as $n_{str} > 1$ . . . . .	66
5.1	Statistics of the Hessian eigenvalues in the halo candidates . . . . .	88
5.2	Number of $\lambda_3$ -haloes identified at smoothing of $n_{str}(\mathbf{x})$ at different scales. . . . .	90

5.3 Local maxima and minima of $n_{str}$ in each of 4492 haloes. The highest $n_{str}$ values in the interior of haloes span over a large range of values. Low values of $n_{str}$ in haloes, which are generally near halo boundaries, have a median of 3. . . . .	94
5.4 Number of haloes, $N_H$ detected by the three halo finder algorithms in the two simulations of $L = 100h^{-1}\text{Mpc}$ with different mass resolutions. Values shown for $\lambda_3$ -haloes are done in the multistream fields with refinement factor of 2, and smoothing scale equal to the diagnostic grid size. . . . .	97

# Chapter 1

## General Introduction

### 1.1 Large Scale structure

### 1.2 Observation history

### 1.3 Cosmic fields: Characterizing non-linear growth of Cosmic structure

The most fundamental attributes of particles in the N-body simulations are their position and velocity co-ordinates, and their masses. Due to the lack of numerical tools for direct analysis these raw data, fields such as mass density  $\rho(\mathbf{x}, t)$ , velocity  $\mathbf{v}(\mathbf{x}, t)$  or gravitational potential  $\phi(\mathbf{x}, t)$  fields are often computed numerically. Mass density fields were calculated using Cloud-in-Cell (CIC) algorithm (cf. Hockney & Eastwood 1988), which is numerically equivalent to counting the number of particles on each cell of a regular grid. Alternatively, the density field also generated on irregular grids by applying Delaunay (For example, Icke & van de Weygaert 1991 and the Delaunay Tessellation Field Estimator (DTFE) by Schaap & van de Weygaert 2000 and van de Weygaert & Schaap 2009) or Voronoi tessellations (See Schaap & van de Weygaert 2000 and references therein) to the particle coordinates. Another parameter ‘linking length’, using distances between

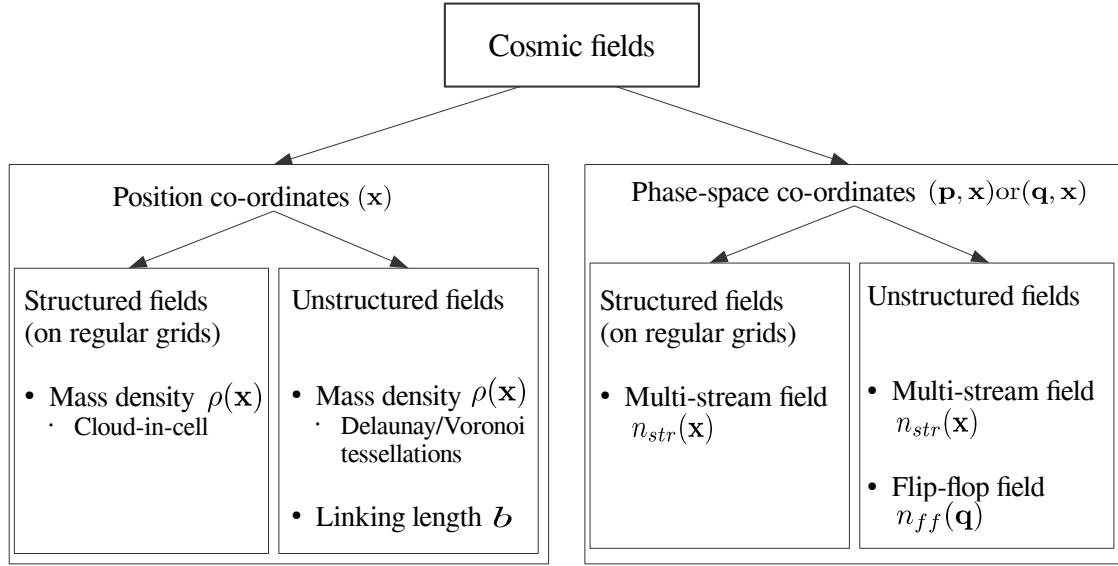


Figure 1.1: Classification of some of the fields/parameters used in cosmological analyses. Some fields utilise position co-ordinates only whereas others use the full phase-space information. In addition, the fields may be defined on a regular grid, or may be defined on an unstructured grid (For instance, flip-flop is a number-valued field defined on each dark matter particle). Fields like mass density and multi-streams can be defined on either grids, depending on the numerical technique. List is obviously not exhaustive- velocity and potential fields are not included, and discussions of correlation functions are excluded as well.

nearest neighbouring particles, was used for percolation analyses and identifying super-clusters of galaxies ( Zeldovich et al. 1982, Shandarin 1983a and Shandarin & Zeldovich 1983) for identifying halos Davis et al. 1985 in the cosmological simulations. Left panel in Figure 1.1 shows some of the popular fields/parameters that use particle mass and positions. It has to be noted that the density fields or linking-lengths are not dynamical descriptions that invoke the initial field of density fluctuations or the velocity of the particles.

An obvious advantage of methods based on particle coordinates, both on structured and unstructured grids, is their applicability to redshift catalogues. The redshift catalogues like SDSS and 2dF provide only two angular coordinates and distances in redshift space. But cosmological N-body dark matter simulations provide the full dynamical information in six-dimensional phase space. This additional information is very valuable providing a greater opportunity for understanding the physics of the web and developing a better theory of the web.

The velocity fields in the simulations of collision-less cold dark matter particles can become multi-valued under the action of gravity. This phenomenon was first discussed by Zeldovich (1970), where he predicted the formation of non-linear structures (also see Shandarin & Zeldovich 1989 for discussion on formation of multi-stream). The primordial oblate structures were later known as ‘Zel’dovich pancakes’. These pancakes grow from initial perturbations in a continuous mass distribution, where the velocities are single-valued (also referred to as single-stream) everywhere in the configuration space. Multiple values in the velocity field  $\mathbf{v}(\mathbf{x}, t)$  or ‘multi-streams’ can also be seen in the dynamically equivalent Lagrangian sub-manifold -  $(\mathbf{q}, \mathbf{x})$ , where  $\mathbf{x}$  and  $\mathbf{q}$  are co-moving Eulerian and Lagrangian co-ordinates respectively. Shandarin (2011) and ? studied this  $\mathbf{q} \mapsto \mathbf{x}$  mapping in N-body simulations to quantify the number of streams using phase-space tessellations. Shandarin (2011) define a multi-stream field  $n_{str}(\mathbf{x})$  as a field taking discrete values that are equal to the number of streams at every evaluation point in configuration space. Ordered sign-reversal of each elementary volume element in the Lagrangian sub-manifold was measured by ?. Their *flip-flop* field  $n_{ff}(\mathbf{q})$  in Lagrangian space demonstrates a very rich sub-structure of the cosmic web, especially in a halo environment.

Fields computed from a complete dynamical information (either  $(\mathbf{q}, \mathbf{x})$  or  $(\mathbf{p}, \mathbf{q})$ ) could provide valuable contributions to our understanding of the cosmic structure. Falck et al. (2012) have recently delineated archetypal web structures by counting the number of foldings in the sub-manifold for each dark matter particle along different directions. Another study by Ramachandra & Shandarin (2016) explored some of the global topological and local geometrical properties of the web in the context of multi-streaming. The applications of these analyses is certainly not limited to diagnostic tools; the multi-streaming phenomenon can be used in improving N-body simulations (Hahn et al., 2013), and studying galaxy evolution and star formation as well (?).

## 1.4 Chapter Organization

# Chapter 2

## Background

### 2.1 Zel'dovich Approximation

ZA is an elegant analytical approximation to describe the non-linear gravitational evolution of collisionless particles in continuous media. Technically it is the first order Lagrangian perturbation theory, however Zeldovich suggested to extrapolate it to the beginning of the non-perturbative nonlinear stage and predicted the formation of caustics which are the boundaries of the first very thin multistream regions dubbed by him 'pancakes'. ZA describes a dynamical mapping from the initial Lagrangian coordinates  $\mathbf{q}$  to Eulerian positions at time  $t$ . In comoving coordinates,  $\mathbf{x} = \mathbf{r}/a(t)$  (where  $\mathbf{r}$  is the physical coordinate and  $a(t)$  is the scale factor; (?) assuming normalization  $a(z=0) = 1$ ,  $\mathbf{r}$  are the physical coordinates of particles at present), ZA takes the form:

$$\mathbf{x}(\mathbf{q}, D(t)) = \mathbf{q} + D(t)s(\mathbf{q}) \quad (2.1)$$

where  $D(t)$  is the linear density growth factor, and the initial density perturbation field  $\psi(\mathbf{q})$  determines the potential vector field  $\mathbf{s}(\mathbf{q}) = -\nabla_q \psi(\mathbf{q})$ . Mass conservation formalism implies  $\rho(\mathbf{x}, t)d\mathbf{x} = \rho_0 d\mathbf{q}$ , so the density field at  $t > 0$  in terms of Lagrangian coordinates is given as

$$\rho(\mathbf{q}, t) = \rho_0 J \left[ \frac{\partial \mathbf{x}}{\partial \mathbf{q}} \right]^{-1} \quad (2.2)$$

where the Jacobian  $J \left[ \frac{\partial \mathbf{x}}{\partial \mathbf{q}} \right]$  is calculated using Equation 2.1. Moreover, diagonalization of the resulting real, symmetric deformation tensor  $d_{ij} = -\nabla_q \mathbf{s}(\mathbf{q}) = \partial^2 \psi(\mathbf{q}) / \partial q_i \partial q_j$  in terms of its eigenvalues  $\lambda_1(\mathbf{q}), \lambda_2(\mathbf{q}), \lambda_3(\mathbf{q})$  gives the contraction or expansion along the three principal axes. This reduces the mass density to a convenient form:

$$\rho(\mathbf{q}, t) = \frac{\rho_0}{[1 - D(t)\lambda_1(\mathbf{q})][1 - D(t)\lambda_2(\mathbf{q})][1 - D(t)\lambda_3(\mathbf{q})]} \quad (2.3)$$

Since the deformation tensor  $d_{ij}$  and its eigenvalues depend only on the initial density field, the ordered eigenvalues defined in Lagrangian space  $\lambda_1(\mathbf{q}) \geq \lambda_2(\mathbf{q}) \geq \lambda_3(\mathbf{q})$  determine collapse condition for masses in Eulerian space (see Doroshkevich (1973) and Lee & Shandarin (1998) for the PDFs of the eigen values as well as several other parameters in the case of Gaussian random fields). In the context of this paper, formation of caustics is of much interest: with increasing  $D(t)$ , the mass density rises until reaching singularity at  $D(t) = 1/\lambda_3(\mathbf{q})$ . In **Lagrangian space, the caustics stem from these points and their counterparts in Eulerian space** were proposed by Zel'dovich as the '**birthplaces**' of the first collapsed structure by gravitational clustering. Now **the regions bounded by the caustic surfaces** are referred to as *Zel'dovich pancakes*. The collapse along other principal axes correspond to formation of filaments and knots (Arnold et al. (1982), Shandarin and Klypin (1984), however simultaneous collapses along all three eigen axes never happen in the case of generic flows.)

Furthermore, the analytical understanding of these structures are thoroughly complicated: In a 2-dimensional ZA, for example, there are **only two types of fundamental singularities that exist at generic instants of time** ( $A_2$ , which are lines and  $A_3$ , which are the cusp points of  $A_2$  lines). In addition there are two singular points ( $A_4$ , and  $D_4$ ) that exist only at particular instants of time: at  $A_4$  two cusp  $A_3$ -points are formed and a smooth part of an  $A_2$  line is transformed in self-crossing line. Moreover there are **In addition there are several transient forms that exist only at particular times.** Each of these correspond of formation, mergers, branching and other dynamical processes involving pancakes. Arnold et al. (1982) and Hidding et al. (2014) studied of singularities in 2-dimensional collapse) in exhaustive detail, but similar analytical characterization of 3-dimensional

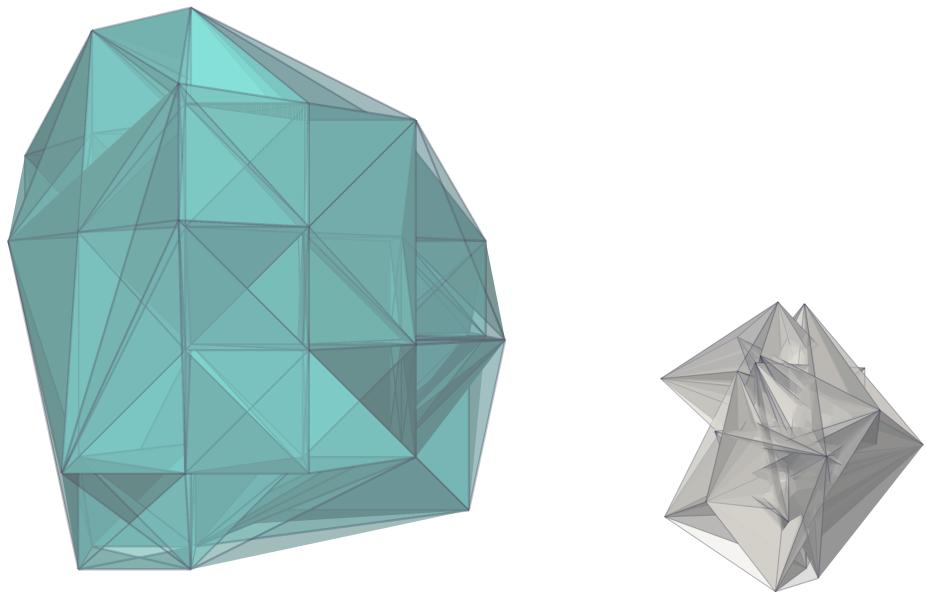


Figure 2.1: This plot shows the idea of Lagrangian tessellations. Left: The distribution of Dark matter particles in Lagrangian space is on the regular grid. The tetrahedra surfaces in this case are mostly regular (except at the edges). Right panel shows the Eulerian positions of the same particles at  $z = 0$ . The particles clearly have undergone multiple flip-flops, and the intersections of Lagrangian tetrahedra signifies locations of caustic surfaces. Specific tessellation schemes can be utilized separately to identify these surfaces.

ZA has not been satisfactorily done yet.

Complexities in 3-dimensional caustics is partly due the intricate mapping in the hypersurface  $\mathbf{x}(\mathbf{q})$  called the Lagrangian submanifold (See Figure 2.1). The Lagrangian submanifold  $\mathbf{x}(\mathbf{q})$  is a single valued, smooth and differentiable function in it's 6-dimensional space  $(\mathbf{q}, \mathbf{x})$ , however the projection onto 3-dimensional Eulerian space is entangled with creases, kinks and folds (Note that this submanifold is very different than the phase space  $(\mathbf{x}, \mathbf{v})$ , even though they are connected by a canonical transformation). However, delineating the Lagrangian submanifold reveals several properties of the dark matter dynamics not inferred from position-space analyses. Two fields related to tessellating the Lagrangian submanifold – The Multistream field  $n_{str}(\mathbf{x})$  in Eulerian space and the Flip-Flop field  $n_{ff}(\mathbf{q})$  in Lagrangian space (check Shandarin et al. (2012), Ramachandra & Shandarin (2015), Shandarin & Medvedev (2017)) are closely related.

## 2.2 Dynamics: Lagrangian sub-manifold

Multi-stream flows, Caustics, Flip-flips

# Chapter 3

## Multi-stream portrait of the Cosmic web

### 3.1 Introduction

The problem of objective identification and quantitative characterization of anisotropic structures in the distribution of galaxies in space emerged after the first evidences of their existence (see the review by Oort 1983 and the references therein). The first theoretical model predicting highly anisotropic concentrations in the mass distribution coming into existence at the non-linear stage of gravitational instability is known as the Zeldovich Approximation (the ZA) (Zeldovich 1970, for further developments see also Shandarin & Zeldovich 1989 and the references therein). The ZA predicted the formation of ‘pancakes’ also known as the walls in the currently popular jargon. The later development of the model by Arnold et al. (1982) predicted the formation of filaments along with the pancakes. Klypin & Shandarin (1983) and Shandarin & Zeldovich (1984) demonstrated that the filaments emerge in the cosmological N-body simulation in three-dimensional space. However, they failed to identify the pancakes at  $z = 0$ . Both the existence of filaments connecting compact clumps of matter and absence of pancakes were confirmed by Frenk et al. (1983). Puzzled by the absence of the pancakes Klypin & Shandarin (1983) speculated that insufficient mass resolution of the simulation was the cause of the negative outcome. This has been unambiguously confirmed by recent simulations using a better numerical technique of computing a density field from the

particle coordinates in cosmological N-body simulations ( Shandarin et al. 2012 and Abel et al. 2012). Klypin & Shandarin 1983 also stressed that the most of filaments are incorporated in ‘a single three-dimensional web structure’<sup>1</sup>. They admitted that their simulation did not allow them to confirm the existence of pancakes between the filaments predicted by the ZA (Arnold et al., 1982).

Although the four archetypical elements of the cosmic web: voids, walls/pancakes, filaments and haloes were predicted by ZA and confirmed in cosmological N-body simulation their identification and quantitative characterization remains under vigorous debate (see e.g. Colberg et al. 2008, Elahi et al. 2013, Knebe et al. 2013, Hoffmann et al. 2014). The dark matter haloes are arguably the easiest objects to identify in N-body simulations. They can also be reliably associated with observed objects like galaxies and clusters of galaxies. But even in this case Knebe et al. (2013) refer to almost thirty different halo finders suggested after 2000. Identifying filaments and pancakes/walls is far more controversial in both N-body simulations and galaxy catalogues. For instance, even estimating the global parameters of the web in N-body simulation such as the fractions of volume and mass in voids, walls/pancakes, filaments and haloes produced quite different results. The estimates of volume fractions of voids range from 13 to 86% (Cautun et al. 2014, Falck & Neyrinck 2015, Forero-Romero et al. 2009 , Hahn et al. 2007 , Aragon-Calvo et al. 2010). Similar estimates for walls/pancakes, filaments and haloes are respectively 5-56% 2-26% and 0.1-1%. Estimates of the mass content vary in large ranges as well.

Large differences in the estimates of volume and mass fractions made by different groups are not surprising if we recognize considerable differences in the definitions of the components of the cosmic web and numerical methods used in the estimates. Without trying to provide an exhaustive review of all definitions and techniques used for the quantitative morphological analysis of the web we just briefly describe a few approaches in order to illustrate how different they could be. Some groups study the web morphology using only coordinates of simulation particles, while others use the particle velocities too. Transforming data from point sets to the density and other

---

<sup>1</sup>The term ‘cosmic web’ was coined by Bond et al. (1996).

fields on a grid is also often used because fields allow to use a variety of mathematical techniques not available for the particle sets. However this step can be done by a variety of methods some as simple as cloud-in-cell (CIC), or more complicated as smoothed-particle hydrodynamics (SPH), or using Voronoi and Delaunay tessellations as in the Delaunay Tessellation Field Estimator (DTFE) method ( van de Weygaert et al. 2009, Cautun et al. 2014). Recently a new method called a discrete persistent structure extractor (DisPerSE Sousbie 2011, ? allowing to identify haloes and other components of the web directly from the particles has been designed. This method can be applied to the galaxy catalogues, for instance ? applied it to SDSS catalogue and extracted the filaments (which are made available online).

An obvious advantage of methods based on particle coordinates, both based on the density field and directly on particle coordinates, is their applicability to redshift catalogues. The redshift catalogues like SDSS and 2dF provide only two angular coordinates and distances in redshift space. But cosmological N-body dark matter simulations provide the full dynamical information in six-dimensional phase space. This additional information is very valuable providing a greater opportunity for understanding the physics of the web and developing a better theory of the web.

Dark matter distribution in phase space is highly degenerate because it is cold. Practically, it occupies a three-dimensional sub-manifold in six-dimensional phase space. In the linear regime, the dark matter sub-manifold is a single-valued function of Eulerian coordinates which means that at each point the dark matter is represented by a single stream flow. As the density perturbations in dark matter grow with time the number of streams jumps to three at the regions of shell crossing. Then five stream regions emerge inside of the three stream regions and so on. Number of streams remains an odd integer in generic points. The corresponding parts of the three-dimensional dark matter sub-manifold form complicated folds in six-dimensional phase space.

The regions with multi-stream flow constitute the web while the regions with only one stream form voids (Shandarin 2011, Shandarin et al. 2012, Abel et al. 2012, Falck et al. 2012). This definition of voids states that in a given N-body simulation, no haloes can be formed before the first shell crossings have occurred and the smallest haloes cannot be smaller than the mass cor-

responding to the small scale cut-off in the initial power spectrum regardless of the cause of the cut-off: physical or due to numerical limitations (see e.g. Angulo et al. 2013b). This definition of voids is physical by nature and thus has no free parameters. In addition, it does not speculate on the sub-grid physical processes. The first three-stream flow regions are similar to the pancakes in the ZA. They quickly grow and merge into a complicated three-dimensional structure; filaments making the framework of the web manifest themselves at the pancake crossings, and haloes emerge at the filament crossings. At later times different parts of the web participating in the large-scale motion overlap which increases the web complexity further.

Using the full six-dimensional information allows one to generate new fields which provide additional useful information about the evolution and morphology of the web. One of them is a multi-stream field in Eulerian space, which will be the focus of this paper. Another example is the flip-flop field in Lagrangian space. In cosmological context it was first used in the ZA. Vogelsberger & White (2011) used it in a study of multi-stream structure of galaxy size haloes. Shandarin & Medvedev (2014) applied it for identifying subhaloes in dark matter haloes. A similar although somewhat simplistic realization of this idea has been revealed in the ORIGAMI method used for the analysis of the web ( Falck et al. 2012, Falck & Neyrinck 2015 ). Although these fields cannot be used directly on observational data because the full phase-space information is not available, they provide much deeper insight into non-linear clustering of collision-less dark matter and reveal new features of the web.

In order to compute the multi-stream field we will use the tessellation scheme described in Shandarin et al. (2012) which is also briefly discussed in Section 3.3. Using this methodology on the entire simulation box, we discuss the global behaviour of the number of streams in the cosmic web in Section 3.4. The tessellation technique we have utilized can be used to find multi-stream fields in smaller Eulerian boxes with very high resolution too. In Section 3.5 we study the local behaviour of multi-streams flows in regions around haloes the detected using friends-of-friends (FOF) technique.

## 3.2 The simulation

We have utilized the data from cosmological N-body simulations by Gadget-2 (Springel, 2005) for  $100 h^{-1}$  Mpc and  $200 h^{-1}$  Mpc box sizes with  $128^3$ ,  $256^3$  and  $512^3$  grids. Each particle is between  $10^9 - 10^{12} M_\odot$ . The initial conditions and cosmological parameters are consistent with the Planck cosmology. We utilize the initial Lagrangian box and do a three-dimensional mapping onto corresponding evolved simulations. In addition, for local multi-stream analyses around haloes, we have utilized halo catalogues for each of these simulation boxes. These haloes are detected using FOF method considering objects with more than 20 particles found at linking length,  $b = 0.2$ .

## 3.3 Multi-stream field calculation

Phase space tessellation considers the dynamics of the particles similar to that of a standard N-body code. However the particles are nodes of the tessellation, and are just massless tracers of the flow. Assuming that the uniform state is modelled by a simple rectangular grid, the particles are the nodes of the grid. Each elementary cube of the grid is tessellated by five tetrahedra (Shandarin et al. 2012<sup>2</sup>) of which the vertices are the vertices of the cube. Mass is assumed to be uniformly distributed within each tetrahedron and the tessellation remain intact at all times. The tetrahedra of the tessellation change their shapes and volumes, the latter are used for computing the densities of the tetrahedra. Despite the complicated deformations experienced by the three-dimensional sub-manifold tessellated by the tetrahedra, it remains continuous. Projected on three-dimensional configuration space, the tetrahedra may form complicated structures. The number of streams at a chosen point  $\mathbf{x}$  is simply the number of tetrahedra that contain the point. The diagnostic points are computationally convenient to choose on a regular grid which can be significantly finer than the original grid in Lagrangian space. The ratio of separation of particles on the initial unperturbed grid to the separation distance of points in the diagnostic grid  $l_{\text{part}}/l_{\text{dg}}$  will be referred to as the refinement factor in the rest of the paper.

---

<sup>2</sup>For the description of an alternative type of the tessellation see Abel et al. (2012).

Number of streams are odd-valued in the entire configuration space, except in a set of points of measure zero where caustics are formed. A single-stream flow implies that the tetrahedra do not overlap in the corresponding region and thus defined as a void region. The web is defined as a set of non-void regions, i.e. the set of regions where the number of stream is equal to or more than three. The level of non-linearity in the web can be quantitatively characterized by using ‘number of streams’ as a parameter. As shown in (Shandarin et al., 2012) there is no simple local relation between the number of streams and density, however the both fields are obviously correlated.

### 3.4 Global statistics of Cosmic web

The 3-dimensional multi-stream field for the entire simulation box exhibits cosmic web structure with void, walls, filaments and haloes. We propose the number of streams, ‘ $n_{str}$ ’ as a parameter for characterizing and distinguishing structures in the universe. This is different from Falck & Neyrinck (2015), where the authors have identified voids, walls, filaments and haloes by particles which have undergone any number of flip-flops along 0, 1, 2 or 3 axes respectively. Their description of voids is close to ours except that some particles that have experienced no flip-flops might be in the region of multi stream flow formed by other particles. Thus we expect that the mass fraction in voids defined as the regions with  $n_{str} = 1$  is somewhat lower than that defined as the particles with flip-flops = 0 only at the final state. This is because some particles may have already fallen in the web but have not experienced flip flop yet and some particles that have experienced an even number of flip flops may come back to the original order. However the above arguments are valid only if the thickness of the web is the same in the both approaches. As we discuss below the thickness of the web in our analysis is about  $(100\% - 84\%)/(100\%-93\%)=2.3$  times thinner than in the analysis by Falck & Neyrinck (2015) (See Table 3.4 for details). However Falck & Neyrinck (2015) discussed these effects and claimed that they were small.

Whereas for non-linear structures, our parameter space has more freedom in terms of number of streams. Similar to the density threshold, the number of streams - used as a local parameter -

cannot distinguish unambiguously whether a point is in a wall, filament or halo. Only some parts of walls where there are only three streams can be identified locally without confusion. This is because the formation of a filament requires at least five streams. A flip-flop along one axis would produce a three-stream region which may be only a pancake. Therefore another flip-flop along the other axis in one of the streams from previous stage is required to transform it into a three-stream flow. Thus the total becomes five. However, if the second flip-flop happens along the same axis the resultant structure will remain a wall. Therefore some points in the five-stream flows can be within walls while the other in filaments. The present simulations have no information about the evolution of the flip-flop field therefore we rely on visual impressions initially to understand the transformation of walls into filaments and parts of filaments into haloes. By inspection, we have identified all the regions with three streams as walls. Unfortunately, walls are difficult to display on paper since they essentially block the view in two-dimensional projection. Nevertheless, we have demonstrated and analysed walls on a smaller Eulerian box around haloes in Section 3.5 using a simple and reasonably effective approximation.

For a multi-stream field calculated on a simulation box of size  $100h^{-1}$  Mpc and  $128^3$  particles, it is visually observed that with the increase of  $n_{str}$  from 3 to 15, the corresponding occupied regions increasingly belong to filamentary structure rather than the membrane like walls, until at the level  $n_{str} \gtrsim 17$  we observe that the number of wall points become negligible.

The filamentary structure of regions with 17 or more streams (denoted as 17+) is shown for a slice of simulation box of size  $100h^{-1}$  Mpc and  $128^3$  particles in Fig. 5.2. It has to be noted that all the regions with 17+ streams are regions with 3+ streams. Thus, the filaments are just interior parts of walls with higher  $n_{str}$ . These are visually observed mostly at the intersections of walls. Further, at the intersections of multiple filaments, there are regions with locally maximum number of streams, signifying the most dense regions in the simulations i.e. the dark matter haloes as Fig. 3.5 illustrates. By superimposing the positions from the FOF-halo catalogue, it is visually confirmed that the FOF haloes reasonably coincide with these high-streaming intersections, as Fig. 5.2 illustrates.

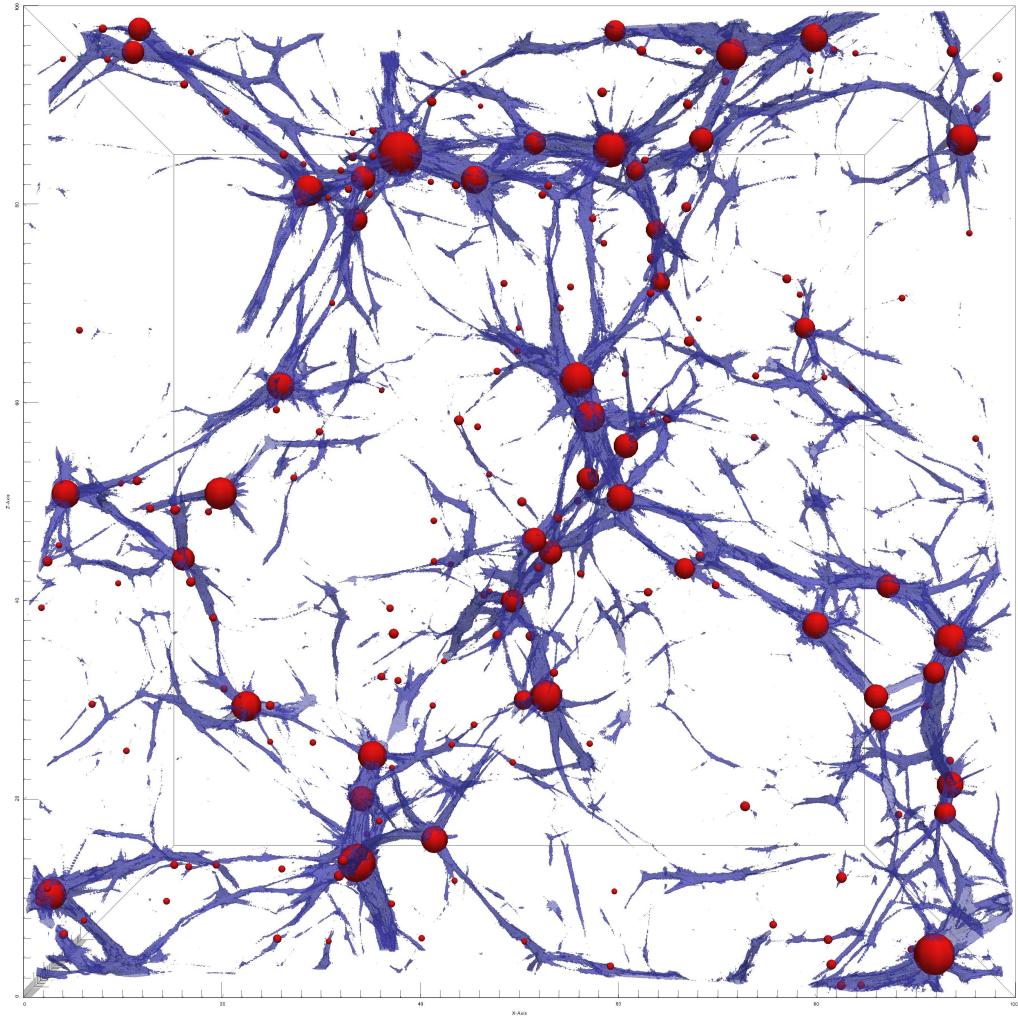


Figure 3.1: The cosmic web structure in a slice of  $30h^{-1}\text{Mpc} \times 100h^{-1}\text{Mpc} \times 100h^{-1}\text{Mpc}$  in a simulation box of size  $100h^{-1}\text{ Mpc}$  and  $128^3$  particles. Regions with 17 or more streams (blue) form a filamentary structure. The haloes determined by FOF (red) are predominantly embedded in the intersections of the filaments.

### 3.4.1 Volume and mass fractions

The single-stream flow, which corresponds to the void, occupies majority of volume of the simulation box (Fig. 3.2). As mentioned in Section 3.4, higher multi-streaming flow regions are nested inside the lower streaming regions. Thus the volume occupied by higher number of streams monotonically decreases with the number of streams. This relation is approximately found to be a power law. For the box of size  $L = 100 h^{-1} \text{ Mpc}$  and  $N = 128^3$  particles ( $L/N = 0.78 h^{-1} \text{ Mpc}$ ), the volume fraction corresponding to each value of number of streams,  $f_{vol}(n_{str})$  in the multi-stream field calculated with refinement factor of 8 (i.e. the multi-stream field was computed on  $1024^3$  grid as described in Shandarin et al. (2012)) is

$$f_{vol}(n_{str}) = 0.69n_{str}^{-2.5} \quad (3.1)$$

This is a good fit for the range of number of streams  $n_{str} \geq 5$ . In multi-stream field for the simulation box mentioned above, about 93% of the volume is occupied by 1-stream. With an increase in  $n_{str}$ , the corresponding volume fraction reduces. Physically, however, the number of streams reflect the advancement of non-linearity. Hence the higher  $n_{str}$  regions are typically regions with higher densities. In effect, the mass fraction can also be approximated by a decreasing power law function of  $n_{str}$ ,

$$f_{mass}(n_{str}) = 0.61n_{str}^{-1.3} \quad (3.2)$$

This is also a good fit for the range of number of streams  $n_{str} \geq 5$ . For the same range of number of streams, the mean density in the regions with particular number of streams, given by the ratio of corresponding mass and volume fractions, increases as expected.

$$\frac{\bar{\rho}(n_{str})}{\langle \rho \rangle} = 0.89n_{str}^{1.2}, \quad (3.3)$$

where  $\langle \rho \rangle$  is the mean density of the universe. This also quantifies our previous claim that very high multi-streams correspond to the most dense areas in the Universe, i.e. the condensed haloes.

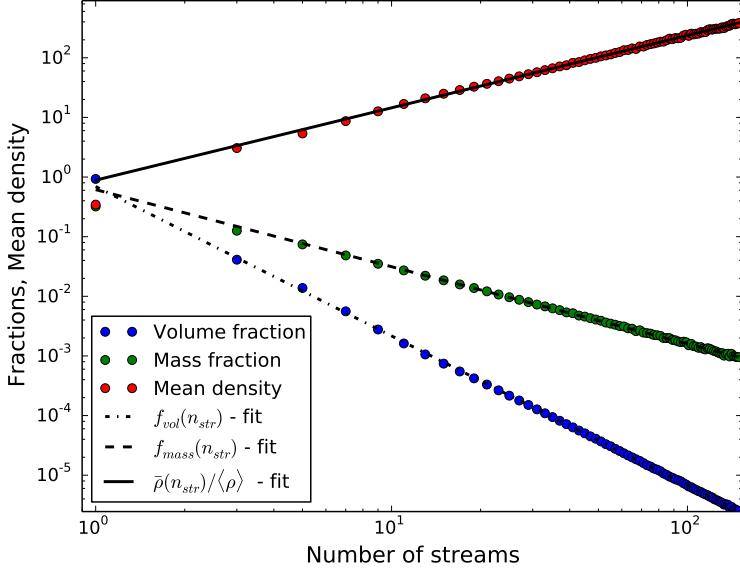


Figure 3.2: Volume and mass fraction of each stream, mean density of each stream in a box of size  $100 h^{-1}$  Mpc and  $128^3$  particles. Exact values of fractions, density and their curve fit for the range  $n_{str} \geq 5$  from Eq. 3.1, Eq. 3.2 and Eq. 3.3 are shown. Multi-streams are calculated with refinement factor of 8. The void ( $n_{str} = 1$ ) occupies 93% of the volume and 55% of the mass.

The common over-density threshold of 200 using virial equilibrium corresponds to roughly 90 streams in Fig. 3.2 and Eq. 3.3.

Comparing the volume fractions of various simulation boxes in Fig. 4.3 and corresponding power law dependences in Table 3.1 (also, specifically for the volume fraction of voids in Table 3.2), we find that the profile is similar for boxes with same inter-particle resolution; i.e., equal box length to grid size ratio( For e.g.,  $L/N = 0.78 h^{-1}$  Mpc for the simulation box of  $100 h^{-1}$  Mpc -  $128^3$  particles and  $200 h^{-1}$  Mpc -  $256^3$  particles). The box with minimum inter-particle resolution in the data, hence the best raw resolution (  $L/N = 0.19 h^{-1}$  Mpc for  $100 h^{-1}$  Mpc,  $512^3$  particles), has higher volume fraction for each multi-stream compared to lower resolution boxes. Additionally, it has a more non-linear stage advanced over time resulting from the initial small scale perturbations. The advancement of non-linearity manifests itself in higher number of streams. Box with the least raw inter-particle resolution ( $L/N = 1.56 h^{-1}$  Mpc for  $200 h^{-1}$  Mpc,  $128^3$  particles), occupies lower volumes than other boxes for each  $n_{str}$ . It is also prone to noise at very high streaming regions.

Table 3.1: Comparison of the approximate power law dependences of curve fits in Fig. 4.3 and Fig. 3.4. Power law relations for volume fraction  $f_{vol}(n_{str})$  and mass fractions  $f_{mass}(n_{str})$  as a function of number of streams at  $n_{str} \geq 5$  are shown (amplitudes are not shown). The boxes of size  $100 h^{-1}$  Mpc on  $128^3$  grids and,  $200 h^{-1}$  Mpc on  $256^3$  grids have same  $L/N = 0.78 h^{-1}$  Mpc. Similarly,  $L/N = 0.39 h^{-1}$  Mpc for boxes of size  $100 h^{-1}$  Mpc on  $256^3$  grids and,  $200 h^{-1}$  Mpc on  $512^3$  grids.

$L/N$	0.19	0.39	0.78	1.56
$f_{vol}(n_{str})$ Vs. $n_{str}$	-2.1	-2.3	-2.5	-2.9
$f_{mass}(n_{str})$ Vs. $n_{str}$	-1.1	-1.2	-1.4	-2.0

One of the advantages of using tessellation is the freedom to compute densities at very high resolutions (Abel et al. (2012), Shandarin et al. (2012)). We remind that the parameter ‘refinement factor’ denotes the ratio of separation of the particles to the separation distance of points in the diagnostic grid as defined in Sec. 3. High refinement factors are extensively used in understanding stream behaviour not only in the halo environment, but inside the halo too (Section 3.5). The volume fractions of resulting number of streams are similar for all refinement factors as shown in bottom of the Fig. 4.3 and in Table 3.3. Multi-stream fields calculated on low refinement factors are more noisy at high number of streams.

With same refinement factors, the mass fractions exhibit similar pattern for same  $L/N$  as well (Fig. 3.4). The simulation box with highest inter-particle distance (thus least mass resolution) has more mass particles in single streaming region, as tabulated in Table 3.2, but decreases steeply thereafter (Table 3.1). Unlike the volume fraction, the behaviour of mass fraction has a systematic variation across different refinement factors. The mass fractions given in Table 3.3 show that the single-streaming regions in the multi-stream fields with refinement factors 1 and 2 have higher mass fraction than in the fields with refinement factor of 8. Decreasing the resolution from refinement factor of 8 to 1 effectively introduces smoothing of the structure. This results in growth of mass fraction in voids ( $n_{str} = 1$ ) and decreasing it in the web ( $n_{str} > 1$ ). The multi-stream field is more robust as one can see in Fig. 4.3. and 3.4. In addition, the less refined multi-stream grids are prone to noise at high  $n_{str}$  as usual.

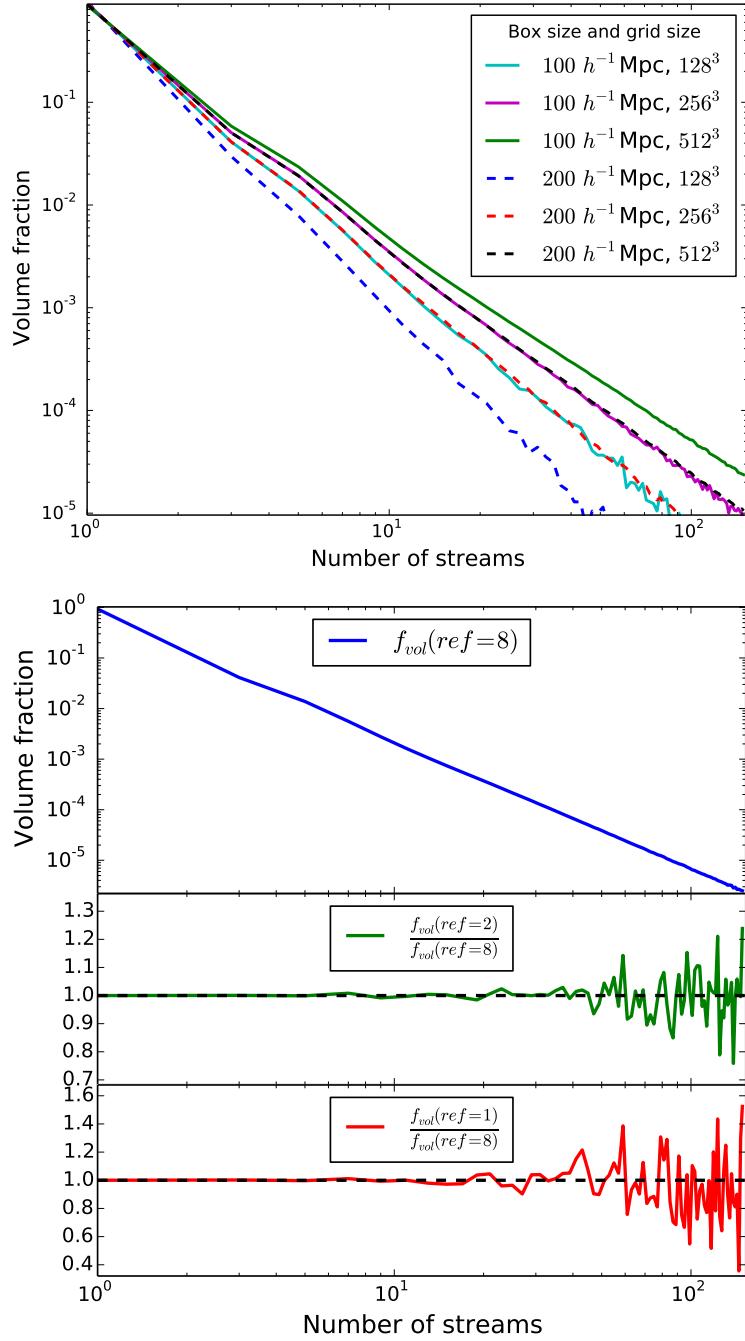


Figure 3.3: Top: Volume fraction distribution of streams in 6 simulation boxes of size  $100 h^{-1}$  Mpc,  $200 h^{-1}$  Mpc and  $128^3$ ,  $256^3$ ,  $512^3$  grids (with refinement factor of 1). Volume fractions are similar for simulation boxes with same inter-particle resolution. Slopes of the curve fits are shown in Table 3.1. Bottom: Volume fraction distribution for different refinement factors for  $100 h^{-1}$  Mpc,  $128^3$  box. A considerably smoother volume fraction distribution is obtained at high number of streams in multi-stream fields with high refinement factor.

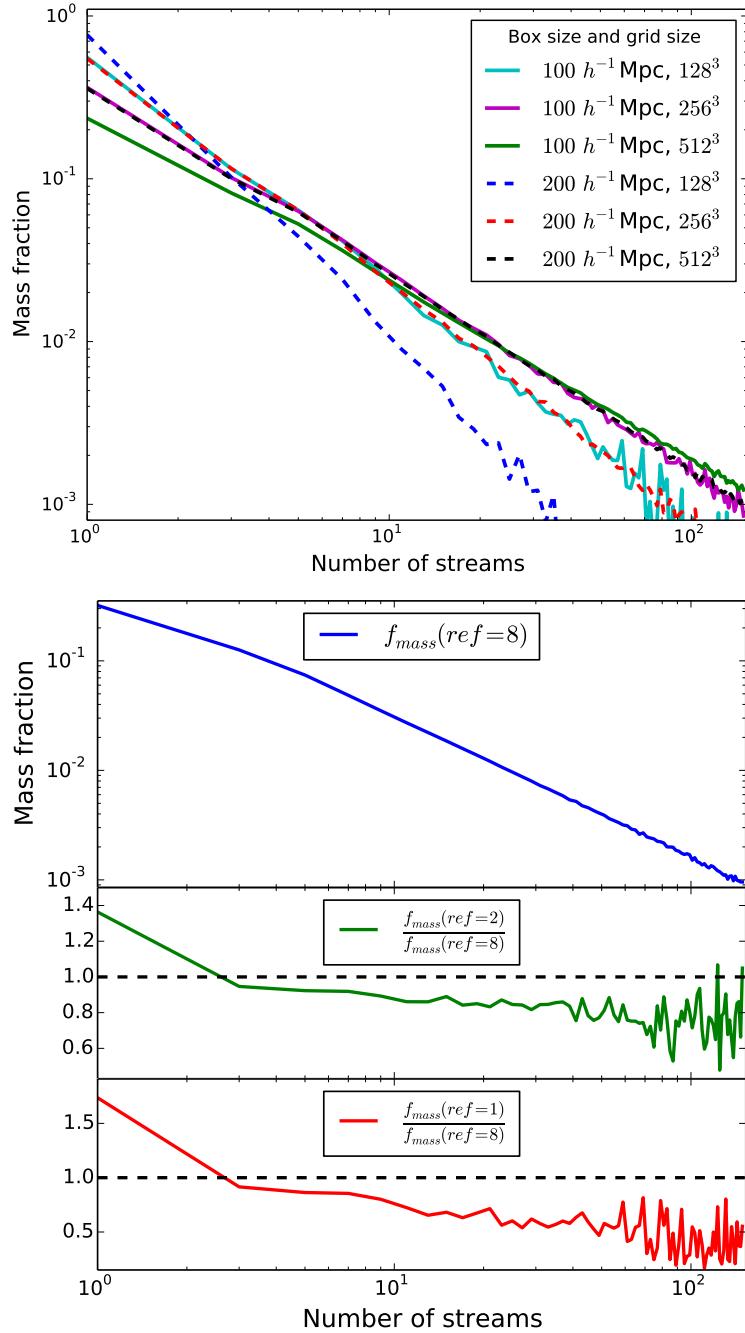


Figure 3.4: Top: Mass fraction distribution of streams in 6 simulation boxes of size  $100 h^{-1}$  Mpc,  $200 h^{-1}$  Mpc and  $128^3$ ,  $256^3$ ,  $512^3$  grids (with refinement factor of 1). Mass fractions are similar for simulation boxes with same inter-particle resolution. Slopes of the curve fits are shown in Table 3.1. Bottom: Mass fraction distribution for different refinement factors for  $100 h^{-1}$  Mpc,  $128^3$  box. Single-streaming void regions have more fraction of mass particles in multi-stream fields calculated at low refinement factors of 1 and 2. This effect is minimized when calculation is done at better refinement. In addition, mass fraction distribution is less noisy.

Table 3.2: Comparison of the volume and mass fractions of the void ( $n_{str} = 1$ ) regions of the cosmic web for various simulation boxes at refinement factor of 1. Mean density is the ratio of mass fraction to the volume fraction. It is given in units of the mean density of the universe. The boxes of size  $100 h^{-1}$  Mpc on  $128^3$  grids and,  $200 h^{-1}$  Mpc on  $256^3$  grids have same  $L/N = 0.78h^{-1}$  Mpc. Similarly,  $L/N = 0.39h^{-1}$  Mpc for boxes of size  $100 h^{-1}$  Mpc on  $256^3$  grids and,  $200 h^{-1}$  Mpc on  $512^3$  grids.

$L/N$	0.19	0.39	0.78	1.56
Volume Fraction (%)	88	90	93	96
Mass Fraction (%)	24	36	55	77
Mean density	0.27	0.40	0.59	0.80

Table 3.3: Comparison of the volume and mass fractions of the void ( $n_{str} = 1$ ) regions of the cosmic web for a simulation box at different refinement factors. Mean density is the ratio of mass fraction to the volume fraction. All the multi-streams for simulation box of length  $100 h^{-1}$  Mpc on raw resolution of  $128^3$  grids ( $L/N = 0.78h^{-1}$ )

Refinement factor	1	2	8
Volume Fraction (%)	93	93	93
Mass Fraction (%)	55	44	32
Mean density	0.59	0.47	0.35

### 3.5 Stream environment around haloes

Multi-stream field can be easily computed for a small Eulerian box with higher refinement factor. This can be utilized to analyse the phase-space behaviour inside and around haloes. In this section, we have used the halo coordinates identified by the FOF method, and selected Eulerian boxes around it. A reasonable correspondence between FOF halo centres and local maxima of multi-stream field is visually examined in Fig. 5.2.

Since each multi-stream region is surrounded by lower number of streams, the walls sandwich filaments within themselves (Fig. 3.5). The filaments are embedded with haloes at various intersections. These high-streaming haloes are completely covered by relatively low-stream filaments and hence surrounded by walls too. This result differs considerably from the several void finder methods, which find existence of haloes within void regions (See Colberg et al. 2008 and references therein). By our classification, we distinguish configuration space of the simulation box as void and non-void or web regions. Further, we have made an attempt to classify the web into

walls, filaments and haloes based on multi-stream thresholds. This classification based on number of stream threshold provides only a very crude description of visual impression from the richness, complexity and fundamentally multi-scale character of the web. The heuristic numbers we use in this paper are by no means universal, but may provide limited use in the discussion of these particular simulations.

Visual inspection of Fig. 3.5 reveals that the multi-stream environment of a halo is a highly intricate. Though the haloes are surrounded by filaments and walls, it can be surprisingly close to the voids in particular directions. Filaments defined by the multi-stream field are quite elongated, but the cross-sections are not circular or elliptical, and moreover, they branch-out and intersect at several regions. Finally, the haloes defined by contours of the multi-stream field look neither spherical nor ellipsoidal. We use a simple geometrical technique of projecting the number of streams onto a diagnostic spherical surface around a halo to visualize and quantify their environments.

### 3.5.1 Technique

Motivated by the complicated morphology of multi-stream field around a dark matter halo, we devised an empirical statistical tool to quantify the multi-stream environment of the FOF haloes. The method is geometrical and non-local. We randomly select a large number of points on a diagnostic spherical surface centred at the FOF centre of the halo and compute the number of streams at these points. By increasing the radius of the sphere from inside of the halo to several times the halo radii, we estimate the fractions of the area on the diagnostic spherical surface occupied by the regions with different numbers of streams:  $3+$ ,  $5+$ , ..., where  $n+$  corresponds to  $n$  or higher number of streams.

The geometry of a filament can be crudely approximated by a cylinder and that of a wall by a sheet with a small constant thickness ' $d$ '. Upon intersecting with the spherical surface, these geometries occupy certain cross-sectional area,  $Area_{c/s}$ , on the sphere (See Fig. 3.6). The ratio of this area to the surface area of the sphere is given by Equation 3.4 and Equation 3.5,

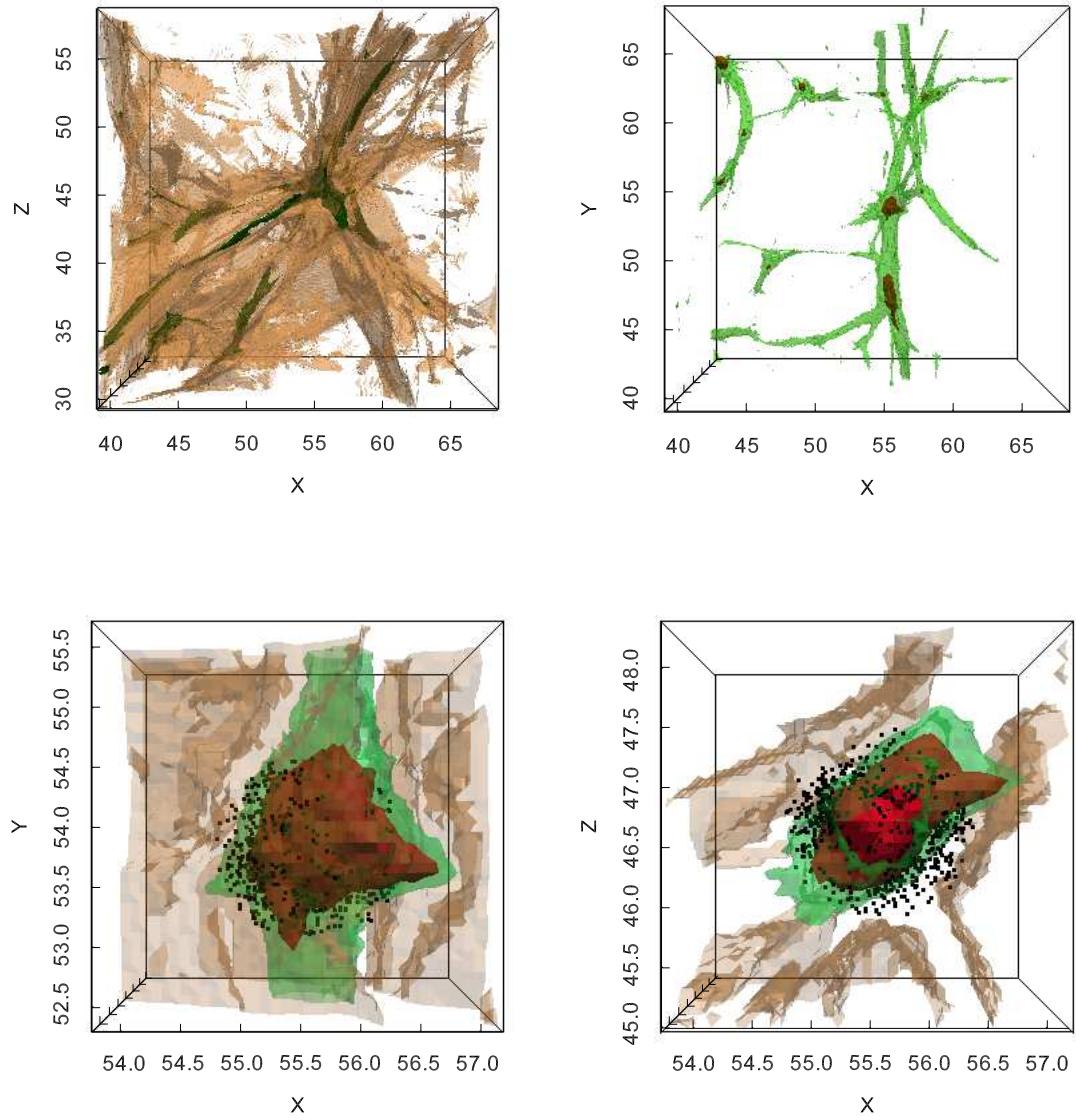


Figure 3.5: Multi-stream flow regions in a small box of the simulation. Top left: regions with more than 3-stream flow are identified as walls (brown). Intersection of multiple walls have higher  $n_{str}$  regions (green, 17+ streams). Single-streaming voids (white) occupy large volume and are very close to the filaments in some directions. Top right: 17+ streams (green) form filamentary structures with nodes at the intersections (red, 90+ streams). Bottom left and right: Closer look at the highly non-linear region reveals that a filament is sandwiched between the walls (brown). The 90+ stream region (red) forms a compact structure and is entirely contained within the filament. The black dots show the particles around the FOF halo within linking length of 0.2.

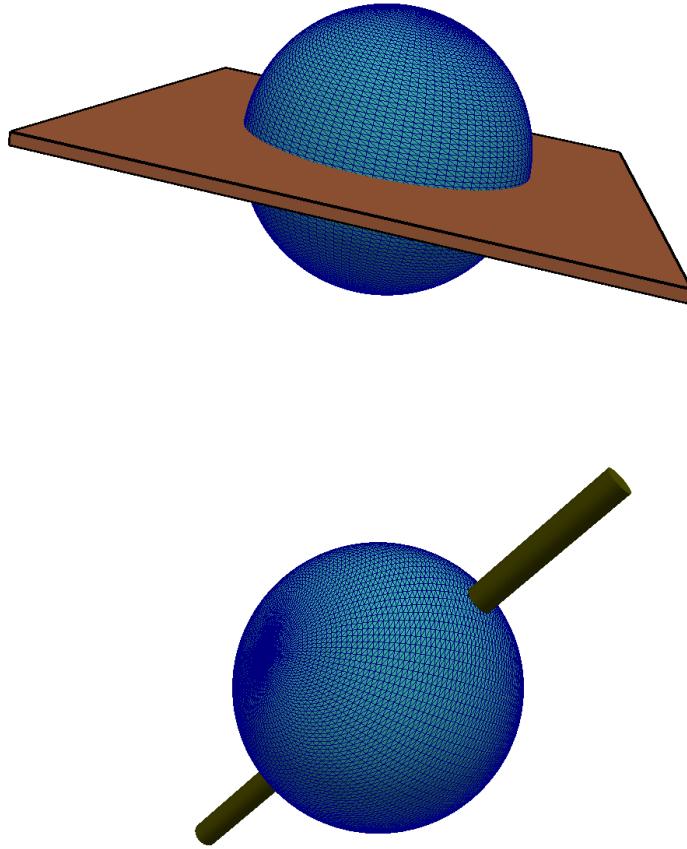


Figure 3.6: Modelling a wall and a filament. A diagnostic spherical surface is intersected by a cylinder and plane.

$$f_{wall}(r) = \frac{Area_{c/s}}{4\pi r^2} = \frac{2\pi r d}{4\pi r^2} \propto r^{-1} \quad (3.4)$$

$$f_{fil}(r) = \frac{Area_{c/s}}{4\pi r^2} = \frac{\text{const.}}{4\pi r^2} \propto r^{-2} \quad (3.5)$$

The fractions of points on the surface of the sphere by multiple number of intersecting sheet-like walls or cylindrical filaments also scale proportional to  $r^{-1}$  and  $r^{-2}$  respectively.

For the diagnostic spheres of different radii, the scaling of multi-streams at the intersections is calculated. By checking the variation in the fraction of area occupied, we associate the number of

streams with wall or halo.

Each of the Mollweide projections in Fig. 3.7 - 3.10 shows projection of the multi-stream field on to the spherical surface, and provide useful insight into the multi-stream structure around a halo. In a Mollweide projections, each filament stemmed from the halo looks as a compact patch. If the physical area of the cross section of the filament remains approximately constant, then the size of the patch on the Mollweide projection would decrease with the increasing radius of the diagnostic sphere. The cross section of a wall with the diagnostic sphere has a well known ‘S’-shape (similar to the ecliptic plane in the galactic coordinates) and the width decreases with the growth of the diagnostic sphere. Both patterns are clearly seen in Fig. 3.7 - 3.10.

### 3.5.2 Voids, filaments and walls around haloes

From the technique described above, we arrive at quantitative thresholds for the different components of the web i.e., all regions where  $n_{str} \geq 3$ . We stress that this method is only a practical tool in arriving at heuristic thresholds of cosmic web structures. The analysis done here are for the simulation box of  $100 h^{-1}$  Mpc,  $128^3$  particles, with refinement factor of 8.

The scaling of fraction of points with 3+ streams is closest to  $r^{-1}$ , where  $r$  is the radius of the diagnostic sphere around the halo (Fig. 3.7 - 3.10; top figures). Since  $r^{-1}$  variation is geometrically identical to that of a wall, it is identified as a flow region with 3+ stream flow. In this simulation the volume fraction of the web is dominated by 3-stream flows:  $f_{vol}(3) \approx 0.04$  while  $\sum_5^\infty f_{vol}(n_{str}) \approx 0.02$ .

The deviation from the exact slope is expected, since assuming the filaments and walls as uniform cylinders and planes is rather crude. In the simulation, the filaments and walls have a far more complicated structure, where they branch out, and do not correspond to regular geometrical shapes. Detailed explanations for deviations are illustrated using Mollweide projections in the next section.

Variation of multi-streams regions of 5+ to 17+ streams steadily changes from  $r^{-1}$  to  $r^{-2}$ . This smooth transition implies that finding an exact cut-off of  $n_{str}$  for a filament is possible neither  $n_{str}$

threshold nor by density. At 17+ stream regions scaling is closest to  $r^{-2}$ , the approximate filamentary geometry. In fact,  $n_{str} = 19+$  regions also exhibit similar variation in spherical projections, but our choice of the threshold based on the lowest  $n_{str}$  value that scales close to  $r^{-2}$ . Again, unlike the threshold for voids and walls, the threshold for filaments and haloes are not unambiguous. Our seemingly arbitrary choice of definition of filaments as regions with 17+ streams (in Section 3.4, Fig. 5.2 and Fig. 3.5) was motivated by this observation. Thus projections on a diagnostic sphere is a convenient tool for classifying regions in the simulation as belonging to void, wall, filament or a halo.

For illustrations, we have picked 4 haloes from different mass ranges:  $3.7 \times 10^{14} M_\odot$ ,  $5.0 \times 10^{13} M_\odot$ ,  $7.0 \times 10^{12} M_\odot$  and  $1.1 \times 10^{12} M_\odot$  from the simulation box of  $100 h^{-1}$  Mpc length and  $128^3$  particles. Multi-stream field with a high refinement factor of 8 is calculated for a greater resolution on scales of the halo volume. Diagnostic spheres of radii  $0.1 h^{-1}$  Mpc to  $5 h^{-1}$  Mpc are drawn for each of these haloes (Fig. 3.7 - 3.10; bottom figures), with the multi-stream field projected onto the surface. In the Mollweide projections of these spheres, the white space refers to single-stream voids. For the largest halo (Fig. 3.7) with FOF radii  $1.2 h^{-1}$  Mpc, the voids already appear in sphere of radius  $1.5 h^{-1}$  Mpc and in the smaller haloes (Fig. 3.10) it appears as early as  $0.5 h^{-1}$  Mpc.

Up to  $1 h^{-1}$  Mpc from halo center of the largest halo, the surfaces are uniformly covered with high number of streams (red, 17+). This shows that the most non-linear regions are close to centres of haloes. A similar trend is seen for the halo of radius  $0.7 h^{-1}$  Mpc (Fig. 3.8). However, for smaller haloes (Fig. 3.9 and 3.10) lower number of streams (even the wall forming 3+ streams; blue) start occupying the spherical surface at radii lesser than FOF-radius. In the case of the smallest halo of  $1.1 \times 10^{12} M_\odot$  mass, the 17+ streams are seen at scales as low as  $0.1 h^{-1}$  Mpc. The distribution of multi-streams on the surface seems do not have a symmetry of any kind, signifying a complex morphology of the web in the vicinity of the haloes. Regions with 5+ to 15+ streams form structures intermediary to filament-like and wall-like behaviour, as seen by scaling of fraction of total points on the space with distance from halo center.

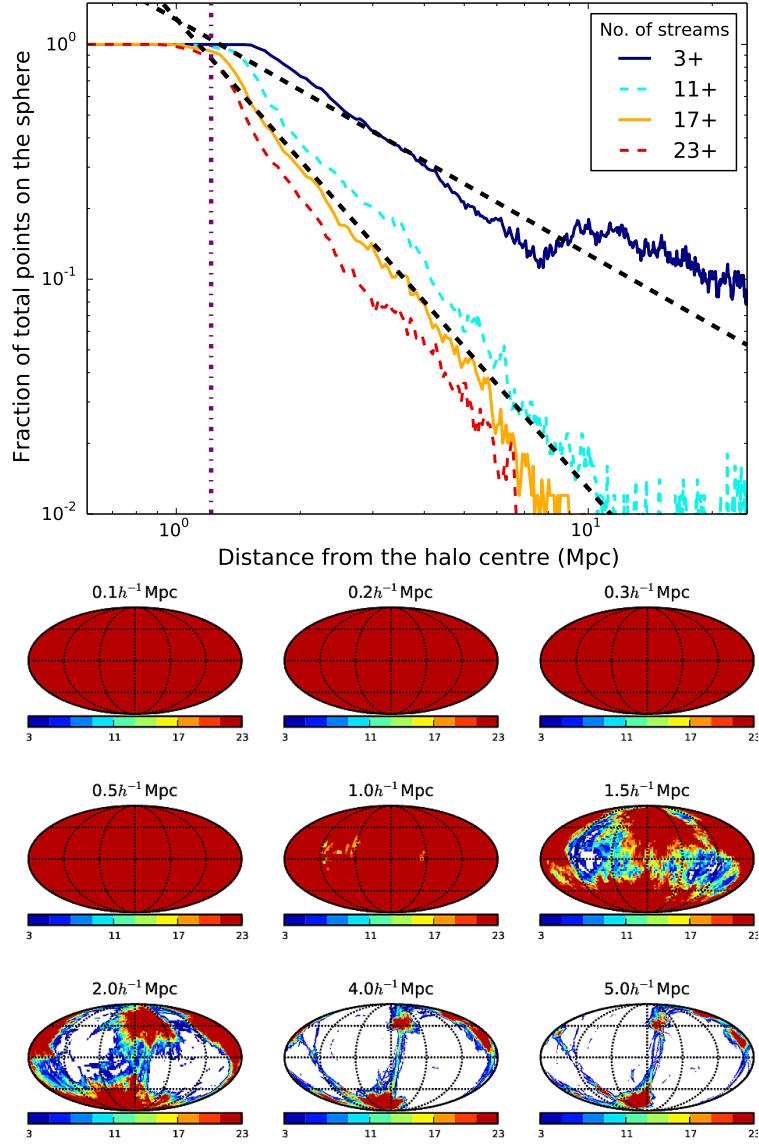


Figure 3.7: Large halo of mass  $3.7 \times 10^{14} M_{\odot}$  and FOF radius  $1.2 \text{ } h^{-1} \text{ Mpc}$  (dotted-violet line in the top figure). Top: Fractional distribution of streams on the surface of spheres of increasing radii. Dashed-black lines are for  $r^{-1}$  and  $r^{-2}$  scaling. 3+ streams are closest to  $r^{-1}$  and 17+ scales close to  $r^{-2}$ . for higher thresholds, the fractional distribution departs smoothly from  $r^{-2}$ . Bottom: Mollweide projection of multi-streams on the surface of the sphere.

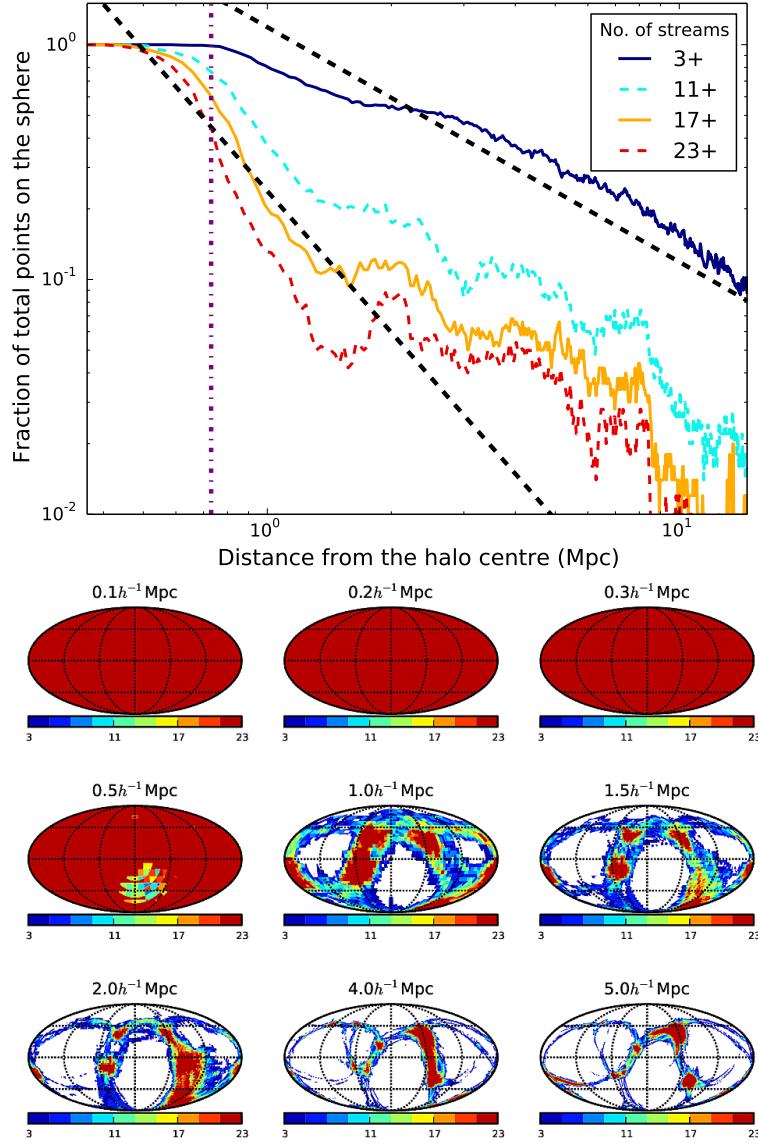


Figure 3.8: Halo of mass  $5.0 \times 10^{13} M_{\odot}$  and FOF radius  $0.7 h^{-1} \text{ Mpc}$ . Top: Fractional distribution of streams deviates from  $r^{-1}$  and  $r^{-2}$  scales since the high stream filament passes along the surface of the sphere. Bottom: Filament passing through the surface is seen from 2 to 5 times the halo radius.

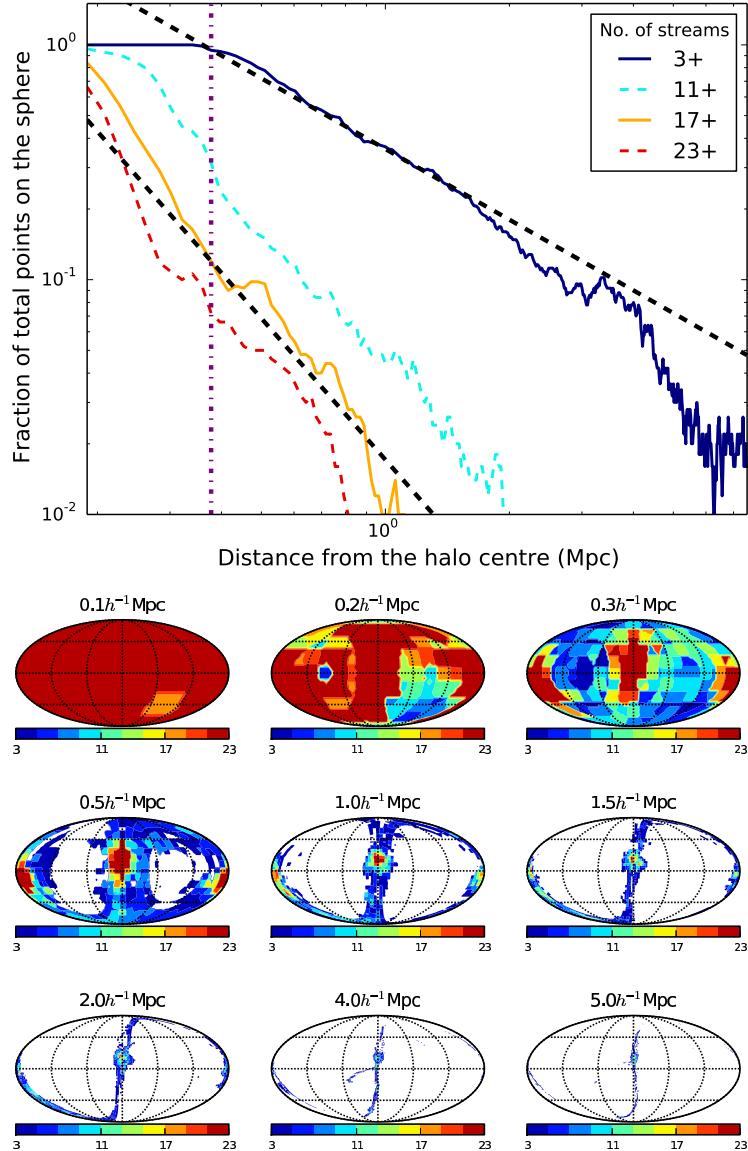


Figure 3.9: Halo of mass  $7.0 \times 10^{12} M_\odot$  and FOF radius  $0.4 h^{-1} \text{ Mpc}$ . Top: All lines clearly scale between  $r^{-1}$  and  $r^{-2}$ . Bottom: The filament is passing through the centre. It persists from radius of halo to  $4 h^{-1} \text{ Mpc}$ . It is also surrounded by a single wall appearing as a line in the middle.

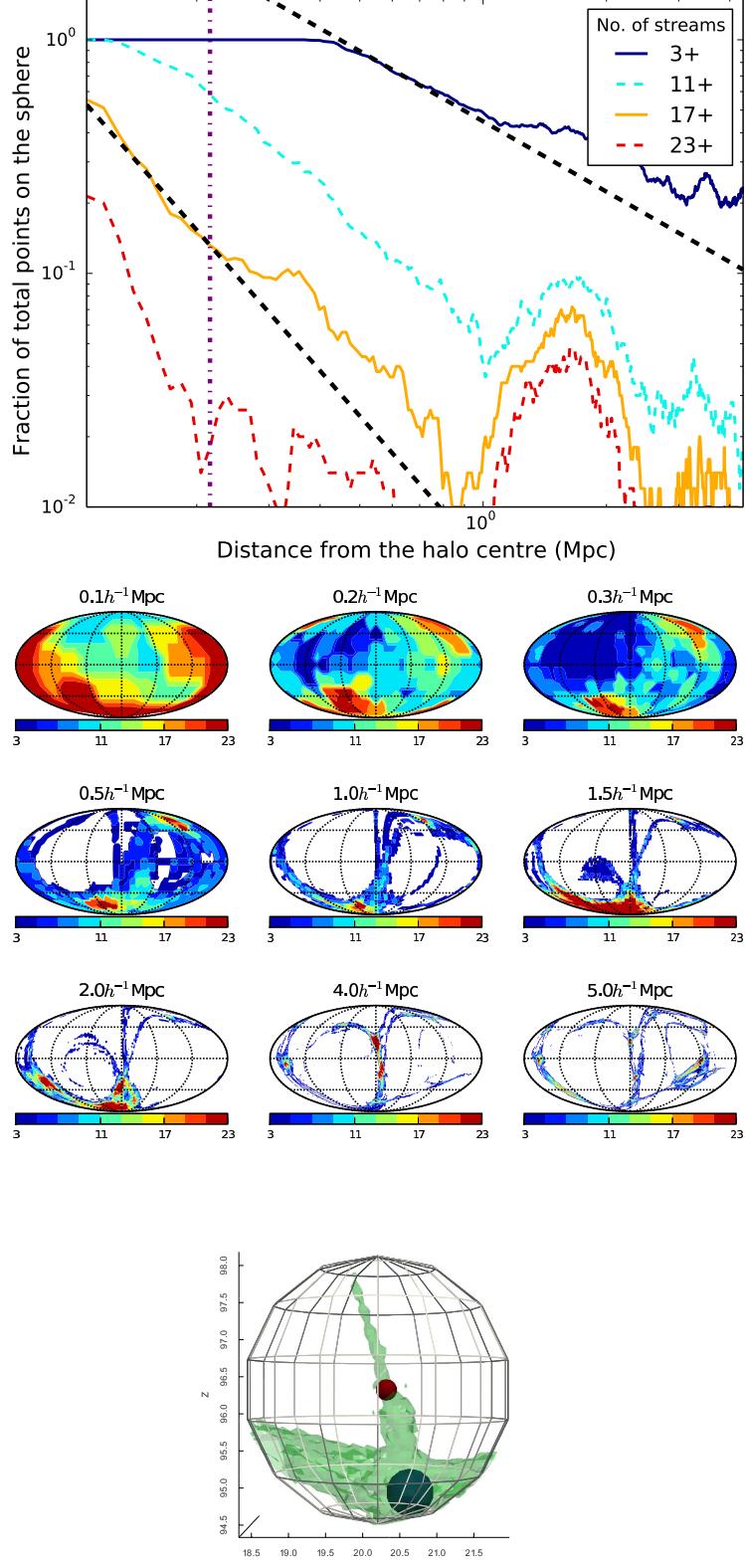


Figure 3.10: Halo of mass  $1.1 \times 10^{12} M_{\odot}$  and FOF radius  $0.2 h^{-1}$  Mpc (dotted-violet line). This halo just has 26 particles, hence the resolution is lesser than previous haloes for surfaces with low radii. Top: There is a bump in fraction of each of 3+ streams a little over  $1 h^{-1}$  Mpc. This is due to the presence of an additional halo nearby, as seen in the projections. The Mollweide projections from  $0.4 h^{-1}$  Mpc to  $2 h^{-1}$  Mpc have high stream flow regions near the lower surfaces of diagnostic spheres. Bottom: Corresponding FOF halo (red, at the centre) has a more massive neighbouring FOF halo (blue) within distance of  $2 h^{-1}$  Mpc. The 17+ stream regions (green) are

Halo environment at distance over twice the FOF radius reveals interesting morphological features. The walls intersect the sphere, and in the Mollweide projections, appear like a thin strip. We also note that a filament oriented tangentially to the diagnostic surface may occasionally appear as a strip too (like in Fig. 3.8, see the corresponding discrepancy in fraction of streams), but upon inspecting the spheres at various radii, we can clearly identify the persisting line-like structures, and they correspond to the walls. Similarly, a filament is projected as a compact patch structure, which occurs due to an intersection of a cylinder-like geometry with the spherical surfaces. It is clearly observed at the distance of  $4 - 5 h^{-1}$  Mpc in Fig. 3.7 and in between  $0.5 - 5.0 h^{-1}$  Mpc in Fig. 3.9.

Hence we conclude that the 3+ stream regions constitute predominantly walls and the regions with 17+ streams correspond mostly to filaments. The higher  $n_{str}$  shells must be surrounded by the layers with lower  $n_{str}$ . Thus, the filaments are within the walls, and do not exist independently. We remind that the radius of diagnostic sphere varies from  $0.1 - 5 h^{-1}$  Mpc, whereas the Mollweide projections shown here are of the same size. Hence the walls and filaments appear more narrow and smaller in larger spheres due to zooming-out effects. In some cases (Fig. 3.8, 3.10), the Mollweide projections display the walls and filaments as a complicated network with patches of high number of streams.

The high peak shown by the curves corresponding to the numbers of streams from 11+ to 17+ in the top panel of Fig. 3.10 is mostly due to the presence of another halo nearby (seen at bottom of Mollweide projection of  $1 - 2 h^{-1}$  Mpc). Fig. 3.8 also has a deviation from usual scaling, and this due to the intricate shape of 17+ stream filament, which appears to be branching out after  $1 h^{-1}$  Mpc.

Generally the transitions from haloes to filaments then to walls and finally to voids appear to be rather smooth. However occasionally sharp features as the one seen in Fig. 3.10 may emerge when the diagnostic sphere hits a neighbouring halo.

The friends-of-friends analysis identifies haloes as spherical structures. The distribution of multi-streams projected onto these surfaces of the FOF haloes can be utilized for a statistical anal-

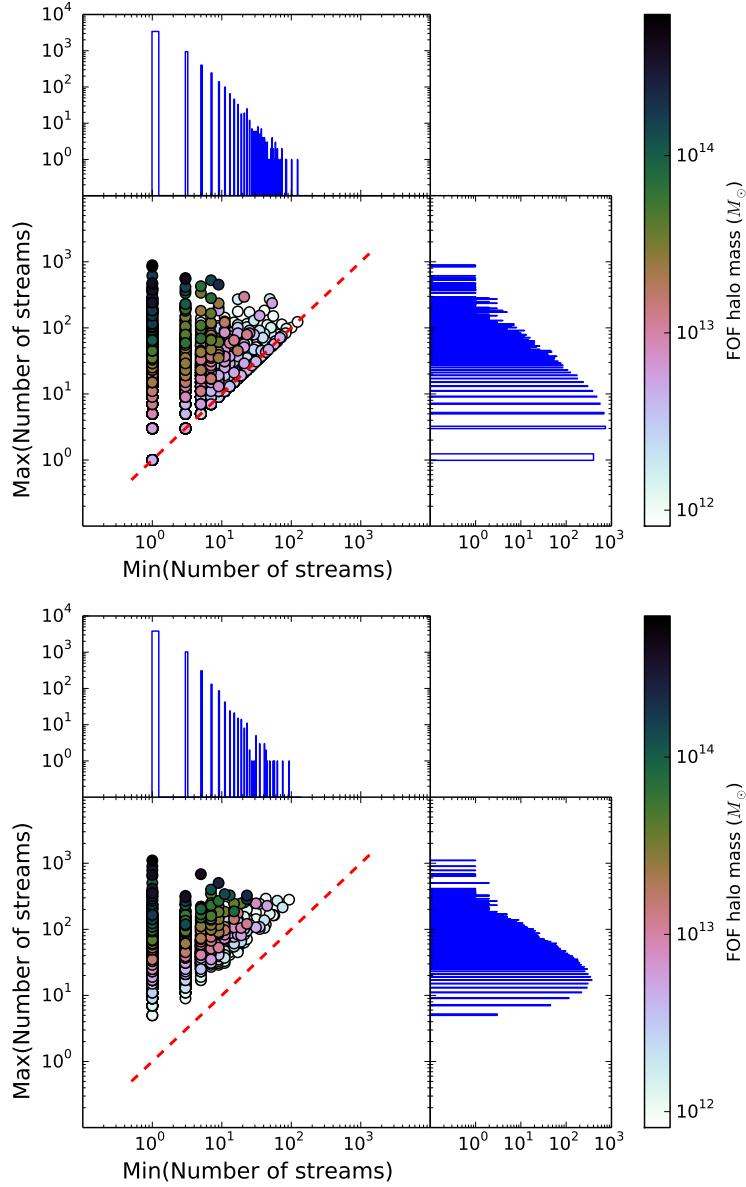


Figure 3.11: Scatter plot of minimum and maximum number of streams on the surface of FOF radius around all haloes in FOF catalogue. For the analysis, we use a total of 5521 haloes that are identified using the FOF technique with linking length,  $b = 0.2$ . Since several of the haloes coincide, distributions of number of haloes corresponding to minimum and maximum  $n_{str}$  on their FOF radii are shown above and beside the scatter plot respectively. Top: Full box of  $L = 100 h^{-1} \text{Mpc}$  and  $N = 128^3$  (i.e.  $L/N = 0.78 h^{-1}$ ), with a low refinement factor of 1 is utilized for multi-stream field calculation. Bottom: Same simulation box; multi-stream field calculated with higher refinement factor of 8.

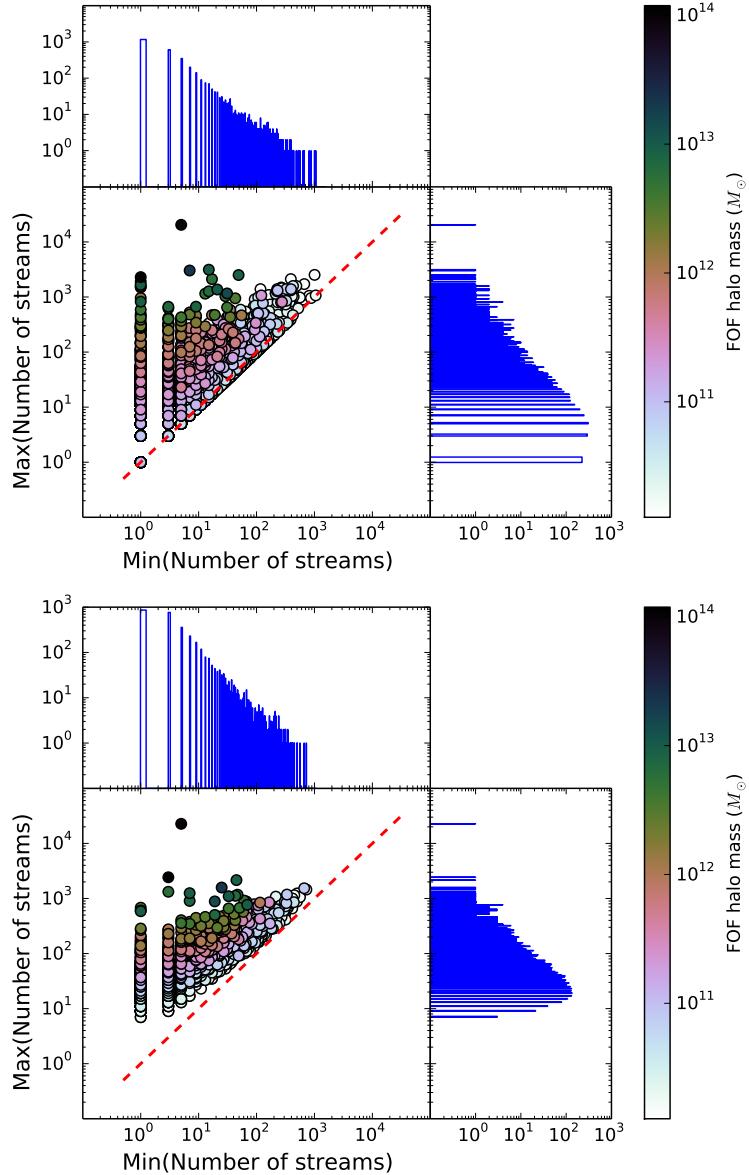


Figure 3.12: Scatter plot of minimum and maximum number of streams on the surface of FOF radius around all haloes in FOF catalogue. Distributions of number of haloes corresponding to minimum and maximum  $n_{str}$  on their FOF radii are shown above and beside the scatter plot respectively. Top: In the simulation box of  $L = 100 h^{-1} \text{ Mpc}$  and  $N = 512^3$  (i.e.  $L/N = 0.19 h^{-1}$ ), multi-stream field is calculated on a smaller slice of  $25 h^{-1} \text{ Mpc}$  with a refinement factor of 1. Multi-stream field is projected onto surfaces of 3448 FOF haloes within this small box. Bottom: Same simulation box; the multi-stream field calculated on the the same small box, but with a refinement factor of 8.

ysis of the haloes (Fig. 3.11 and 3.12). We have utilised FOF catalogues of haloes more than 20 particles found at linking length  $b = 0.2$ . The  $n_{str}$  ranging from as low as 1 to higher than  $10^3$  are seen on the halo surfaces. Haloes which have the minimum  $n_{str} = 1$  are in contact with the void. However, if the maximum  $n_{str}$  is also 1, then the halo is completely within the void. In calculations with low refinement factor of 1, only 7.3% (for  $L/N = 0.78 h^{-1}$ ) and 6.5% (for  $L/N = 0.19 h^{-1}$ ) of the haloes are completely within single streaming voids. This is solely due to low resolution of multi-stream field, since none of the FOF haloes are found in high resolution multi-stream calculations on both  $L/N = 0.78 h^{-1}$  and  $L/N = 0.19 h^{-1}$  boxes. At high refinement factors, none of the haloes are entirely embedded in a region with just one multi-stream value (i.e.,  $\max(n_{str}) = \min(n_{str})$ , along the dotted-red lines in Fig. 3.11 and 3.12). However, there are significant number of haloes whose FOF surfaces are in contact with the void region: in calculations with refinement factor of 8, 62% of the haloes in  $L/N = 0.78 h^{-1}$  and 34% of the haloes in  $L/N = 0.19 h^{-1}$  are in contact with void on their FOF radii. Rest of the haloes are completely within non-void regions.

Statistical analysis of FOF haloes in Fig. 3.11 and 3.12 show that massive haloes tend to have low  $\min(n_{str})$  and high  $\max(n_{str})$ , hence a very diverse multi-stream environment on their spherical surface. The heuristic multi-stream threshold for haloes mentioned in Section 3.4 results in virialized haloes with a constant  $n_{str}$  value. These halo surfaces far from sphere (see Fig. 3.5), whereas, the FOF surfaces are spherical and have a large range of number of streams on their surfaces. The probability distribution function of number of FOF haloes has an approximately exponential tail monotonically decreasing with  $\min(n_{str})$ .

## 3.6 Summary

In this paper, we explore for the first time, the multi-stream environment of dark matter haloes in cosmological N-body simulations. The visualization and quantitative characterization of generic non-linear fields in three-dimensional space represent a serious challenge from both conceptual and computational points of view. The complexity of the problem requires diverse tools for analysing

the results of cosmological simulations as well as galaxy catalogues.

This study is different from the most previous works in a few aspects. Firstly, we consider the representation of the cosmic web in the form of a multi-stream field rather than a density field. The multi-stream field contains a different information about the web than the density and velocity fields and thus represents a complimentary characterization of the web revealing new dynamical features of the web (Shandarin 2011, Shandarin et al. 2012, Abel et al. 2012). Secondly, for computing the multi-stream field we use the tessellation of three-dimensional Lagrangian sub-manifold  $\mathbf{x} = \mathbf{x}(\mathbf{q}, z = 0)$  in six-dimensional  $(\mathbf{x}, \mathbf{q})$  space which allows to significantly increase the spatial resolution (Shandarin et al. 2012, Abel et al. 2012). The Lagrangian sub-manifold is more convenient since  $\mathbf{x}$  is a single-valued function of  $\mathbf{q}$  at any stage including a highly non-linear regime while the phase space sheet projected on  $\mathbf{x}$ - or  $\mathbf{v}$ -space is not. If the initial state of the simulation is represented by a uniform three-dimensional mesh, then storing the Lagrangian sub-manifold does not require additional space for Lagrangian coordinates. And thirdly, in the study of the multi-stream environment of dark matter haloes we use the Mollweide projection of the multi-stream field computed on a set of diagnostic spherical surfaces centred at the FOF haloes and having radii from  $0.1 h^{-1}\text{Mpc}$  to  $5 h^{-1}\text{Mpc}$ .

Most of the results are obtained for a simulation in  $L = 100h^{-1}\text{Mpc}$  box with  $N = 128$  particles along each axes although we report some of the results for the simulations in  $100 h^{-1}\text{Mpc}$  box with  $256^3$  and  $512^3$  particles as well as in  $200 h^{-1}\text{Mpc}$  box with  $128^3$ ,  $256^3$  and  $512^3$  particles.

Using the tessellation of the three-dimensional Lagrangian sub-manifold  $\mathbf{x} = \mathbf{x}(\mathbf{q}, t)$  (Shandarin et al., 2012), we compute the multi-stream field i.e. the number of streams on a regular grid in the configuration  $\mathbf{x}$ -space,  $n_{str}(\mathbf{x})$  for estimating global parameters or on selected set of points in the study of the haloes environments.

The multi-stream field takes odd whole numbers everywhere except at a set of points of measure zero where it takes positive even whole numbers. This property is very useful for debugging the numerical code.

The multi-stream field allows one to define physical voids as the regions with  $n_{str} = 1$ . The

rest of space with  $n_{str} \geq 3$  can be called the non-void or web. This division of the space into two parts is unique and physically motivated: no object can form before shell crossing happens. It is worth emphasizing that the division of space into voids and web is based on the local parameter, the number of streams at a single point (Shandarin et al. 2012, Abel et al. 2012). Falck et al. (2012) and Falck & Neyrinck (2015) defined the web as a set of particles that experienced the 'flip-flop' at least once along any axis. We discussed potential problems with this definition in the beginning of Section 3.4.

The further division of the web into walls, filaments and haloes is not straightforward although haloes can be defined using dynamical parameters related to the requirement of virialization of haloes. One of the simplest is the famous density threshold  $\rho/\bar{\rho} \approx 200$ . Identifying filaments and walls is significantly more tricky (see e.g. Hahn et al. 2007, Forero-Romero et al. 2009, Aragon-Calvo et al. 2010, Cautun et al. 2014, Falck & Neyrinck 2015) and require non-local parameters.

The large part of walls can also be identified locally since the regions where  $n_{str} = 3$  can be neither filaments nor haloes. For instance, in the simulation  $100 h^{-1}\text{Mpc}$  box with  $128^3$  particles the web occupies about 6% of the volume, the three-stream flow regions occupy about 4% and the rest of the web remaining 2% of the total volume.

In this study we introduced an empirical statistical criteria which very crudely distinguish wall, filament and haloes. We have found empirically that in the studied simulation the transition from wall points to filament points takes place approximately at  $5 \leq n_{str} \lesssim 15$ . Using the virial over-density threshold of 200 in Eq. 3.3, we have also estimated that the haloes correspond to the regions with  $n_{str} \gtrsim 90$ . Thus, the transition from filament to haloes takes place in the range  $17 \leq n_{str} \lesssim 90$ . The above critical values for transition from walls to filaments and from filaments to haloes were shown to be approximately correct for the simulation with  $L/N = 0.78h^{-1} \text{ Mpc}$ . This technique can be also applied to simulations with different  $L/N$  ratio and multi-stream grid of different refinement factors but the classification based on the threshold applied locally will remain only a very crude estimator. A more sophisticated morphological analysis will require non-local geometric and topological methods, which is beyond the scope of this paper.

Table 3.4: Comparison of the volume and mass fraction of the elements of the cosmic web between our analysis and Falck & Neyrinck (2015).  $L/N = 0.78h^{-1}$  Mpc for simulations used in both techniques. We use a refinement factor of 8 for the multi-stream grid. The mean density is given in units of the mean density of the universe.

	Multi-stream analysis (This work)				ORIGAMI (Falck & Neyrinck, 2015)			
	Voids	Walls	Filaments	Haloes	Voids	Walls	Filaments	Haloes
Volume Fraction (%)	93	7	< 1	< 0.1	84	12	3	< 1
Mass Fraction (%)	32	35	17	14	26	19	19	35
Mean density	0.34	5	> 17	> 140	0.31	2.2	6	> 35

We have found that the volume and mass fractions in the voids are approximately V.F./M.F. = 96/76, 93/32 and 88/24, where each number is the percentage, for the simulations with  $L/N = 1.56h^{-1}$ ,  $0.78h^{-1}$  and  $0.19h^{-1}$  Mpc respectively. As the ratio  $L/N$  gets smaller both volume and mass fractions in voids monotonically decrease. This is fairly consistent with the results of Falck & Neyrinck (2015), considering the differences between our numerical methods. We compare the fractions of the volumes and masses in other components of the web in Table 3.4, and the results are in a good qualitative agreement. Our estimates systematically higher for both volume and mass fractions for voids and thus systematically lower for the web. One general conclusion seems to be obvious: the web defined by the multi-stream flows is about  $(100\% - 84\%)/(100\%-93\%)=2.3$  times thinner than that defined by the ORIGAMI method.

In conclusion we would like to outline the major aspects of the web revealed by the study of the multi-stream field. The multi-stream field is a fundamental attribute of the structures formed in cold collision-less dark matter. Its properties are of great importance for the detecting dark matter directly in a laboratory setting or indirectly via astronomical observations. The dark matter web described by a multi-stream field represents a nested structure, consisting of layers with increasing number of streams. The number of streams are odd integers almost everywhere except on caustics where they are even integers. The caustics occupy infinitesimal volume. The most of the volume is occupied by one stream flow regions which are dark matter voids. The regions with three streams are the regions occupying the second largest volume. They form very thin membrane type structures (often referred to as walls or pancakes) most of which are connected in one huge connected

formation. The three-stream regions form the external shell of the web. All other structures filaments and haloes are within the three-stream shell. The membranes are attached to each others by the filaments which locally consist of regions with higher number of streams than the neighbouring membranes. The filaments form the framework or a skeleton of the dark matter web. Similar to a real skeleton, the filamental structure has joints where most of the the dark matter haloes are located. The haloes are the local peaks of the multi-stream field.

## Acknowledgements

The authors are grateful to Gustavo Yepes and Noam Libeskind for providing the simulations used in this study and the Sir John Templeton Foundation for the support. The authors would also like to thank the anonymous referee for careful review and useful suggestions.

# Chapter 4

## Topology and geometry of the dark matter web: a multistream view

### 4.1 Introduction

Large scale structures with highly anisotropic shapes were first theoretically predicted by Zeldovich approximation (hereafter ZA) (Zeldovich, 1970). The model based on ZA suggested that the eigenvalues of the deformation tensor dictate the shapes of the *collapsed* structures at the beginning non-linear stage of gravitational instability (Arnold et al. 1982, see also Shandarin & Zeldovich 1989 and Hidding et al. 2014). These structures were found to be crudely characterised as two-, one- and zero- dimensional which actually meant that three characteristic scales of each structure ( $L_1 \geq L_2 \geq L_3$ ) are approximately related as  $L_1^{(p)} \approx L_2^{(p)} \gg L_3^{(p)}$  or  $L_1^{(f)} \gg L_2^{(f)} \approx L_3^{(f)}$  or  $L_1^{(h)} \approx L_2^{(h)} \approx L_3^{(h)}$  respectively. In addition it implied that  $L_1^{(p)} \approx L_1^{(f)}$  and  $L_3^{(p)} \approx L_2^{(f)} \approx L_1^{(h)}$ .<sup>1</sup> At present these generic types of structures are referred to as walls/pancakes/sheets/membranes, filaments and haloes. Although the accuracy of the Zeldovich approximation deteriorates from pancakes to filaments and especially to halos on qualitative level there are no more types of structures. Altogether these structures contain the most of mass in the universe nevertheless they occupy

---

<sup>1</sup>The multi-scale character of the cosmic web was not discussed until 1990s.

very little space. The most of space is almost empty and is referred to as voids.

Klypin & Shandarin (1983) (firstly reported in Shandarin 1983a) were the first to identify a ‘three dimensional web structure’ in the N-body simulation of the hot dark matter scenario. The simulation with  $32^3$  particles used Cloud-in-Cell (CIC) technique on equal mesh revealed that the gravitationally bound clumps of mass – haloes in the present-day terminology – were linked by the web of filamentary enhancements of density which spanned throughout the entire simulation box with the side of about  $150h^{-1}\text{Mpc}$  in co-moving space. In addition Klypin & Shandarin (1983) suggested that pancakes must be considerably less dense than the filaments since they were not detected in the simulation. These results were quickly confirmed by Centrella & Melott (1983) and Frenk et al. (1983). In addition Centrella & Melott (1983) who ran the simulation on similar mesh but with 27 times more particles also detected pancakes at  $\rho/\bar{\rho} = 2$  level. At present this picture is widely accepted, and is referred to as the ‘cosmic web’ (Bond et al. 1996 and van de Weygaert & Bond 2008b).

Galactic distributions in redshift surveys have also revealed distinct geometries and topologies of the cosmic web. One of the first indications of the connection of the clusters of galaxies by filaments was demonstrated by Gregory & Thompson (1978) who discovered a conspicuous chain of galaxies between Coma and A1367 clusters using a sample of 238 galaxies. Later this result was confirmed by de Lapparent et al. (1986) who used a significantly greater redshift catalogue of 1100 galaxies of the same region. Zeldovich et al. (1982) compared the percolation properties of the redshift catalogue of 866 local galaxies provided by J. Huchra with three theoretical distribution of particle in space: a Poisson distribution, the hierarchical model by Soneira & Peebles (1978) and the particle distribution obtained from N-body simulation by Klypin & Shandarin (1983). They found that the both the galaxy sample and the density field obtained in N-body simulation percolated at considerably smaller filling factors than the Poisson distribution. On the other hand the hierarchical model percolated at higher filling factors than the Poisson distribution. Further studies confirmed that the galaxies and the particles in the hot dark matter model are arranged in the web-like structures Zeldovich et al. (1982), Shandarin (1983a), Shandarin & Zeldovich (1983),

Shandarin & Zeldovich (1984). This result was confirmed in more detailed analysis by Einasto et al. (1984). Melott et al. (1983) also found similar percolation properties in the mass distribution in the N-body simulation of a CDM model.

Thus by the early 1990s it was clearly demonstrated that the web like structure is a generic type for a wide range of initial conditions in both two-(Melott & Shandarin 1990, Beacom et al. 1991) and three- dimensional (Melott & Shandarin, 1993) cosmological N-body simulations. However it also was demonstrated that the quantitative parameters of the web structures depend on the initial power spectrum. Remarkably the simulations also showed that adding small scale perturbations does not ruin the large scale structures if the slope of the power spectrum is negative in both two- and three- dimensional simulations.

All aspects of these studies have been experiencing great advancements in three decades passed since the discovery and first studies of the geometry and topology of the large-scale structures. The galaxy redshift catalogues have grown by thousands of times (by surveys such as Sloan Digital Sky Survey (SDSS) Tegmark et al. 2003 and Albareti et al. 2016 and the 2MASS Redshift Survey Huchra et al. 2012), the sizes of cosmological N-body simulations (modern large scale simulations like Millennium Springel et al. 2005 and Q-Continuum Heitmann et al. 2015) by more than a million times. The number of various methods for identifying structures has also grown practically from one method<sup>2</sup> to several dozens (Colberg et al. 2008, Knebe et al. 2011, Onions et al. 2012, Knebe et al. 2013 and references therein). Measuring or quantifying the structures always has been a difficult problem and many sophisticated techniques both mathematically and computationally have been proposed and investigated (see reviews by van de Weygaert & Bond 2008b, van de Weygaert & Bond 2008a).

Cosmic web structures have been characterized using several geometrical and topological indicators such as genus curves (Gott et al. (1986)). In an attempt to characterize the shapes of individual regions in the excursion sets of the density field, Sahni et al. (1998) suggested to use

---

<sup>2</sup>FOF was used for the topological studies via percolation technique and identifying super clusters of galaxies (Zeldovich et al. 1982, Shandarin 1983a, Shandarin & Zeldovich 1983 on the one hand and for identifying halos Davis et al. 1985 on the other.

partial Minkowski functionals. They developed the method labelled SURFGEN and applied it to CIC density field obtained in N-body simulations (Sathyaprakash et al. 1998, Sheth et al. 2003, Shandarin et al. 2004). Aragon-Calvo et al. (2007) have developed the multi-scale MMF (Multi-scale Morphology Filter) detection technique based on the signs of three eigenvalues of the Hessian computed for a set of replicas of the density field filtered on different scales. Similar multi-scale approaches to identifying structures is adopted in NEXUS and its extensions to velocity shear, divergence, and tidal fields Cautun et al. (2013). More recently, persistence and Morse-Smale complexes in the density fields are analysed by Sousbie (2011), ? and Shivshankar et al. (2015) to detect multi-scale morphology of the cosmic web.

There is also an increasing interest in the measures for detecting filaments in large astronomical surveys. Topology in the large scale structure was analysed by Betti Numbers for Gaussian fields (Park et al., 2013) and SDSS-III Baryon Oscillation Spectroscopic Survey (Parihar et al., 2014). Sousbie et al. (2008) detected skeleton of filaments of the SDSS and compared to the corresponding galaxy distribution. In smoothed density of mock galaxy distribution, Bond et al. (2010a) studied the projection of eigenvalues. The Hessian eigenvector corresponding to the largest eigenvalue is used by Bond et al. (2010b) to trace individual filaments in N-body simulations and the SDSS redshift survey data. Majority of the above analyses, however, ignore the dynamical information from the velocity field.

On the other hand, detection of voids and study of their morphological properties are done via numerous methods too. Traditional detection of void regions using just the particle coordinates differ based on the various methods used to identify them (see comparison of void finders in Colberg et al. 2008 and references therein). Some methods involve using under-density thresholds. Blumenthal et al. (1992) proposed that the mean density in voids is  $\delta = -0.8$  by applying linear theory argument. Similar threshold was used by Colberg et al. (2005) to identify voids. Under-dense excursion set approach was used by Shandarin et al. (2006) to identify percolating voids. ? used the excursion set formalism to develop an analytical model for the distribution voids in hierarchical structure formation (also see the excursion set approaches applied to voids by Paranjape et al. 2012,

Jennings et al. 2013 and Achitouv et al. 2015). Voids are also detected by isolating regions around local minima of density fields. For instance, the watershed transform is used by WVF-Platen et al. (2007), ZOBOV-Neyrinck (2008) and VIDE-Sutter et al. (2015) for segmentation of under-dense regions.

The unfiltered density field was generated using DTFE-Delaunay Tessellation Field Estimator (Schaap & van de Weygaert 2000, van de Weygaert & Schaap 2009 and Cautun & van de Weygaert 2011) by applying it to the particle coordinates. Earlier it was shown that DTFE is superior to CIC techniques (Schaap 2007 and van de Weygaert & Schaap 2009) in generation of the density field with high spatial resolution. In a new approach to the analysis of the shapes of the large-scale structures, Sousbie (2011) introduced DIScrete Persistent Structure Extractor (DisPerSE) based on Morse-smale complex. By implementing it on realistic cosmological simulations and observed redshift catalogues Sousbie et al. (2011) found that DisPerSE traces very well the observed filaments, walls and voids.

An additional dimension to the scope of the structure shapes is related to the question whether the density distribution (regardless of its form: continuous or discrete) is the only physical diagnostic of the cosmic web shapes or not. If not, then whether it is the best of all or not. And even if it is the best, then whether the other fields or distributions can provide a valuable contribution to understanding the shapes of the cosmic web or not. The answer to the latter question seems to be positive. In fact there are examples of attempts to bring new players into the field. For instance Hahn et al. (2007) and Forero-Romero et al. (2009) studied the relation between the geometry of structures and the Hessian of the gravitational potential. Shandarin (2011) demonstrated that the study of the multistream field reveals some features of the structures that cannot be easily seen in the density field. This has become even more evident when Shandarin et al. (2012) and ? showed that the full dynamical information in the form of three-dimensional sub-manifold in six-dimensional phase space can be easily obtained from the initial and final coordinates of the particles in DM simulations. Hahn et al. (2015) showed that this method provides extremely accurate estimates of the cosmic velocity fields and its derivatives. It has been shown that the

multistream field provides a physical definition of voids in N-body DM simulations by the local condition  $n_{str} = 1$  (Shandarin et al. 2012 and Ramachandra & Shandarin 2015). Falck et al. (2012) proposed the ORIGAMI method of assigning particles to structures based on the number of axes along which particle crossing has occurred. Void, wall, filament, and halo particles are particles that have been crossed along 0, 1, 2, and 3 orthogonal axes, respectively. Shandarin & Medvedev (2017) identify the void particles as the ones that do not undergo any *flip-flop* through the evolution. Each of above definitions completely independent of any free parameters, with small differences in the physical implication.

Tracing the Lagrangian sub-manifold also provides rich insights into caustics (Arnold et al. (1982) and Hidding et al. 2014) and halo collapse ?. Recently, there are attempts to improve N-body simulations (see Hahn et al. (2013), Angulo et al. (2013c), Angulo et al. (2013a), Sousbie & Colombi (2015) and Hahn & Angulo (2016)) by solving the Vlasav-Poisson equation using tessellations in the Lagrangian sub-manifold. Galaxy evolution and star formation in the context of multi streaming phenomenon are studied by ?.

Despite the considerable improvements in simulating, identifying and measuring the cosmic web – briefly discussed above – many aspects remain unsettled and are vigorously debated. The intention of this work is to further investigate the strengths and weaknesses of the multistream field as a complimentary diagnostic of the shapes in the DM web. Multi-stream filed is simply the number of DM streams at every point of Eulerian space. Thus it is an odd positive integer at a given point (Arnold et al. 1982, see also Shandarin & Zeldovich 1989 and Hidding et al. 2014). We estimated it on a regular mesh of a chosen resolution from the tessellation of of the simulation particles in Lagrangian space and the particle coordinates at a chosen time Shandarin et al. (2012). The external boundaries of the cold DM web are the caustics in the density field which are clearly seen in the simulations with adequate resolution of the density field (see e.g. Fig 7 in Hahn et al. (2015)). However the exactly same boundaries of the DM web can be identified as the boundaries of a single-stream flow which is a local parameter. The multistream field even a better indicator of the boundaries of the DM web than caustics because caustics are present everywhere the number

of streams varies (from 1 to 3, from 3 to 5, etc) but the boundary of the web are only the one where the number of stream changes from 1 to 3.

In particular we would like to discuss the differences in defining voids in density and multistream fields. It is closely related to the definition and distinguishing of linear and non-linear structures or regimes. One simple statistical definition that often used is as follows: after defining the std of the density contrast  $\sigma_\delta \equiv <(\rho(x)/\bar{\rho} - 1)^2>^{1/2}$  one can roughly separate the linear and non-linear regimes by the boundary  $\sigma_\delta = 1$ . This is obviously very crude characteristic which does not say much about the geometry and topology of the non-linear structures. The parameter  $\sigma_\delta$  is frequently evaluated for filtered fields  $\sigma_\delta = \sigma_\delta(R_f)$ . Unfortunately the transition from ‘non-linear’ field at small  $R_f$  to ‘linear’ field at large  $R_f$  is smooth and thus choosing a particular value of  $R_f$  is remarkably subjective.

A related but different question is how to select individual non-linear structure, like halos, filaments and walls by using a local parameter. In particular the density threshold has been used on numerous occasions especially for identifying halos and voids. As a rule the choices of particular values have not been justified by solid physical evidences. The virial mass and virial radius of a halo are often used as direct indicators of gravitationally bound objects but they are determined by a nonlocal quantity – the mean overdensity of the halo. An interesting compaison of several kinds of boundaries of halos was provided by More et al. (2015). In particular they considered the virial radius  $R_{\text{vir}}$ ,  $R_{200m}$ , the splashback radius  $R_{\text{sp}}$ , and  $R_{\text{infall}}$ . The splashback radius is defined as an average distance from the center of the halo to the most external caustic if it was resolved. The authors argue that it is “a more physical halo boundary choice” than “commonly defined to enclose a density contrast  $\Delta_{m,c}$  relative to a reference (mean or critical) density. This is the boundary where the number of streams falls from three to one in the multistream field.

Gravitationally bound structures could be defined as linear in the sense that  $\delta(\mathbf{x}) \ll 1$  for all points in the structure. A simple example is a progenitor of large halo at linear stage. However one cannot accurately identify such an object at linear stage using a local criterion like a density threshold. Even at the nonlinear stage of N-body simulation one cannot predict when a particular

Table 4.1: Parameters for the simulation boxes: Side length  $L$ , number of particles  $N_p$ , mass of each particle  $m_p$ , and the gravitational softening length  $\epsilon$  for the GADGET simulations are shown.

$L$	$N_p$	$m_p$	$\epsilon$
$100h^{-1}Mpc$	$128^3$	$3.65 \times 10^{10}h^{-1}M_\odot$	$20h^{-1}kpc$
$100h^{-1}Mpc$	$256^3$	$4.57 \times 10^9h^{-1}M_\odot$	$10h^{-1}kpc$

fluid element with a given value of  $\delta$  in a void will be accreted to a wall or filament. Among other factors the size of the void and proximity to a wall would play significant roles. In addition the walls accrete expanding fluid elements as well thus the velocity divergence on the fluid element would not help.

The rest of the paper is organised as follows: we describe the cosmological simulations in Section 5.3. Some of the important features of the multistream field are described in Section 4.2.1. Topology of the single-streaming voids is discussed in 4.3 and that of the multistream structure is investigated using percolation theory in Section 4.4. . Discussion of the local geometry of multistream field using Hessian matrices is done in Section 4.5.

## 4.2 The simulation

In this analysis, we use cosmological N-body simulations generated by the tree-PM code GADGET-2 (Springel 2005 and Springel et al. 2001b). The periodic side lengths  $L$ , number of particles  $N_p$ , masses of each particle  $m_p$  and the gravitational softening length  $\epsilon$  for the two simulations are tabulated in Table 4.1. Initial conditions at redshift of  $z_{ini} = 80$  are generated by MUSIC (Hahn & Abel, 2011) with the transfer function from Eisenstein & Hu (1998). We adopt the  $\Lambda$ CDM cosmological model with cosmological parameters  $\Omega_m = 0.276$ ,  $\Omega_\Lambda = 0.724$ , the Hubble parameter,  $h = 0.703$ , the power spectrum normalization,  $\sigma_8 = 0.811$  and the spectral index  $n_s = 0.961$ .

### 4.2.1 Multi-stream field at $z = 0$

The multistream field objectively characterizes the level of non-linearity in the cosmic web. The ‘number-of-streams’ field or  $n_{str}(\mathbf{x})$  is computed from the Lagrangian sub-manifold  $\mathbf{x}(\mathbf{q})$ , which is

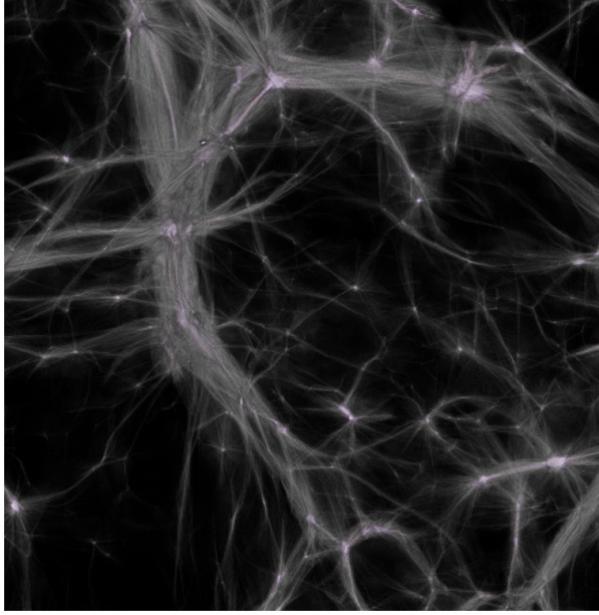


Figure 4.1: 3D rendering of the multistream field: the cosmic web structure of a  $50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc}$  slice in a simulation box of side length  $100h^{-1}\text{ Mpc}$  and  $128^3$  particles. The multistream field is calculated at 8 times the native resolution. void(black) is a percolating structure with  $n_{str} = 1$ . Regions  $n_{str} \geq 17$  show a filamentary structure (gray) and the bright spots at the intersections of the filaments are regions with  $n_{str} \geq 100$ .

a continuous three-dimensional sheet in a six-dimensional  $(\mathbf{q}, \mathbf{x})$  space. In this paper, we utilize the tessellation implementation by Shandarin et al. (2012) to calculate the multistream flow field on the GADGET-2 snapshot at  $z = 0$ . This implementation only requires initial and final coordinates of the dark matter particles.

The  $n_{str}(\mathbf{x})$  values are mostly odd-numbered since each folding in the Lagrangian sub-manifold results in an increase of  $n_{str}$  by 2. Exception to this are only at caustics - which have volume measure zero, then the  $n_{str}$  is even-valued number. The particles in  $n_{str} = 1$  have not experienced orbit crossings and thus these regions are unambiguously identified as void (Shandarin et al., 2012). Foldings in the Lagrangian sub-manifold generally occur one-by-one. For example, a contour of  $n_{str} = 7$  will be within a region of  $n_{str} \leq 5$ . Hence the multistream field commonly has nesting shells, i.e.,  $3 \supseteq 5 \supseteq 7 \supseteq 9 \supseteq 11 \dots$ . Some of the important features of the multistream field are discussed in Appendix B.

The first non-linear DM structures that reach non-perturbative stage of gravitational evolution

have  $n_{str} = 3$ . By visual inspection, these regions generally form a fabric-like open structures that resemble walls. N-body simulations suggest that a DM fluid element after the first crossing of a caustic never returns in a single-streaming state. Therefore the *local* condition  $n_{str}(\mathbf{r}_{\text{f.e.}}) \geq 3$  (where  $\mathbf{r}_{\text{f.e.}}$  is the position of the fluid element) is sufficient for the fluid element to be bound to the DM web.

All particles that have fallen into a wall will never return to any single-streaming regions, therefore they can be labeled as gravitationally bound to pancakes/walls. The surface contours of higher  $n_{str}$  are embedded within the walls. Figure 5.2 shows a filamentary structure of the multistream web at  $n_{str} \geq 17$ . The figure also shows regions around local maxima of the multistream field, which are generally located at the intersections of filaments.

The multistream field can be computed at arbitrary resolutions of diagnostic grids. The parameter ‘refinement factor’ denotes the ratio of separation of the particles in Lagrangian grid,  $l_l$ , to side length of diagnostic grid  $l_d$ . In a simulation of  $128^3$  particles, for instance, multistream field computed on a diagnostic grid of size  $256^3$  would have a refinement factor of  $l_l/l_d = 2$ .

## 4.3 Voids in the multistream field

Gravitational instability results in movement of the collision-less fluid particles in the Universe from voids to walls, walls to filaments, and filaments to haloes. As we mentioned above in the multistream portrait, the entry of mass particles from single-streaming regions into  $n_{str} > 1$  region is irreversible. The converse is obviously not true, that is, the particles in  $n_{str} = 1$  regions may move to multistreaming region at a later time in the evolution. At a given cosmic time, sufficient condition for dark matter particles to be bound to non-perturbative and non-linear structures like walls/filaments/haloes is being in multistream regions. Therefore, a single-stream flow implies that gravitationally bound structures haven’t yet formed, and thus defined as a void region. This definition of void is unambiguous and physically motivated, as demonstrated by Shandarin et al. (2012). It is worth stressing that while the density in voids varies, the number-of-streams is uniformly

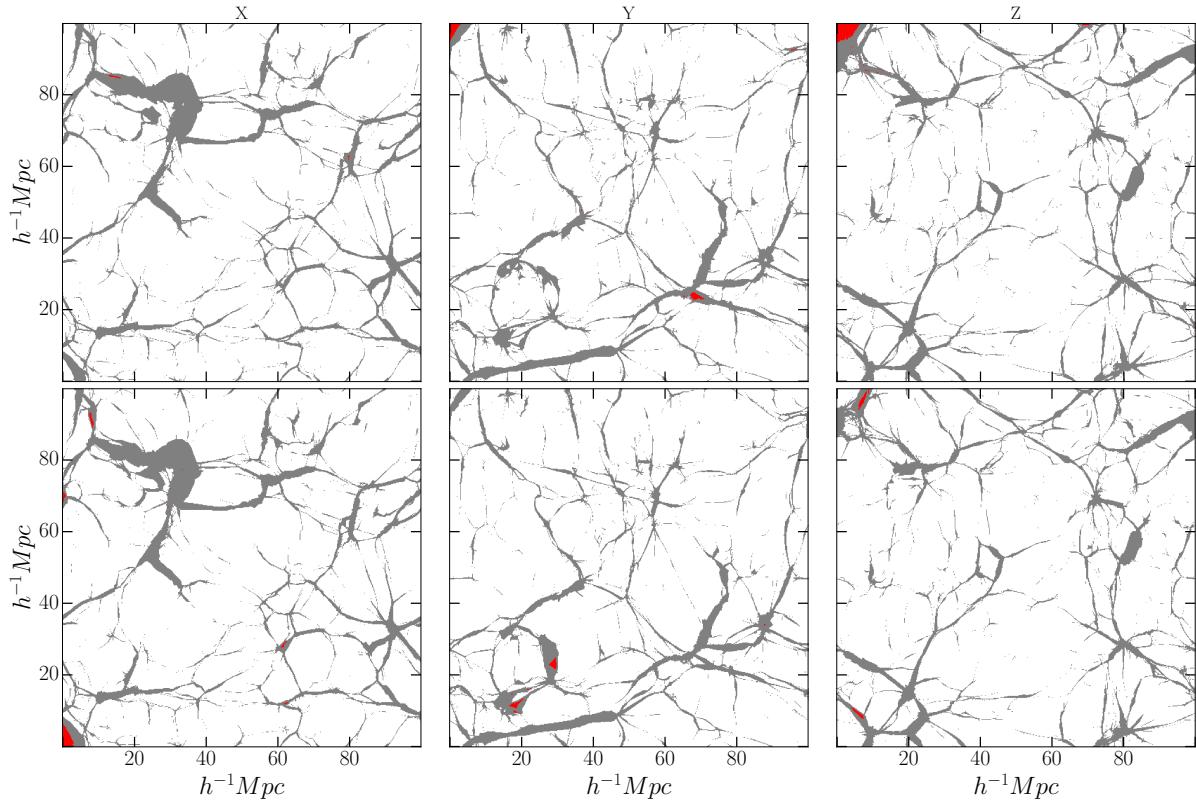


Figure 4.2: Opposite faces of the multistream field for the simulation box with  $N_p = 128^3$ . Non-void regions (gray) have  $n_{str} > 1$ . The largest void (white) in the entire field spans over the entire box. Rest of the smaller isolated voids (red) occupy very small volume fraction.

Table 4.2: Volume fraction  $VF_V$  of the voids, total number of isolated voids  $N_V$  and the filling fraction of the largest void  $FF_1/VF_V$  at different refinement factors  $l_l/l_d$ . The filling fractions of the largest void at each refinement factor show that most of the  $n_{str} = 1$  region is almost entirely a single percolating structure.

$N_p$	$l_l/l_d$	$VF_V$	$N_V$	$FF_1/VF_V$
$128^3$	1	93.46%	1	100%
$128^3$	2	93.44%	11	99.999%
$128^3$	4	93.44%	113	99.999%
$128^3$	8	93.44%	914	99.997%
$256^3$	1	90.80%	11	99.999%
$256^3$	2	90.80%	97	99.999%
$256^3$	4	90.80%	1029	99.997%
$256^3$	8	90.80%	7259	99.964%

equal to unity.

For simulation box with  $128^3$  particles,  $n_{str} = 1$  regions have a large volume fraction of  $VF_V \approx 93$  per cent regardless of the value of refinement factor (shown in Table 4.2). Multi-stream web structure in the simulation with higher mass resolution ( $N_p = 256^3$ ) is better enhanced, and the single streaming void occupies around 90 per cent of the volume. Figure 4.2 shows the single streaming voids occupying large volume of the simulation with  $128^3$  particles at refinement factor of 4.

### 4.3.1 Connectivity of the voids

In order to find whether the void regions of the multistream field are connected or not, we isolate three-dimensional segments with  $n_{str} = 1$  and separately label them. The number of disconnected voids in the simulation with  $N_p = 128^3$  range from 1 (for refinement factor,  $l_l/l_d = 1$ ) to about 900 (for  $l_l/l_d = 8$ ) as shown in Table 4.2. Number of isolated voids increases similarly in the simulation with  $N_p = 256^3$  particles as well.

Smoothing of the structure at lower resolution of the multistream field results in increased connectivity of single-streaming regions. In Figure 4.2, opposite faces on each axes of the multi-field, show a large connected void (white). This means that the largest void percolated throughout the

multistream field in all directions. This result is in agreement with Falck & Neyrinck (2015), who studied percolation of ORIGAMI-voids in simulations with side lengths of 100 and  $200h^{-1}\text{Mpc}$ . In addition to the percolating the field, the largest void also fills most of the void volume: the ratio of filling fraction of the largest void  $FF_1$  to the volume fraction of  $n_{str} = 1$  regions in the simulation is close to unity (see Table 4.2). This phenomenon is seen at each of the refinement factors in our analysis. Hence, over 99.9 per cent of the single-streaming sites are connected throughout the simulation box, and they form a single empty region.

As previously mentioned, the multistream web structures of  $n_{str} = 3$  form the first gravitationally collapsed structures. These tiny structures are better resolved in higher refinement factors, and they tend to enclose greater number of pockets of single-streaming voids inside them. The red regions in Figure 4.2 some of the small voids on faces of the simulation box with  $128^3$  particles. Despite increase in the number of small voids at each of the refinement factors, these void regions (i.e., the single streaming regions excluding the largest void) collectively occupy less than 0.1 per cent of the total void volume in both the simulations. It is also likely that the small voids are simply due to numerical noise. However, the major conclusion regarding small voids remains the same up to refinement factor of 8. We do not pursue further investigation due to tiny effects.

### 4.3.2 Halo boundaries within the void

Dark matter haloes are the most non-linear objects in the cosmic web. With the exception of ORIGAMI (Falck et al., 2012), most of the halo finders do not consider multistreaming in the configuration space for finding haloes. Potential haloes found by several such halo finding methods, hence, may have boundaries that intersect with the single-streaming void, which is the least non-linear structure in the dark matter universe. Colberg et al. (2008) even mention existence of ‘void-haloes’ in several halo finder algorithms.

We studied the  $n_{str}$  environment of the haloes detected using the Friends-of-Friends method (FOF-Davis et al. 1985) as illustrated in Figure 4.3. FOF-haloes with more than 20 particles are detected using linking-length of  $b = 0.2$  in the simulation with  $128^3$  particles. We implement the

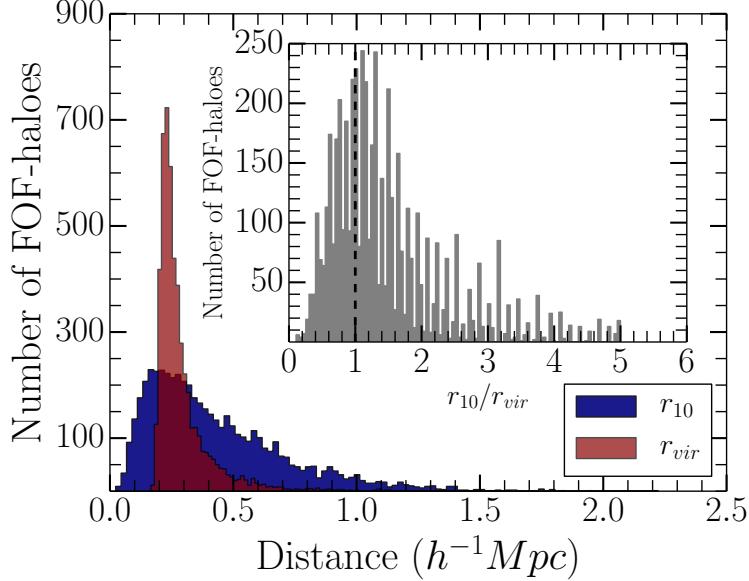


Figure 4.3: Single-streaming void distribution on diagnostic spheres around FOF-haloes are considered. At radius  $r_{10}$ , each diagnostic sphere has  $n_{str} = 1$  on 10 per cent of its spherical surface. Distribution function of  $r_{10}$  (blue) and FOF-radii  $r_{vir}$  (red) are shown. Inner plot shows the distribution function of  $r_{10}/r_{vir}$ . The haloes within the dashed line have at least 10 per cent of their virial-surfaces in contact with  $n_{str} = 1$  regions.

diagnosis method prescribed in Ramachandra & Shandarin (2015): a large number of points are randomly selected on diagnostic spherical surfaces centred at the FOF-centre of the halo. Multi-stream values are iteratively calculated at these spherical surfaces of various radii. We define the distance from centre of a halo,  $r_{10}$ , where  $n_{str} = 1$  at 10 per cent of the surface of the diagnostic sphere. Distribution of this void-distance parameter is compared to the virial radii  $r_{vir}$  of the FOF-haloes. Surprisingly,  $r_{10}$  distribution peaks at slightly lower values than the  $r_{vir}$  distribution. This implies a large number of FOF-haloes are in the vicinity of the void.

For specific examples of some FOF-haloes, Ramachandra & Shandarin (2015) showed that single-stream may appear within their virial radii too. The distribution of  $r_{10}/r_{vir}$  in the inner plot of Figure 4.3 shows the same phenomenon. The FOF-haloes within  $r_{10}/r_{vir} < 1$  (represented by the vertical dashed line) have  $n_{str} = 1$  on 10 per cent of their virial surfaces. The figure illustrates that a large number of FOF-haloes satisfy this condition, thus are in contact with the void surfaces. Hence not all the FOF particles have undergone a gravitational collapse during their evolution.

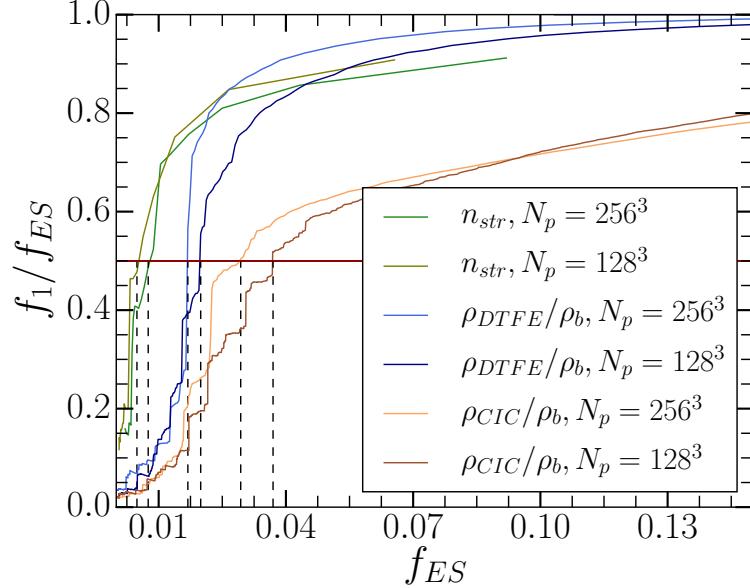


Figure 4.4: Percolation plot in the multistream field and mass density. Two density estimators - CIC and DTFE are shown. Percolation transition (at  $f_1/f_{ES} = 0.5$  shown by the horizontal red line) occurs at smaller excursion set volumes for the multistream field, as seen by the dashed lines for both the curves. It is worth stressing that the percolation curves for  $n_{str}$  field are bounded by conditions  $f_{ES} < 0.1$ .

For methods such as FOF, there is no unambiguous linking-length criterion for voids. Similarly for the density fields, a range of under-densities are prescribed by various void finder methods (cf. Colberg et al. 2008). On the other hand, the multistream field unambiguously identifies all the regions without a single gravitational collapse as voids. Haloes detected on the multistream field may address the issue of haloes being in contact with voids.

## 4.4 Percolation in the multistream web

A single percolating void fills the  $n_{str} = 1$  regions almost entirely, as discussed in Section 4.3.1. Disconnected pockets of void may exist, but they collectively occupy very small volume fraction (less than 0.1 per cent of the total volume as tabulated in Table 4.2). Whereas, the non-void structure in the multistream field has a different topological structure. The regions selected with a lower bound on  $n_{str}$  could be isolated (generally for high  $n_{str}$  thresholds) or connected in a

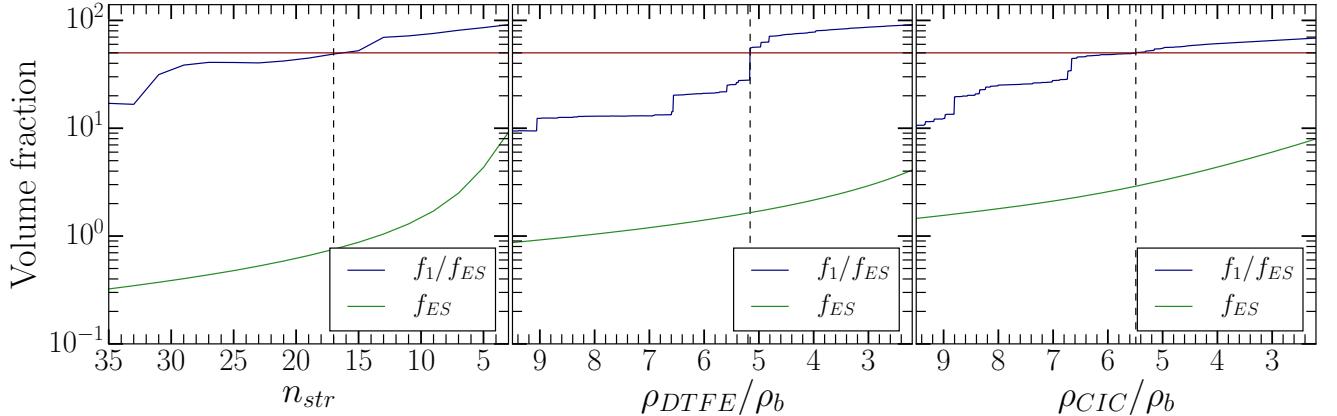


Figure 4.5: Percolation threshold in the multistream (left panel) and matter density fields. Matter density is calculated using DTFE (middle panel) and CIC (right panel) with a refinement factor of 2 in the simulation with  $256^3$  particles. The volume fraction of excursion set and the filling fraction of the largest structure is shown. Percolation transition in multistream field at  $n_{str} = 17$  is shown by the dashed vertical line. Percolation at  $\rho_{DTFE}/\rho_b = 5.16$  and  $\rho_{CIC}/\rho_b = 5.49$  are shown by the dashed vertical line.

percolating region (for low  $n_{str}$  thresholds). We investigate the topological transitions in these excursion sets of multistream field.

The volume fraction as a function of number-of-streams decreases according to a power law in the  $n_{str} > 1$  structure (Shandarin et al. 2012 and Ramachandra & Shandarin 2015 report  $\text{VF}(n_{str})$  decreasing as  $n_{str}^{-2.8}$  and  $n_{str}^{-2.5}$  respectively for their simulations). The volume fraction of the excursion set  $f_{ES}(n_i)$  is the ratio of volume of all the regions with a lower bound  $n_i$  on the multistream field to the total volume  $V_{tot}$  of the simulation box, i.e.,  $f_{ES}(n_i) = \frac{V_{ES}}{V_{tot}} = \sum_{n_{str} \geq n_i} \text{VF}(n_{str})$ . Since volume fraction of each  $n_{str}$  rapidly increases with a decrease in multistream value, so does the  $f_{ES}$ .

The excursion set may have number of isolated segments of different volumes. A measure of connectivity in the excursion set regions can be given by the filling fraction,  $f_1/f_{ES}$ , where  $f_1$  is the volume fraction of the largest isolated region in the excursion set.  $f_1$  can be computed numerically in the simulations. If the value of  $f_1/f_{ES}$  is close to 0, then none of the isolated regions dominate the excursion set. This implies absence of percolation. If  $f_1/f_{ES}$  is close to one, it implies a single connected structure dominates most of the excursion set.

The filling fraction  $f_1/f_{ES}$  grows from 0 to 1 occurs rapidly  $f_{ES}$  during percolation phase transition. A practical robust definition of the percolation transition is at  $f_1/f_{ES} = 0.5$ , i.e, when the largest region occupies more than 50 per cent of the excursion set volume. The percolation plot in Figure 4.4 reveals this phenomenon. Excursion volume fraction  $f_{ES}$  at this transition,  $f_{ES}^{(p)} = 0.48$  and 0.75 per cent for the simulations with with  $128^3$  and  $256^3$  particles respectively (although the numbers were obtained in one simulation each. The difference may be well within the range of statistical errors for this size of simulation box). After the percolation transition, the filling fraction of the largest structure stabilizes towards unity.

The nature of the transition in mass density field is similar to that in multistream field. For the simulation simulation with  $256^3$  particles, the density is calculated using CIC method at  $256^3$  and  $512^3$  grid points. In Figure 4.4, the percolation phenomenon in both mass density fields is shown along with that of multistream fields. The excursion set volume fraction at percolation transition,  $f_{ES}^{(p)}$  is lower for multistream field, because the filaments in the multistream field are thinner than that of density picture. Volume fraction of the largest structure detected in the density field also tends to unity with decreasing  $f_{ES}$ , albeit less rapidly as that of the multistream field. This means that while the largest structure in a multistream web occupies most of the structure, the over-density excursion set is more fragmented.

The excursion volume fraction of the multistream web structure is limited to a small fraction of less than 10 per cent since rest of the volume is void. The excursion set volume fraction increases with decreasing number-of-streams and reaches it's maximum at  $n_{str} = 3$ . At this limit the filling fraction  $f_1/f_{ES}$  is still less than unity, about 95 per cent. These two peculiar properties of the multistream field explain the shape of the percolation curves in Figure 4.4. Since the multistream flow field is a discrete data field, the percolation transition is seen to occurs at a particular value of  $n_{str}$  rather than a large range of values. For  $n_{str} = 17$ , the largest structure in the excursion set occupies more than half the volume of the entire excursion set. At this multistream threshold, the largest segment starts spanning large volume of the simulation box (as observed in the left panel of Figure 4.5). The volume fraction of the excursion set at this percolation transition is  $f_{ES}^{(p)} = 0.75$

per cent for simulation with  $256^3$  particles.

The percolation transition at  $n_{str} = 17$  could be used as a criterion for detecting filaments in the cosmic web. Since the largest  $n_{str} \geq 17$  region occupies more than 50 per cent of the excursion set, it is essentially the ‘backbone’ of the cosmic web (Shandarin et al., 2010). Heuristic analysis as discussed by Ramachandra & Shandarin (2015) also arrived at the same threshold for identifying filaments. That analysis was based on a multistreams variation in halo environments, hence a local value. From our percolation analysis, we see that it is also justified globally.

In the simulation with  $256^3$  particles, percolations in the density field occurs at  $\rho_{DTFE}/\rho_b = 5.16$  and  $\rho_{CIC}/\rho_b = 5.49$  for densities calculated with DTFE and CIC respectively. Here  $\rho_b = 256^3/100^3 M_\odot h^{-3} \text{Mpc}^{-3}$ , the background density. Notice that these values correspond to the density as calculated by the CIC and DTFE algorithms, and it might be different for other density finding methods. The volume fraction of the excursion set of over-densities at the percolation,  $f_{ES}^{(p)} = 2.7$  per cent, is considerably higher than the corresponding  $f_{ES}^{(p)}$  value in the multistream field. This implies that the percolation occurs at larger values of filling fraction in mass densities.

## 4.5 Local geometry of the multistream field

The multistream field has a constant value of 1 for around 90 per cent of the simulation box. At least one gravitational collapse occurs in the remaining 10 per cent of the volume. In these non-void regions, the  $n_{str}$  value varies from 3 to very high values, often in the order of thousands. In the multistream field of refinement factor of 2 for simulation with  $N_p = 128^3$  particles, maximum  $n_{str}$  is 2831. Within the non-void structure, the multistream field may have several local maxima, minima and saddles. Variation of  $n_{str}$  is especially high inside halo boundaries, where the particles in their non-linear stage of evolution have undergone a large number of flip-flops.

Local second order variation in a scalar field  $f$  like the multistream field can be found using the Hessian matrix  $\mathbf{H}(f)$ . An element of the Hessian matrix is given in Equation 5.1, where  $i$  and  $j$  can be any of  $x$ ,  $y$  or  $z$  directions.

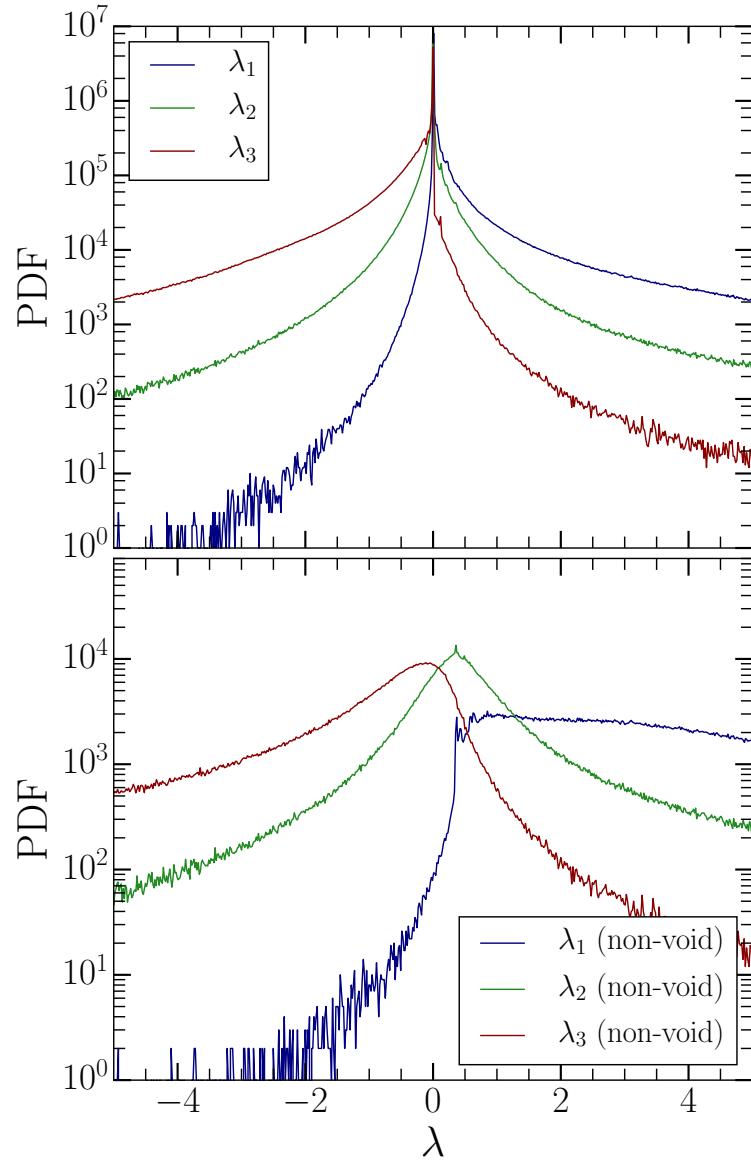


Figure 4.6: Probability distribution function of the sorted eigenvalues of the Hessian  $\mathbf{H}(-n_{str})$  in the simulation box with  $N_p = 128^3$ . Top panel: Distribution in the entire simulation box. The multistream field is calculated at refinement factor  $l_l/l_d = 2$  and smoothing scale of equal to  $l_d$ . All the three eigenvalue data fields have a highest number of points where their value is 0. Bottom panel: Hessian eigenvalues for the non-void region ( $n_{str} > 1$ ) is shown. Total number of eigenvalue triplets are less than 10 per cent of that of the full simulation box. Eigenvalues close to zero in non-void regions are notably fewer than in the entire simulation box.

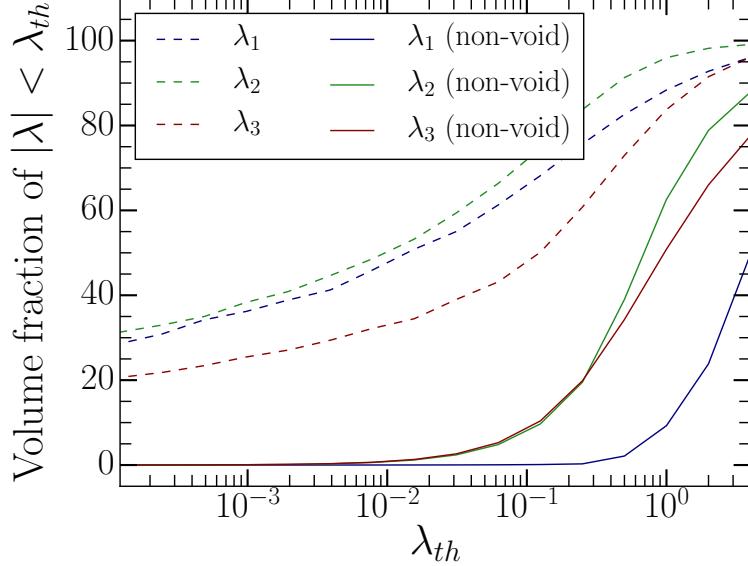


Figure 4.7: Comparison between small eigenvalues of the multistream Hessian  $\mathbf{H}(-n_{str})$ . Percentage of eigenvalues with absolute values less than a cut-off,  $\lambda_{th}$  are shown for full simulation box (dashed lines) and the multistream web structure (solid lines). The multistream web has fewer eigenvalues below  $\lambda_{th} = 0.1$ . The void seems to have most of the small eigenvalues.

$$\mathbf{H}_{ij}(f) = \frac{\partial^2 f}{\partial x_i \partial x_j} \quad (4.1)$$

In our analysis, we have chosen  $f = -n_{str}(\mathbf{x})$  for understanding local variations of the multistream field. The resulting Hessians at each point on the configuration space are always symmetric matrices, as illustrated in Appendix C. The eigenvalues of these Hessian matrices are always real, and depending on if their values are positive or negative, one may infer local geometrical features in the multistream field.

Within the void, there is no variation in the multistream values. Hessians  $\mathbf{H}(-n_{str})$  are zero matrices in large volume fraction of the simulation box (around 90 per cent in both the simulations) due to the constant value of  $n_{str} = 1$  in this percolating void. Eigenvalues of these Hessian matrices, sorted as  $\lambda_1 \geq \lambda_2 \geq \lambda_3$  are close to 0 at a large number of regions as shown in the top panel of Figure 4.6. In the simulation with  $128^3$  particles, the median values of each eigenvalue are 0.09,  $-3 \times 10^{-10}$  and  $-0.11$  for  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  respectively. By selecting just the non-void region by  $n_{str} > 1$ , notably fewer number of eigenvalues have small absolute values. The median values

of each of the eigenvalues in the non-void regions are 4.01, 0.48, and  $-0.85$  respectively for  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$ . Bottom panel in Figure 4.6 shows a significant change in the probability distribution of Hessian eigenvalues around 0, the distribution pattern at the tails are mostly identical to the distribution pattern in the entire simulation box.

A large fraction of eigenvalues in non-void regions are still around 0, but their percentage is quite less compared to that of the entire box. For instance, nearly 66 per cent of  $\lambda_1$ 's, 72 per cent of  $\lambda_2$ 's and 48 per cent of  $\lambda_3$ 's are within in the range of  $0.0 \pm 0.1$  in the entire simulation box. However, with the exclusion of void regions, these volume fractions drops to 0.1, 7.7 and 8.4 per cent respectively (Figure 4.7). Hence most of the eigenvalues at the void region have small absolute values.

Hessian eigenvalues in multistream fields differ from that in density, gravitational potential or velocity shear tensor. Constant scalar value of  $n_{str}$  facilitates the Hessian  $\mathbf{H}(-n_{str})$  matrices to be presumptively close to zero. On the other hand, in density field manifests in a range of low values in the voids, resulting in non-zero Hessian matrices. Eigenvalues of velocity shear tensor do not peak at zero either Libeskind et al. (2013). For the deformation tensor, morphological characterization of the cosmic web using Zel'dovich formalism shows that each eigenvalue must be negative in voids.

The eigenvalues of  $\mathbf{H}(-n_{str})$  span a large range of values in our cosmological simulation. The largest eigenvalue of the triplets,  $\lambda_1$  having large positive values throughout the multistream web structure ( see Figure 4.8 ). Absolute values  $|\lambda_1|$ ,  $|\lambda_2|$  and  $|\lambda_3|$  peak around the neighbourhood of intersections of filaments. These junctions are usually high streaming regions due to shell crossing from multiple directions. Ramachandra & Shandarin (2015) observed that these regions with intersecting filaments are in the vicinity of large FOF haloes.

If the Hessian matrices are positive definite in a region, i.e., if all the eigenvalues are strictly positive, then the interior of this convex region has at-most one minimum. For our choice of  $-n_{str}(\mathbf{x})$  as the domain of Hessian, this means that the convex neighbourhoods around local maxima of the multistream field are isolated by the positive definite Hessian matrices. Closed surface

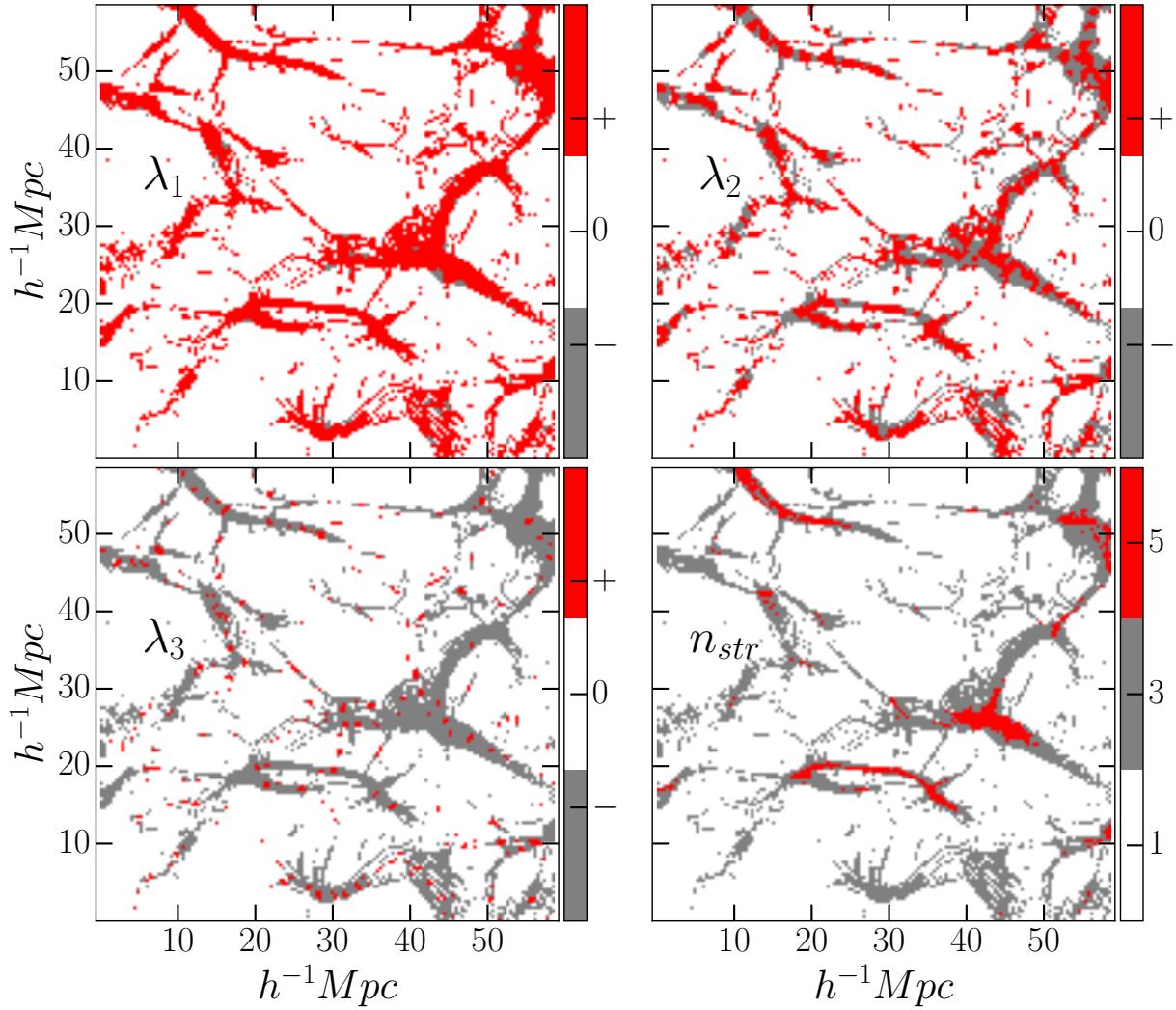


Figure 4.8: Eigenvalues of the Hessian matrix  $\mathbf{H}(-n_{str})$  in a slice of  $50 h^{-1} \text{ Mpc} \times 50 h^{-1} \text{ Mpc}$  slice of the simulation box of  $128^3$  particles. Variation in the eigenvalues in the multistreaming web structure is shown. The largest eigenvalue  $\lambda_1$  (top left panel) has positive values throughout the structure. The smallest eigenvalue  $\lambda_3$  (bottom left) has negative values surrounding positive definite regions of the  $n_{str}$  field. Corresponding multistream field is shown in the bottom right panel for single, three and more than five streams.

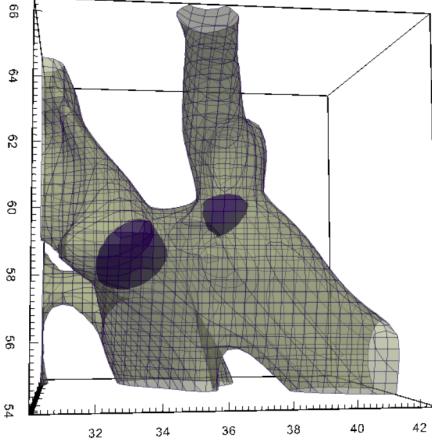


Figure 4.9: Surfaces identified in the multistream field. Blue regions are closed regions with  $\lambda_3 > 0$ , which we identify as two haloes. Other surface has an open curvature along one direction, with  $\lambda_1 > \lambda_2 > 0$  and  $\lambda_3 < 0$ .

contours at high streaming or the most non-linear regions are selected. These regions may indeed be the regions of dark matter haloes.

The smallest eigenvalue,  $\lambda_3$  has lowest volume fraction of all the eigenvalues in the positive tail of the distributions in Figure 4.6. Since the condition  $\lambda_3 > 0$  ensures the Hessian matrix to be positive definite, we may use it as a primary criterion in isolating compact regions of dark matter haloes. These regions also roughly correspond to isolated globs as seen in Figure 4.9. Local geometry analysis is pertinent for halo detection due to compact geometry of the haloes. In principle, other components of the cosmic web could also be detected. Tubular structures in filaments could be detected, as shown in Figure 4.9, using conditions on the eigenvalues as  $\lambda_1 > \lambda_2 > 0$  and  $\lambda_3 < 0$ . Fabric-thin walls could be detected by  $\lambda_1 > 0$  and  $\lambda_3 < \lambda_2 < 0$ .

### 4.5.1 Softening of the multistream field

Hessian eigenvalues are generally defined on continuous functions. Although our domain of the Hessian is an inherently integer-valued field, it describes the multistream structure at the level of diagnostic grid. Hence it may be considered to be numerically equivalent to a continuous function where the numerical approximation of differentiation is a valid operation. This can be verified mathematically by finding that Hessian  $\mathbf{H}(-n_{str})$  is symmetric (Appendix C shows the numerical

approximation of the Hessian matrix term for generic unfiltered multistream field.)

Smoothing the multistream field (at the refinement level of  $l_l/l_d = 1$  or 2) effectively reduces noise. There is also a systematic variation in the distribution of smoothed  $n_{str}$  values as shown in Figure 4.10. Volume fraction of the single-streaming voids only varies from 90.8 per cent without smoothing to 89.1 per cent for the Gaussian softening length of  $0.39 h^{-1} Mpc$  (twice the length of diagnostic grid  $l_d$ ). On the other hand,  $n_{str} = 3$  regions gain volume fraction from 4.9 per cent in un-smoothed field to 7.1 per cent for  $0.39 h^{-1} Mpc$ . This is seen in the multistream structures of smoothing scales of  $0.39 h^{-1} Mpc$  in Figure 4.11. Multi-stream regions with  $3 < n_{str} \leq 100$  occupy correspondingly lower volumes for higher smoothing, and the variation is noisy beyond  $n_{str} > 100$ . Figure 4.11 shows the multistream field on a small slice of the simulation at different softening scales, and walls and filaments are resolved better with increasing softening.

Smoother multistream fields result in less noisy PDFs of the Hessian eigenvalues. For instance, the volume fraction of regions with positive curvature (i.e.  $\lambda_3 > 0$ ) is 2.4%, 2.3% and 2.5% for scales  $0.20h^{-1} Mpc$ ,  $0.39h^{-1} Mpc$ ,  $0.78h^{-1} Mpc$  respectively. Further analysis of smoothed positive definite regions is relevant in determining halo boundaries, and will be extensively discussed in the next paper.

### 4.5.2 Resolution dependence

Multi-stream calculation can be done at arbitrarily high resolutions by populating the tetrahedral simplices. For our resolution study, we have chosen a smaller slice of  $50h^{-1} Mpc \times 50h^{-1} Mpc \times 50h^{-1} Mpc$  (grid of size  $64^3$  from the N-body simulation) from the simulation with  $N_p = 128^3$  particles. The multistream field is calculated at 4 different refinement factors, i.e., at diagnosis grids of size  $64^3$  ( $l_l/l_d = 1$ ),  $128^3$  ( $l_l/l_d = 2$ ),  $256^3$  ( $l_l/l_d = 4$ ) and  $512^3$  ( $l_l/l_d = 8$ ) respectively.

Volume fractions of each multistream does not change systematically for different levels of refinement, except at very high  $n_{str}$  values (see Ramachandra & Shandarin 2015 for dependence of  $n_{str}$  variation on refinement of the diagnostic grid). At high multistream values, higher resolutions reveal a considerably less noisy multistream fields.

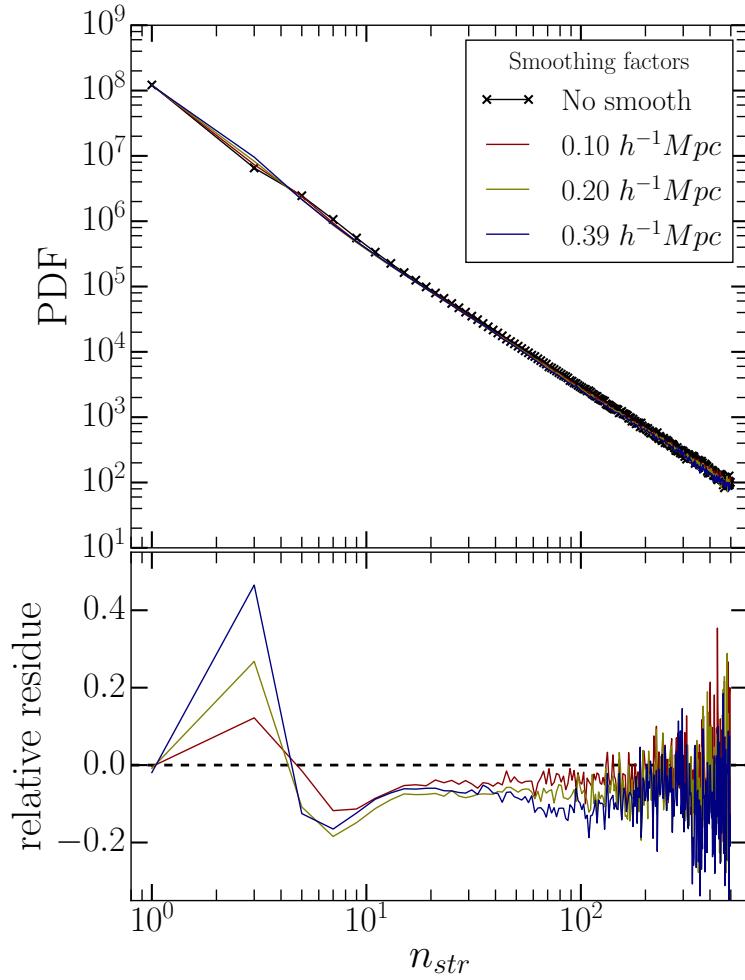


Figure 4.10: Probability distribution function of the multistream  $n_{str}$  values in the simulation box with  $N_p = 256^3$ . The multistream field is calculated at refinement factor  $l_l/l_d = 2$ . Unsmoothed multistream field is compared with different Gaussian filtering scales. Softening scales of equal to 0.5, 1, and 2 times the side length of diagnostic grid  $l_d$  correspond to  $0.10h^{-1}\text{Mpc}$ ,  $0.20h^{-1}\text{Mpc}$ , and  $0.39h^{-1}\text{Mpc}$  respectively.

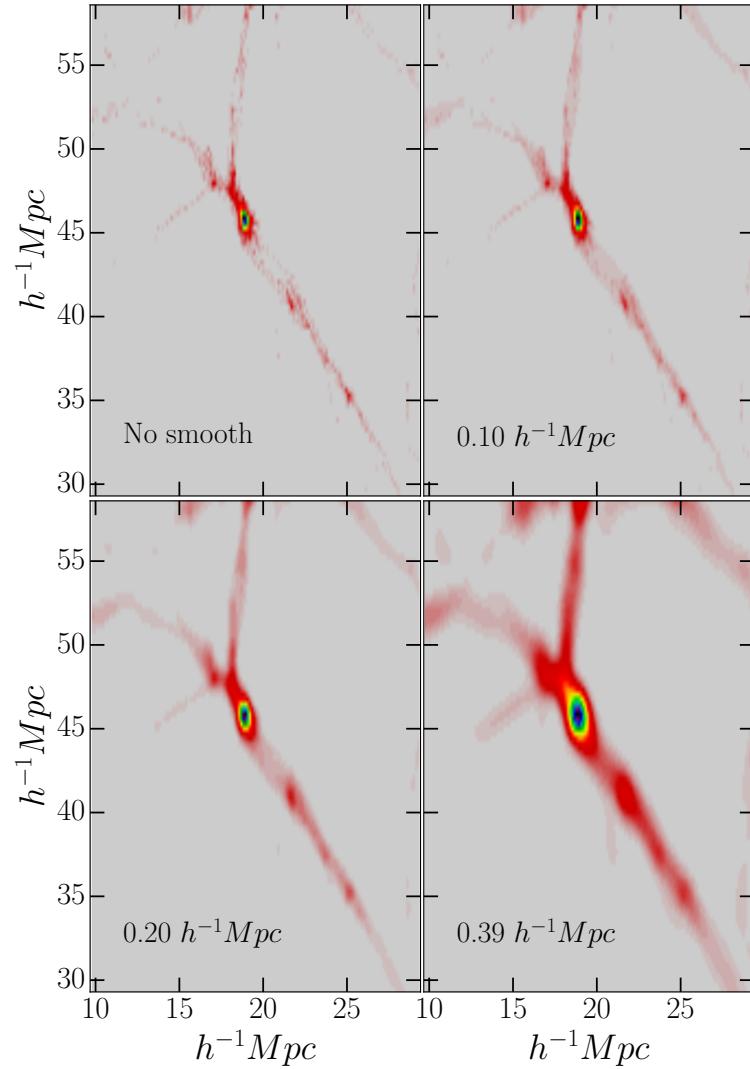


Figure 4.11: Multi-stream field at various softening scales in the simulation box with  $N_p = 256^3$ . The multistream field is calculated at refinement factor  $l_l/l_d = 2$ . Unsmoothed multistream field is compared with different Gaussian filtering scales equal to  $0.10 h^{-1} \text{Mpc}$ ,  $0.20 h^{-1} \text{Mpc}$ , and  $0.39 h^{-1} \text{Mpc}$  respectively.

Table 4.3: Volume fraction (in per cent) of  $n_{str}$  thresholds for cosmic web structures as defined by Ramachandra & Shandarin (2015). Multi-stream field is calculated at 1, 2, 4, and 8 times the native simulation resolution of  $64^3$  grids. Small slice of  $50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc}$  is chosen for the analysis.

Global thresholds	$64^3$	$128^3$	$256^3$	$512^3$
$n_{str} = 1$ (Void)	90.87	90.92	90.94	90.94
$3 \leq n_{str} < 17$ (Wall)	8.71	8.66	8.63	8.64
$17 \leq n_{str} < 90$ (Filaments)	0.39	0.39	0.39	0.39
$n_{str} \geq 90$ (Haloes)	0.034	0.035	0.036	0.036

There are no variations in the volume fractions of the cosmic web components classified using the global  $n_{str}$  thresholds as shown in Table 4.3. Voids ( $n_{str} = 1$ ) occupy about 90 per cent of the volume at each refinement factor. Rest of the heuristic thresholds that identify the structure components (as prescribed by Ramachandra & Shandarin 2015) are constant multistream contours:  $3 \leq n_{str} < 17$  for walls,  $17 \leq n_{str} < 90$  for filaments and  $n_{str} \geq 90$  for haloes. Since the volume fraction of each  $n_{str}$  values are about the same at each refinement factor, the volume fraction of the cosmic web components corresponding to global multistream thresholds do not vary considerably.

However, local geometry analysis of the multistream flow field varies considerably on the resolution of the analysis grid. For our Hessian  $\mathbf{H}(-n_{str})$ , the regions with  $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$  in non-void regions occupy 1.8 per cent of the entire box in native resolution of diagnostic grid, as shown in Table 4.4. This fraction reduces to 1.3 per cent at diagnostic grid of  $512^3$  resolution. Variations with refinement factors are seen in other eigenvalue conditions in the non-void too: volume fraction of  $\lambda_1 > 0 > \lambda_2 \geq \lambda_3$  regions increases from 1.7 per cent at refinement factor of 1 to 3 per cent at refinement factor of 8. Volume fraction of  $\lambda_1 \geq \lambda_2 > 0 > \lambda_3$  regions decreases from 5.6 to 4.6 per cent with the increase of refinement from 1 to 8.

In principle, the conditions for geometric criteria are:  $\lambda_1 > 0 > \lambda_2 \geq \lambda_3$  for locally flat regions,  $\lambda_1 \geq \lambda_2 > 0 > \lambda_3$  for locally tubular structures and  $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$  for clumped blobs. However, the tabulated the volume fractions in Table 4.4 does not correspond to cosmic web components themselves. Identification of the components may require post processing steps.

High resolution studies of multistream fields would play an important role in detection of walls

Table 4.4: Volume fraction of criteria based on  $n_{str}$  and  $\lambda$ s of  $\mathbf{H}(-n_{str})$  calculated at various resolutions. We chose a smaller slice of  $50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc} \times 50h^{-1}\text{Mpc}$  i.e., half the volume of the original GADGET simulation box. The refinement factors are the multiplication factors of 1, 2, 4 and 8 times of the native resolution ( $64^3$ ) of the simulation grid along each axis. Eigenvalues of the Hessian of the field are local geometric parameters. The void is globally defined as  $n_{str} = 1$  and the multistream web structure as  $n_{str} > 1$ .

Global/local conditions	$64^3$	$128^3$	$256^3$	$512^3$
$n_{str} = 1$ (Void)	90.87	90.92	90.94	90.94
$n_{str} > 1; \lambda_1 > 0 > \lambda_2 \geq \lambda_3$	1.72	2.22	2.67	2.96
$n_{str} > 1; \lambda_1 \geq \lambda_2 > 0 > \lambda_3$	5.60	5.28	4.91	4.57
$n_{str} > 1; \lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$	1.81	1.56	1.37	1.26

and filaments. These two components have smaller length scales along at least one direction with respect to others. As seen in Section 4.3.1, walls are more resolved in high resolution of multi-stream fields, enclosing pockets of voids (see Figure 4.2).

However, a Hessian analysis to identify filaments and walls may be considerably different from that of halo finding due to the following reasons: First, a local geometrical analysis is uniquely convenient for detecting dark matter haloes since they are local structures. Filaments and walls, alternatively, are structures that span over large distances. Secondly, we try to find regions around local maxima of multistream field for haloes. Whereas, filaments and walls have much weaker relationship with local multistream maxima. Filaments and walls usually deviate from flat planar or straight tubular geometries: they often have complicated structures several connections and branches. For these reasons, Hessian eigenvalues alone would not be sufficient in detecting walls or filaments.

## 4.6 Discussion

Formation of multiple velocity streams in the context of structure formation has been known in the past, starting from Zel'dovich approximation. Quantification of the multistreams in N-body simulations, however, was recently achieved by Shandarin et al. (2012) and ? using the Lagrangian sub-manifold. In our study, the multistream fields are calculated using the tessellation algorithm

by Shandarin et al. (2012). We have analysed, for the first time, the local geometry and percolation properties of the cosmic web using this multistream field.

Distinguishing the configuration space into void and non-void is one of the uses of the multistream field. Lagrangian sub-manifold has no folds in the beginning, thus  $n_{str} = 1$  uniformly throughout the simulation. Gravitational instability folds the sub-manifold in complicated ways, however, most of the volume has particles without any collapse. Shandarin et al. (2012) and Ramachandra & Shandarin (2015) observed that the single-streaming voids occupy around 85-90 per cent of the simulations at  $z = 0$ . In this study, we found that the void regions are also connected in a way that the largest percolating void occupies more than 99 per cent of the all the single-streaming regions. Recent study by Wojtak et al. (2016) uses a watershed transform method in the density field prescribed by Lagrangian tessellations (Shandarin et al. 2012 and ?) to analyse the evolution of isolated voids. Another recent study by Falck & Neyrinck (2015) on ORIGAMI-voids also reveal a similar percolation at the limit of simulation resolution. They observed persistence of this phenomenon for different resolutions of the N-body simulation. Multi-stream analysis, on the other hand, is not limited to mass resolution of the simulation. Our multistream analysis refined upto 8 times the simulations resolution revealed that the percolation phenomenon still persists. However, at high refinements of the multistream field, we observed small voids that are enclosed by highly resolved non-void membranes.

Walls are the first collapsed structures in the dark matter Universe. At highly refined multi-stream field, thin membranes of the structures are often resolved, revealing small voids enclosed by them (compare two top panels in Figure 4.12). These preliminary structures are separated from the voids by caustic surfaces. These caustics have volume measure zero, which makes detection of their surface harder in the multistream field, even at very high resolutions. On the other hand, caustic surfaces themselves can be detected using the Lagrangian sub-manifold by identifying the common faces of neighbouring tetrahedra with opposite volume signs (Shandarin et al., 2012). They are shown in the bottom panel in Figure 4.12. One can see that increasing the refinement factor from 2 to 8 adds mostly walls but the complete wall structure shown in the bottom panel is still

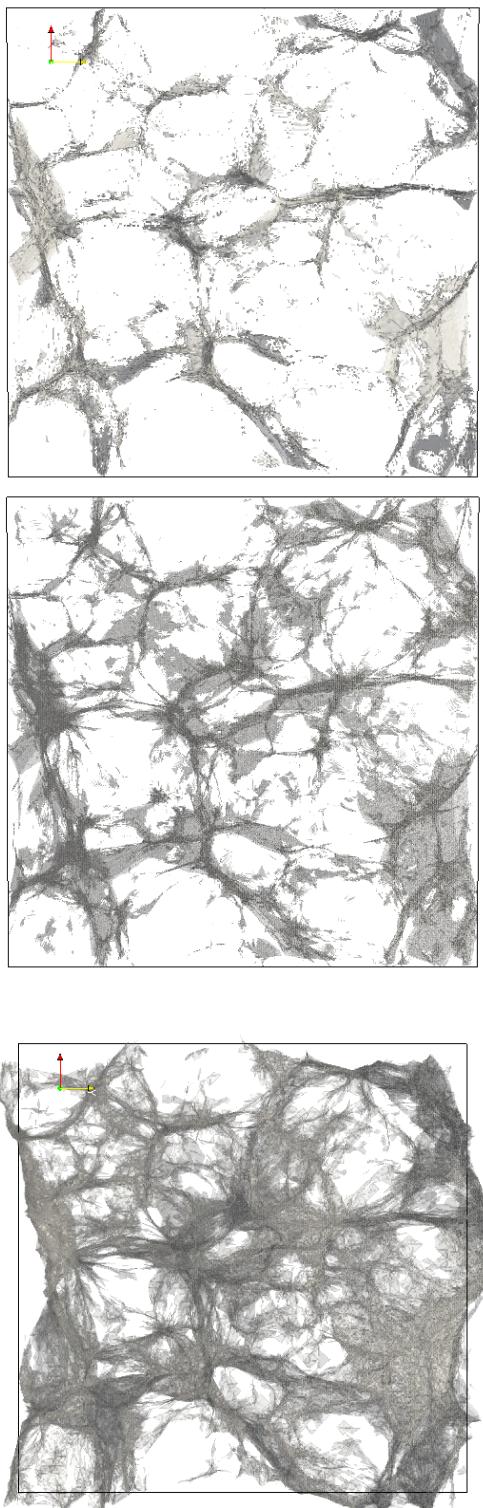


Figure 4.12: Two top panels show three contours ( $n_{str} = 3, 11, 17$ ) in a slice  $100h^{-1}\text{Mpc} \times 100h^{-1}\text{Mpc} \times 10h^{-1}\text{Mpc}$  in the simulation with  $128^3$  particles, computed at two refinement factors: 2 (upper) and 8 (lower). The bottom panel shows the caustic surfaces in the same slice.

considerably greater. Please note that the plots in two top panels adjusted exactly to the simulation box in Eulerian space, and the bottom plot shows the Lagrangian box mapped to Eulerian space without adjusting to the simulation box.

There are extensive number of topological indicators in the context of density fields or spatial co-ordinates - such as alpha shapes, Betti numbers, genus statistics. Although a comparative study of these topological measures in multistream fields may be interesting, it is not the intent of this paper. In this study, we only investigate percolation transitions in excursion sets of multistreams as a preliminary analysis of topological connectivities. Excursion sets in density fields are shown to have quick percolation transitions (Shandarin et al., 2010) and a similar trend in multistream field is investigated here.

Excursion sets of multistream and density field (calculated using CIC and DTFE in this study) reveal some of the topological differences. At any volume fraction of excursion set  $f_{ES}$ , the filling factor of the largest structure  $f_1/f_{ES}$  is lower for mass density (both CIC and DTFE). This concludes that the mass density field is more fragmented than the multistream field. A large number of disconnected segments are seen at high  $n_{str}$  or  $\rho/\rho_b$  thresholds, and the number of connections increase with decreasing  $n_{str}$  threshold.

Global connectivities in the cosmic web is slightly different for multistream field and the density field. The largest structure in the excursion set starts percolating at certain values of excursion volume fraction ( $f_{ES}$ ). As shown in Section 4.4, these percolation transitions occur at  $\rho_{DTFE}/\rho_b = 5.16$ ,  $\rho_{CIC}/\rho_b = 5.49$  for density fields and  $n_{str} = 17$  for the multistream field. The corresponding percolation volume fraction  $f_{ES}^{(p)}$  is smaller for multistream fields ( $f_{ES}^{(p)} = 0.75$  per cent for multistream field and  $f_{ES}^{(p)} = 1.7$  per cent for the CIC-density field  $f_{ES}^{(p)} = 2.9$  per cent for the DTFE-density field). This indicates that the percolating multistream filament is over 2 times thinner than that of  $\rho_{DTFE}$  and over 3 times thinner than  $\rho_{CIC}$  field.

Since the  $n_{str}$  field in this study is calculated on regular grids, the boundaries of the structures are not exactly traced. Outlining foldings in the Lagrangian sub-manifolds exactly as shown in Figure 4.12 or in the *flip-flop* calculations shown in Shandarin & Medvedev (2017) give point datasets

which are considerably more difficult to analyze. However, recent advancements in computational topology - such as the adaptation of the watershed transforms (using SpineWeb - Aragon-Calvo et al. 2008 and Morse theory (using DisPerSe - Sousbie et al. 2011 and Felix - Shivshankar et al. 2015) to inherently discrete datasets may be useful in the topological analyses of flip-flop fields and caustics.

The multistream field is a scalar function of Eulerian coordinates. We have analysed functional variation of the  $-n_{str}(\mathbf{x})$  field using Hessian eigenvalues. The Hessian analysis is generally done for inherently continuous fields, For example, Hessian analysis has been previously studied for smoothed density fields (see Sousbie et al. 2008, Aragon-Calvo et al. 2007, ?, Cautun et al. 2014 etc.), gravitational potential and velocity shear tensor (Hoffman et al. 2012, Libeskind et al. 2013, Hahn et al. 2007, Forero-Romero et al. 2009, Hoffman et al. 2012 and Cautun et al. 2014). Although the multistream field has discrete values by definition, it may be considered smooth for numerical analysis at the scale of grid length of the field. The resulting Hessian eigenvalues characterize the geometry in a four-dimensional hyper-space of  $(-n_{str}, x, y, z)$ . The boundary of a region with  $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$  is a closed convex contour in this hyper-space, and thus its projection onto the three-dimensional Lagrangian space is also closed and convex.

Dark matter haloes, being localised structures, are uniquely convenient for our local Hessian analysis. Conditions of  $\lambda_1 > 0 > \lambda_2 \geq \lambda_3$  and  $\lambda_1 \geq \lambda_2 > 0 > \lambda_3$  also give information about curvature. Hessian eigenvalue analysis at high resolution of multistream fields may be very interesting in understanding the tubular edges of filaments and surfaces of walls at smaller scales.

## 4.7 Summary

We studied certain geometrical and topological aspects of the multistream field in the context of large scale structure of the Universe. Several features were found to be considerably different from traditional density fields. The major findings from our analysis are briefly summarized as follows:

1. We use the multistream field as a proxy for distinguishing of the DM web from DM voids:

the web is defined as the regions with number of streams greater than one and thus voids as a single stream regions. The boundary between them representing a sharp transition from one- to three- stream flow regions would be a caustic surface in the density field if the mass and spatial resolutions were sufficiently high. They were clearly seen in 2D simulations by Melott & Shandarin (1989) as well as in 3D simulations by Angulo et al. (2016), Hahn & Angulo (2016), Hahn et al. (2013) and in velocity fields Hahn et al. (2015).

2. Regions without any folds in the Lagrangian sub-manifold are mostly connected. These single streaming void regions at  $z = 0$  occupy around 90 per cent of both simulations used in this study, most of which belong to a single percolating structure. However at high resolution multistream analysis, we identify a number of isolated pockets that are entirely enclosed by boundary of walls. But these voids are tiny and collectively occupy less than 0.1 per cent of the volume of the simulation box.
3. The Hessian components of the multistream field are universally zero in the interior of the void, due to constant value of  $n_{str}$ . Density field need not have zero Hessians since mass density is not unequivocally constant at  $z = 0$ .
4. We studied the global topology of the non-void ( $n_{str} > 1$ ) structure using percolation analysis. A rapid percolation transition occurred in our multistream field at  $n_{str} = 17$ . The percolating filament in multistream field is thinner than the percolating filament in mass density field.

The Lagrangian sub-manifold contains dynamical information of structure formation. We analysed the multistream field that contains the information of foldings in the sub-manifold. Connectivities in the void and non-void components of the multistream web reveal several details about structure of the Universe that are not probed by traditional density fields. In addition, we demonstrated the use of geometrical features of the multistream field in identifying potential dark matter halo candidates in cosmological N-body simulations.

## Acknowledgements

This work has been funded in part by the University of Kansas FY 2017 Competition General Research Fund, GRF Award 2301155. This research used resources of the Advanced Computing Facility, which is part of the Center for Research Computing at the University of Kansas. We thank Mark Neyrinck and Mikhail Medvedev for discussions and suggestions. We also thank the anonymous referee for insightful comments on improving this manuscript.

# Chapter 5

## Dark matter haloes: a multistream view

### 5.1 Introduction

The web-like distribution of matter initially revealed by redshift surveys (with less than 300 galaxies by Gregory & Thompson 1978 and around 1000 galaxies by de Lapparent et al. 1986) and numerical modeling (using N-body simulations of around 30000 particles by Shandarin 1983a and Klypin & Shandarin 1983) pioneered morphological investigations of the cosmic web structures (see Bond et al. 1996, also reviews by Shandarin & Zeldovich 1989 and van de Weygaert & Bond 2008a). Detailed mapping of the Universe has crossed three million objects today, by catalogues such as the Sloan Digital Sky Survey (SDSS; Albareti et al. 2016). The upcoming Large Synoptic Survey Telescope (LSST; LSST Science Collaboration et al. 2009) is expected to probe the nature of dark matter using several billion galaxies. On the other hand, cosmological simulations have improved immensely in several aspects – numerical techniques, parallelization schemes, inclusion of various physical processes, volume and resolution (some of these developments are summarized in Bertschinger (1998) and Bagla & Padmanabhan (1997). Modern state-of-the-art simulations like the Illustris Project Vogelsberger et al. (2014), the EAGLE project Schaye et al. (2015) and Q-Continuum Heitmann et al. (2015) use more than a billion dark matter particles. Finally, the ever improving data analysis techniques have resulted in new and sophisticated density estimators,

geometrical and topological indicators. A plethora of algorithms for identifying and characterizing dark matter structures have emerged in last two decades (a summary on cosmological data analysis is highlighted in van de Weygaert & Schaap 2009). Considering all these improvements, it is worth noting that the proto-structures detected in the modern simulations are qualitatively similar to the quasi-linear description of clustering by Zeldovich Approximation (ZA; Zeldovich 1970). Location and properties of these structures, i.e., the voids, walls, filaments and haloes maybe inconsistent across different structure finding algorithms, but that is primarily due to varied definitions.

Most of structure finders are halo finders only and majority of them are stemmed from three underlying algorithms. One of them is the SO (Spherical Over-density) halo finder that defines halos as spherical regions whose mass density exceeds the mean density by a specified factor (Press & Schechter, 1974). Another is the FOF (Friends-of-Friends) halo finder describing haloes as the groups of particles separated less than a specified linking length often chosen as 0.2 times the mean particle separation (Davis et al., 1985). The FOF can be also used for identifying filaments and walls/pancakes by increasing the linking length (Zeldovich et al. 1982, Shandarin 1983b, Shandarin et al. 2010). Finally the DENMAX (DENsity MAXimum) halo finder assumes that the halos are the peaks of the density fields and thus selects the particles concentrated in the vicinity of the density maxima (Bertschinger & Gelb, 1991). One of the common features of these techniques is that all three are based on density, in one form or another. And all of them depend on free parameters that are chosen chiefly on the ‘merits principle’ (Forero-Romero et al., 2009) rather than on physics. Over the years all three kinds of the halo finders have been experiencing various modifications and improvements. A few examples from a long list of these modifications may include:

- (i) Improvised techniques of generation of the density field from the particle positions, and finding spherically bound over-densities (Lacey & Cole 1994, Jenkins et al. 2001, Evrard et al. 2002, Weinberg et al. 1997, Neyrinck et al. 2005, Knollmann & Knebe 2009, Sutter & Ricker 2010, Planelles & Quilis 2010 etc.)

(ii) Adaptive methods controlling the linking length in methods using FOF (Davis et al. 1985, van Kampen 1995, Gottlober et al. 1999, Springel et al. 2001a, Habib et al. 2009, Rasera et al. 2010 etc.)

(iii) Adaptive methods for searching the positions of density maxima (For example, Klypin et al. 1999.)

(iv) Generalization of FOF and DENMAX techniques to six-dimensional phase space, and many others (such as 6DFOF by Diemand et al. 2006 and ROCKSTAR by Behroozi et al. 2013 use velocity-position space with parameters analogous to linking-length.)

(v) Computing hierarchical tree of clusters in the phase-space such as the Hierarchical Structure Finder MacLejewski et al. (2009), and the 6-D Hierarchical Over-density Tree Ascasibar (2010).

(vi) Hybrid algorithms: frameworks such as the Hierarchical Bound-Tracing algorithm Han et al. (2012) and SURV Giocoli et al. (2010) identify haloes at multiple time steps from the simulation to find prospective sub-haloes. In addition, there are HOP methods by Eisenstein & Hut (1998), Tweed et al. (2009) and Skory et al. (2010).

A detailed comparisons of several halo/sub-halo finders is provided in Knebe et al. (2011), Knebe et al. (2013), Onions et al. (2012). In a nice summary discussing these developments as well as describing a few new suggestions they concluded that there was no general consensus for a precise definition of a halo or a sub-halo. Consequently, there were different estimates of number of haloes, halo mass functions, halo centre locations, boundaries and other parameters.

There are significant concerns with SO, DENMAX and FOF algorithms - both in terms of underlying mechanisms of halo formation and the parameters used in halo identification. SO is motivated by the analytical toy model of the collapse of a top-hat spherical density perturbation. Parameters of the virial radii  $r_{vir}$  and virial mass  $M_{vir}$  are determined by the regions with density  $\rho_{vir} \geq \Delta_{vir} \times \rho_b$ , where  $\rho_b$  is the background density of the simulation box.  $\Delta_{vir}$  is generally taken around 180 or 200, derived for an isolated spherical collapse model, and it varies for different cosmologies and redshift. The peaks in CDM models not only aspherical, but their collapse is subject to tidal forces, mergers and presence of sub-structures - none of these complexities are

weighed in the spherical collapse model.

For FOF, the free parameter of linking length is generally taken as  $b = 0.2$  times the mean separation of particles at  $z = 0$ . This inter-particle separation corresponds to  $\Delta_{vir} \approx 180$  if the halo has an isothermal density profile,  $\rho \propto r^{-2}$ . Using percolation theory, More et al. (2011) argued that this corresponds to a rather wide range of over-densities depending on halo mass and density profiles. They found out that  $b = 0.2$  corresponds to local over-density  $\delta$  within the enclosed halo to be in the range of 250 to 600. Moreover, the resulting FOF-haloes need not have a compact geometry: often the haloes are irregularly shaped, which is unlikely if the halos are virialized. Hence modern algorithms re-define the halo boundaries by excluding particles using post-processing techniques. In recent simulations with clear delineation of walls and filaments Angulo et al. (2013c),  $b = 0.05$  was used for finding FOF-haloes since the traditional value of  $b = 0.2$  corresponded to structures that percolate into the web structure.

Absence of dynamical traits in the FOF and SO algorithms are arguably more crucial. In phase-space, the halo collapse models show collisionless DM particles in oscillatory motions about a core, at successive foldings of the phase-space sheet. The velocity field within each oscillatory spiral is multi-valued in physical space. Incrementing *multistream* shells, separated by *caustic* surfaces sequentially trace the structures of the cosmic web - walls, filaments and the haloes. Majority of the mass accretion into the haloes along the filaments: from lower multistreams into higher. Thus the DM haloes are not independent of filaments around them, and the hierarchical layers of multistreams represents this portrait precisely. This picture of structure formation was initially theorized using ZA Zeldovich (1970) and in context of caustics Arnold et al. (1982) as well as in the Adhesion Approximation (Gurbatov, S.N., Saichev, A.I., Shandarin et al. 1989, Kofman et al. 1992). Shandarin & Zeldovich (1989) reviewed gravitational evolution of density perturbations in the context.

It has been demonstrated that the multistream field in Eulerian space can be computed directly from the Lagrangian sub-manifold (Shandarin et al. 2012 and Abel et al. 2012). About 90% of the field is single-streaming voids, and the rest of the volume comprises of multistream walls, fil-

aments and haloes. Ramachandra & Shandarin (2015) found the multistream value of  $n_{str} \approx 90$  corresponds to virial density  $\Delta_{vir} = 200$ . On the other hand, DM particles are identified by (Falck et al., 2012) as belonging to haloes if they undergo flip-flop along 3 orthogonal axes. These analyses have opened up a new avenue in studies of halo formation, both qualitative and quantitative. Re-investigations of halo spins, physical radii of the halo, sub-structure in the light of streaming phenomena have shown that the halo structures and formations are more complicated than previously believed. Vogelsberger & White (2011) investigated the distribution of streams in small haloes at various redshifts. They concluded that tracking caustics and streams is better than density, since density fields are noisy in the dense inner regions of haloes. In another study, More et al. (2015) argued that the ‘splashback radius’ - distance from the halo core to the first caustic enumerated from outside - is a better physical indicator of DM halo boundary than the virial radius (also see the discussion on turn-around radius of bound objects by Lee & Yepes 2016). Angulo et al. (2013c) also agree with the view that the locally overdense regions correspond better with the volumes within the first caustic than the virialized DM clumps. Recent toy model of anisotropic halo collapse by Neyrinck (2016) considers intersecting multistream filaments forming spinning nodes. Ramachandra & Shandarin (2017) showed that the virial surfaces of FOF haloes have varying number of streams, including single-streams. Study by Shandarin & Medvedev (2017) delineated the rich sub-structure of haloes using another derivative parameter from the Lagrangian sub-manifold called the ‘flip-flop’ defined on the Lagrangian space.

In this paper, we identify potential haloes by utilizing multistreaming as the governing dynamical phenomena. A review of the DM particle clustering in a one-dimensional dimensional universe is made in Section 5.2, and the concept of multistream field is extended to higher dimensions. The multistream field is computed on the cosmological simulations described in Section 5.3. The halo identification framework in this field is described in Section 5.4. This algorithm isolates convex regions of the multistream field using Hessian eigenvalues, each enclosing a local multistream maximum. Without employing any non-local thresholds that several halo finders generally use, these convex multistream regions are identified as potential halo sites. We also illustrate the sig-

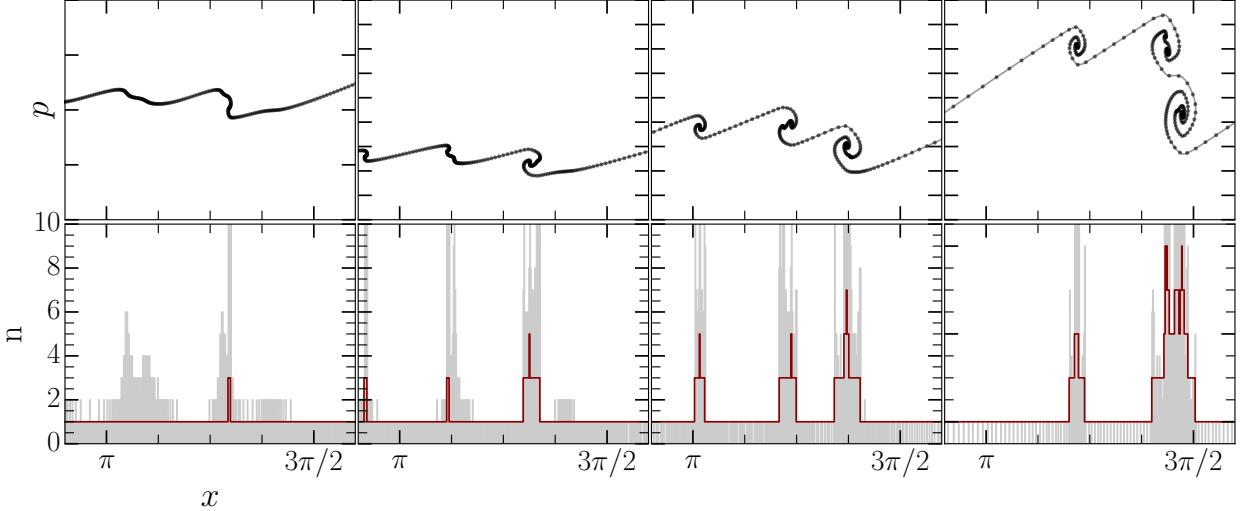


Figure 5.1: Dynamical collapse of dark matter in one-dimensional universe: Top panels show the  $(\mathbf{p}, \mathbf{x})$  phase-space manifold of the dark matter sheet at redshifts  $z_1, z_2, z_3$  and  $z = 0$ . Dots represent the dark matter particles. The momentum values are chosen at arbitrary scales. Bottom panels show the corresponding multistream field multistream field  $n_{str}(\mathbf{x}, z)$  (red) and density field  $\rho(\mathbf{x}, z)$  (gray).

nificance of multistream refinement and softening scales in finding subhaloes. However, this paper does not focus on adaptive multi-scale analyses for substructure studies. A few properties of the multistream haloes are discussed in Section 5.5, and comparison of these haloes with AHF and FOF algorithms is done in Section 5.6. We also discuss the spatial distribution of the dark matter haloes along the percolating web structure.

## 5.2 Phase-space representation of gravitational clustering

We begin with a simple illustration showing the formation of a few haloes in a one-dimensional simulation. Dark matter clustering in a (1+1)-dimensional phase-space  $(\mathbf{p}, \mathbf{x})$  (where  $p$  is the momentum and  $x$  is the co-moving Eulerian coordinate) at four successive time steps is shown in the top panels of Figure 5.1. The lower panels show the corresponding multistream field (Shandarin et al. 2012 and Abel et al. 2012)  $n_{str}(\mathbf{x}, z)$  (red) and density field  $\rho(\mathbf{x}, z)$  (gray). At  $z_1$  (left-most panel), velocity is single-valued in Eulerian co-ordinates shown, except at a small three-stream region near  $\mathbf{x} = 5\pi/4$ . This is the first instance of multistreaming in the region, which was previously

had  $n_{str} = 1$  throughout. The interface of  $n_{str} = 1$  and  $n_{str} = 3$  regions is also the location of the first caustic. On the other hand, the density calculated at a high resolution shows variations, even in the mono-streaming regions. The variations are especially more pronounced around the caustic (near  $\mathbf{x} = 5\pi/4$ ).

The gravitational clustering is more evolved in the two center panels ( $z_2$  and  $z_3$ ) with three prominent phase-space spirals. The regions between the spirals have sparsely distributed dark matter particles, and have  $n_{str} = 1$ . Each spiral corresponds to a location of gravitational collapse with  $n_{str} > 1$  region, and higher density. A few of these regions within three-streaming regions are elevated to  $n_{str} = 5$ . The corresponding density field is not only noisier, but also reaches very high values at the caustics. This is also a primary distinguishing feature between mass density fields and multistream fields: At the locations of caustic, the density (regardless of how it is calculated) is not smooth Vogelsberger & White (2011). Computational limitations on simulation resolutions and refinement of density calculations soften the fields, exceptionally at the zero volume measure regions of caustic surfaces. On the other-hand, multistream values are increased by finite values at caustic surface locations - This property is preserved at higher simulation resolutions and any refinements of multistream field calculations - although  $n_{str}$  may be resolved enough to have intermediate even-values. Multistream fields are also intrinsically discrete valued, which is not true with density fields. Discreteness of multistream fields is discussed in more detail in Ramachandra & Shandarin (2017).

The right-most panel in Figure 5.1 shows the final structure at  $z = 0$ . Two large spirals have spatially merged. These collapse environments are naturally very complex, with an increased number of successive caustic formation and merging. The corresponding velocity streams also show a more complicated structure. Clearly, the multistream field has a saddle point that is not as low as  $n_{str} = 1$ . This poses a bigger problem in the context of most of halo detection algorithms, and we discuss this in Appendix C.

### 5.2.1 Collapse in higher dimensions

Extending the above results of one-dimensional collapse into higher dimension is vital, primarily in the context of halo formation. The individual spiral collapses in the one-dimension happen at a small region (left-most panel in Figure 5.1), and the region grows by volume, whilst increasing the spiral twists within. This is in contrast with the theoretical top-hat spherical model of halo formation when the shell crossing would not happen until the final moment of the collapse of the entire halo into a point-like singularity. Thus all shell crossings happen at a single point and at a single instant of time. The collapse of an isolated, spherically symmetric density peak is a very exceptional case, because every spherical shell feels only the force due to interior mass until it collapses into the caustic region. The collapse of the real peak proceeds in the field generated by the mass distribution - in both the mass within the forming halo, and the mass outside the halo.

The collapse of a uniform ellipsoid also results in a simultaneous collapse of the entire ellipsoid however this time not into a point but into a two-dimensional ellipse (Lin et al. 1965, Icke 1973, Eisenstein & Loeb 1995). Another customarily used spherical model of halo formation by Fillmore, J. A., Goldreich (1984) and Bertschinger (1985) does not consider the initial collapse at all. Instead it assumes self-similar initial conditions and the halo at advanced stage with formally infinite number of spherical caustic shells.

The ‘core’ in a collision-less dark matter collapse (in Figure 5.1) is a region where a multistream region is first formed due to caustic formation. This is conceptually similar to a shell crossing. However, there are successive caustic formations that follow the first shell crossing, and they are not limited to the halo cores. Each caustic increases the multistream value within by a finite number. The cores of the multistream haloes obviously have the local maxima of velocity streams in Eulerian coordinates. On the contrary, mass densities have infinite values at the caustics surfaces, including the core. Discontinuities in densities at these regions of sharp multistream transitions are clearly seen if the mass and spatial resolutions were sufficiently high( see two-dimensional simulations by Melott & Shandarin 1989 as well as in three-dimensional simulations by Hahn et al. 2013, Angulo et al. 2016, Hahn & Angulo 2016 etc.).

In three-dimensional simulations, the Lagrangian sub-manifold twists in complicated ways in a six-dimensional phase space. This is due to complexities involving caustic formations in higher dimensions, which is true even in the ZA, see Arnold et al. 1982 and Hidding et al. 2014 for detailed analyses of caustic formation. The resulting intricate geometrical structures can be characterized by the  $n_{str}$  field. Nearly 90% of the volume in N-body simulations are single-streamed voids at  $z = 0$  (Shandarin et al. 2012, also see Falck & Neyrinck 2015 for a percolation analysis of single-streaming voids). From the visualizations in Ramachandra & Shandarin (2015) and percolation analysis of Ramachandra & Shandarin (2017), we also know that the  $n_{str} = 3$  regions mostly form connected wall-like structures, unlike the isolated patches as seen in one-dimensional simulations of Figure 5.1. The structures become predominantly filamentary at higher thresholds of  $n_{str} \gtrsim 17$ . Subsequently, the regions around the multistream maxima have isolated closed surfaces (for example, in Figure 5.2), which may be identified as halo locations.

Caustic formations and mass accretion are also seen to occur more along the higher streams, which makes the haloes non-spherical, with the alignment generally determined by a complicated interplay of the intensities of the streams from neighboring filamentary structures. Number of streams corresponding to the dark matter halo also has a local environment dependence. The three small haloes in Figure 5.2, where the number of streams are higher than the neighbouring filaments, are aligned along three intersecting filaments. Halo environment studied in Ramachandra & Shandarin (2015) show similar hierarchical variation in  $n_{str}$  values. The halo environments are thus very complicated, and empirical thresholds (on multistream or density fields) may not account for all the haloes uniformly. Hence we use a local geometrical method to identify potential haloes in multistream fields.

The first non-linear structures in the web have  $n_{str} = 3$ . By visual inspection, these regions generally form a fabric-like open structures that resemble walls. The surface contours of higher  $n_{str}$  are embedded within the walls. Figure 5.2 shows a filamentary structure of the web at  $n_{str} \gtrsim 17$ . The figure also shows regions around local maxima of the multistream field, which are generally located at the intersections of filaments.

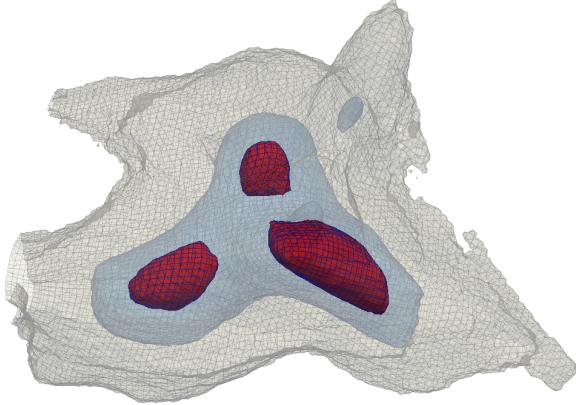


Figure 5.2: Multistream field contours: The multistream field is calculated at 8 times the native resolution. Each inner convex blobs (red) surround local multistream maxima inside. Surrounding outer shell(blue) is not convex throughout the surface, and the outermost gray multistream surface displays a filamental geometry.

### 5.3 Simulations and tools

The emphasis of this paper is to demonstrate the use of multistream field in identifying potential dark matter haloes, and not a full statistical analysis of halo properties. For this purpose, we have run simulations at two different mass resolutions (number of particles  $N_p = 128^3$  and  $256^3$ , and respective mass of particles,  $m_p = 3.65 \times 10^{10} h^{-1} M_\odot$  and  $4.57 \times 10^9 h^{-1} M_\odot$ ), with the same periodic side length  $L = 100 h^{-1} Mpc$ . The gravitational softening length  $\epsilon = 20 h^{-1} kpc$  and  $10 h^{-1} kpc$  for low and high resolution simulation respectively. The initial conditions are generated by MUSIC (Hahn & Abel, 2011) with the transfer function from Eisenstein & Hu (1998) at a redshift of  $z_{ini} = 80$ . The  $\Lambda$ CDM cosmological simulation is run using GADGET-2 (Springel 2005 and Springel et al. 2001b) is similar to the ones used in Ramachandra & Shandarin (2017). The cosmological parameters used in the simulation are  $\Omega_m = 0.276$ ,  $\Omega_\Lambda = 0.724$ , the Hubble parameter,  $h = 0.703$ , the power spectrum normalization,  $\sigma_8 = 0.811$  and the spectral index  $n_s = 0.961$ .

Multistream field  $n_{str}(\mathbf{x})$  is calculated on the GADGET-2 snapshots at  $z = 0$  using the tessellation scheme by Shandarin et al. (2012). The multistream field can be computed at the native resolution of the Lagrangian grid of the simulation, i.e., at refinement factor of  $l_l/l_d = 1$  (where  $l_l$  is the inter-particle separation in Lagrangian grid and  $l_d$  is the side length resolution of diagnostic

grid). Arbitrarily high refinement factors can be utilized in computing multistream fields as well, for example  $l_l/l_d = 8$  for the halo multistream environment shown in Figure 5.2. For analysis of full simulation boxes, we restrict  $l_l/l_d$  to 1 and 2.

Two halo finders are also used to identify potential haloes with 20 or more particles at  $z = 0$ : a classic Friends-of-Friends method (FOF-Davis et al. 1985) using a popular linking length,  $b = 0.2$  (e.g. Frenk et al. 1988 and Lacey & Cole 1994) and the Adaptive Mesh Investigations of Galaxy Assembly (AMIGA halo finder or AHF-Knollmann & Knebe 2009, Gill et al. 2004). Halo catalogue from these halo finders are used to compare with our implementation of halo detection in the multistream field. The haloes candidates from AHF and FOF algorithms are hereafter referred to as AHF-haloes and FOF-haloes respectively.

## 5.4 Haloes in the multistream field

We intend to identify haloes in the  $n_{str}(\mathbf{x})$  field instead of using just the position coordinate data. While the eigenvalue analysis itself is done at a chosen time, the multistream field inherently has data from six-dimensional Lagrangian space ( $\mathbf{q}, \mathbf{x}$ ) that contains the full dynamical information, similar to the phase-space sheet albeit in a different form. Dynamical history that is embedded in the multistream field is crucial in understanding the strength of gravitational binding of the particles. A physically motivated distinction between void and gravitationally collapsed regions – voids are the regions with a single stream – is a unique feature of multistream analysis (Shandarin et al. 2012 and Ramachandra & Shandarin 2017). Thus the haloes detected from local maxima of the  $n_{str}$  field can be ensured to be away from the mono-streaming voids. Methods based on linking-length or density fields may not be able to ensure that all the particles in haloes are away from voids (as shown for FOF haloes in Ramachandra & Shandarin 2017).

Numerical analyses of scalar fields generally depend on resolution as opposed to particle coordinates analysis tools like FOF. The multistream field conveniently has an advantage of being less noisy than mass density (Shandarin et al. 2012, also see the Appendix in Ramachandra &

Shandarin 2017 ).

### 5.4.1 Hessian of multistream fields

Hessian matrix  $\mathbf{H}(f)$  of a scalar field  $f$  involves local second-order variations in three orthogonal directions. Each element of the Hessian matrix  $\mathbf{H}_{ij}(f)$  (where  $i$  and  $j$  can be any of  $x$ ,  $y$  or  $z$  directions) is given by Equation 5.1.

$$\mathbf{H}_{ij}(f) = \frac{\partial^2 f}{\partial x_i \partial x_j} \quad (5.1)$$

By choosing a  $f = -n_{str}(x)$  (smoothed if necessary), local multistream variations can be diagnosed. The Hessian matrices at each point on the configuration space are always symmetric matrices, resulting in real eigenvalues. The Hessian eigenvalues in multistream fields differ from that in density, gravitational potential or velocity shear tensor. We refer the readers to Ramachandra & Shandarin (2017) for an extensive analysis on multistream Hessians and their geometrical significance. Some of the salient features of Hessian eigenvalues of multistream field are as follows:

- (i) Every element of Hessian matrices  $\mathbf{H}(-n_{str})$ , and consequently the eigenvalues  $\lambda_i$ 's are zero in single-stream voids. Even if the multistream field is a smoothed, the eigenvalues peak at zero. This property is unique to multistream fields. Eigenvalues of Hessians of density Aragon-Calvo et al. (2007), velocity shear tensor Libeskind et al. (2013) do not peak at zero, and the eigenvalues of deformation tensor are negative in voids as a result of continuity equation (shown in Zel'dovich formalism as well).
- (ii) The eigenvalues of these Hessian matrices are always real, and depending on if their values are positive or negative, one may infer local geometrical features in the field. For our choice of  $-n_{str}(x)$  as the domain of Hessian, at least in principle, the conditions for geometric criteria are:  $\lambda_1 > 0 > \lambda_2 \geq \lambda_3$  for locally flat regions,  $\lambda_1 \geq \lambda_2 > 0 > \lambda_3$  for locally tubular structures and  $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$  for clumped blobs.

- (iii) Convex neighbourhoods around local maxima of the multistream field are isolated by the positive definite Hessian matrices.
- (iv) The resulting Hessian eigenvalues characterize the geometry in a four-dimensional hyper-space of  $(-n_{str}, x, y, z)$ . The boundary of a region with  $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$  is a closed convex contour in this hyper-space, and thus its projection on to the three-dimensional Lagrangian space is also closed and convex.

Of the three geometries that are characterised by the eigenvalue conditions, we investigate the convexity of multistream contours in the context of halo finding in the section below.

### 5.4.2 Halo finder algorithm

Our goal is to isolate the locations of convex geometries in the multistream flow field. Prospective regions of the halo locations in the web structure are selected by positive definite condition of the Hessian  $\mathbf{H}(-n_{str})$ :  $\lambda_1 > 0$ ,  $\lambda_2 > 0$  and  $\lambda_3 > 0$ , or simply the smallest of the eigenvalues,  $\lambda_3 > 0$ . We also filter out the regions if the multistream values inside do not suggest gravitational collapse into haloes. The sequence of our halo detection framework is listed below:

1. The multistream flow field is calculated on a diagnostic grid. The number of tetrahedra that encompass each vertex in the grid gives the  $n_{str}$  field. Top left panel of Figure 5.3 shows the multistream web structure in a slice of the simulation with  $n_{str} > 1$  in gray and  $n_{str} \geq 7$  in blue.
2. The discrete multistream flow field is smoothed in order to reduce numerical noise. We have used Gaussian kernel for smoothing in our analysis. Effect of smoothing scales in the halo identification is discussed in Section 5.4.3.
3. Second order variations of the smoothed  $-n_{str}(\mathbf{x})$  is computed at each point in the field. This gives symmetric Hessian matrices for this field whose eigenvalues are real. Ordered eigenvalues of the Hessian,  $\lambda_1 \geq \lambda_2 \geq \lambda_3$  are calculated. The  $\lambda_3$  field is shown in the top right panel of Figure 5.3.

4. Using segmentation techniques, each region with  $\lambda_3$  strictly greater than 0 within  $n_{str} \geq 3$  regions of multistream field are isolated and labelled. This condition for each halo candidate guarantees that it is in the region where at least one gravitational collapse happened within the halo boundary. Mass particles belonging to these regions are shown as dark spots in the top right panel of Figure 5.3.
5. The multistream field has a range of values within the isolated sites. We impose constraints on the isolated regions to rule out the labels with low multistream values. The local maxima of  $n_{str}$  inside each halo must be at least 7. By this restriction, it is ensured that the halo sites with three Lagrangian sub-manifold foldings are selected. Bottom left panel of Figure 5.3 shows patches that are discarded in red. The resulting  $\lambda_3$ -haloes are shown in the bottom right.
6. In our comparisons with other halo finders in Section 5.6, we also used an additional constraint on the minimum number of mass particles in the haloes to be 20 - which is generally used as a criteria in several halo finders.

For the illustration halo detection framework in this section, we have calculated the number-of-streams at refinement factor of 2 and smoothing scale of  $0.39h^{-1}Mpc$  (equal to the grid length of the multistream field) for the simulation box of  $128^3$  particles and size  $L = 100h^{-1}Mpc$ . Hessian matrices and eigenvalues are calculated on the same diagnosis grid. Results of the halo detection scheme for simulation box of higher mass resolution, and different smoothing factors are discussed in Sections 5.6 and 5.4.3. Hereafter we refer to the potential dark matter haloes detected from the Hessian analysis of the multistream field as  $\lambda_3$ -haloes for brevity.

Applying the above scheme on the simulation with side length of  $100h^{-1} Mpc$  and  $128^3$  particles (with cosmological parameters mentioned in Section 5.3), we detected approximately 50000 regions satisfying  $\lambda_3 > 0$  within the non-void in the multistream field of refinement factor  $l_l/l_d = 2$  and smoothing scale of grid length, i.e,  $0.39h^{-1}Mpc$ . We filtered out the segments with local maxima of  $n_{str} < 7$ , and around 14000 regions remained as prospective haloes. Majority of these

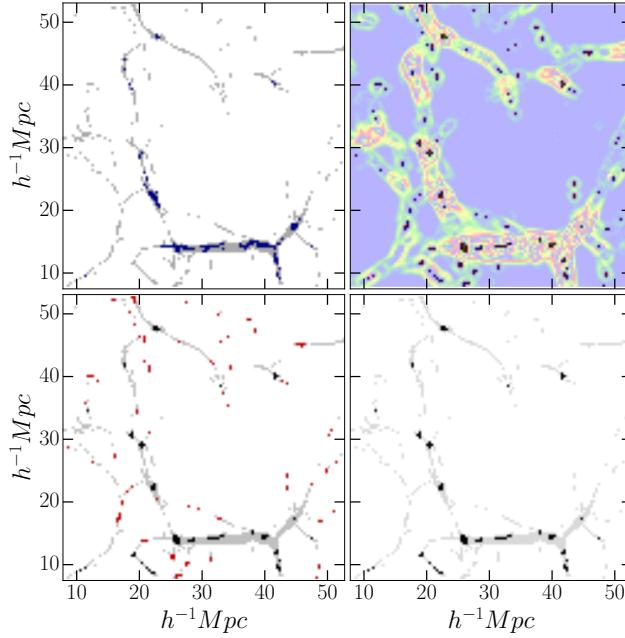


Figure 5.3: Detection of potential halo candidates in the multistream field: algorithm of segmentation and filtering are illustrated in a smaller slice of  $40h^{-1}Mpc \times 40h^{-1}Mpc$  slice of the simulation box. Top left figure shows the multistream field of the slice. Voids (white) are the regions with  $n_{str} = 1$ , rest are non-void structures. Blue patches within the structure (gray) are the regions with gravitational collapses in more than one direction, i.e.,  $n_{str} \geq 7$ . Top Right figure shows the smallest eigenvalue  $\lambda_3$  field. The value of  $\lambda_3$  is close to 0 in most of the regions (yellow), including the voids. Regions with  $\lambda_3 > 0$  and  $n_{str} > 1$ , are isolated (black spots) using image segmentation techniques. Bottom left panel shows the filtering scheme: the red patches do not have maxima of  $n_{str} \geq 7$  in the regions, hence are filtered out. The remaining potential halo regions with more than 20 particles are shown in the bottom right panel.

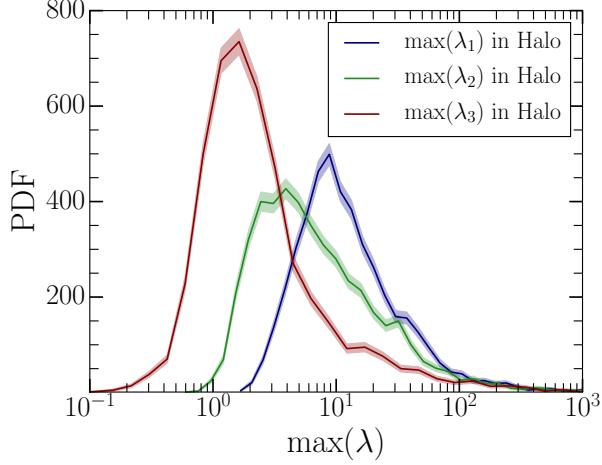


Figure 5.4: PDF of highest  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  values in each of 4492 haloes detected by our algorithm. The peaks of the PDF are in the range 1-10. Shaded regions represent  $1\sigma$  error.

Table 5.1: Statistics of the Hessian eigenvalues in the halo candidates

Statistics	$\lambda_1$	$\lambda_2$	$\lambda_3$
Minimum	1.5	0.5	$1.3 \times 10^{-2}$
Maximum	$1.7 \times 10^3$	$1.5 \times 10^3$	$1.1 \times 10^3$
Median	10.5	5.5	1.9

regions have less than 20 particles, which are excluded in the halo catalogues. On the whole, our algorithm detected about 4500 haloes with more than 20 particles in the entire simulation box. We have not applied virialisation to define the halo boundaries. A more detailed study of halo edges, and comparison with that of FOF-haloes and AHF-haloes is done in Section 5.5. Here we concentrate on the three vital factors in our framework: local geometrical indicators  $\lambda_i$ 's, the softening scale of the field and multistream thresholds.

The maximum values of  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  in each of the haloes have peaks away from 0 as shown in Figure 5.4. The median values of  $\max(\lambda_1)$  and  $\max(\lambda_2)$  are in the range of 1-10 (Table 5.1), in spite of the threshold for  $\lambda_3$  being barely positive, by definition. Hence the interior of the potential halo segments is quite convex, with a local maxima inside. In some haloes, the local maxima of eigenvalue are in the order of thousands, as tabulated in Table 5.1.

With this algorithm, we obtain prospective dark matter haloes - regions with a local maximum of the multistream field in the interior of their closed convex surfaces. The haloes are detected

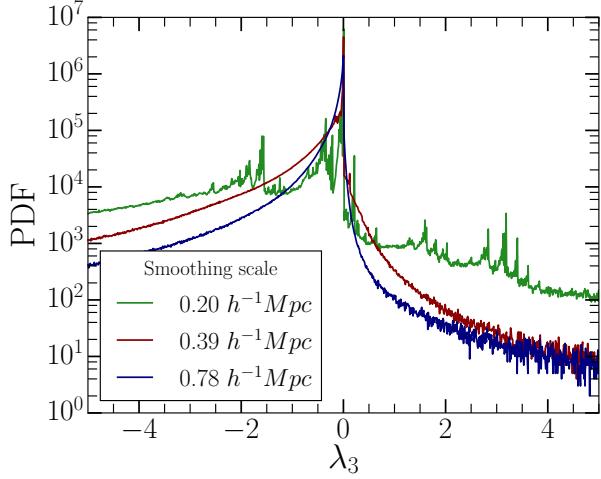


Figure 5.5: The distribution of  $\lambda_3$  in the simulation box of  $128^3$  particles and multistream field of refinement factor  $l_l/l_d = 2$ . Three smoothing scales are shown.

without using density fields or linking lengths between particles. The parameters in the algorithm are entirely based on features of the multistream field and local geometry using Hessian matrices.

### 5.4.3 Effect of smoothing

In order to reduce noise, the field is smoothed for our analysis using a Gaussian filter. The effect of smoothing scale on the distribution of the eigenvalue  $\lambda_3$  in the simulation of  $128^3$  particles is shown in Figure 5.5. Effect of softening on the multistream fields does not alter the distribution of multistream distribution significantly (Seen in Figure 10 of Ramachandra & Shandarin 2017). However, the second order variation (and consequently the Hessian eigenvalues) is significantly changed due to the softening of the edges of structures. PDF of  $\lambda_3$  at multistream smoothing scale of the half the side length of diagnostic grid ,  $0.5 \times l_d = 0.20h^{-1}$  Mpc is noisier than in the scales of  $l_d$  and  $2 \times l_d$ . However, at every scale, the PDF peaks at 0. The volume fraction of regions with  $\lambda_3 > 0$  (i.e. with positive curvature) is 2.4%, 2.3% and 2.5% for scales  $0.20h^{-1}$  Mpc,  $0.39h^{-1}$  Mpc,  $0.78h^{-1}$  Mpc respectively. For the detection of haloes in Section 5.4, we only look at these regions.

In addition to reducing the numerical noise, smoothing of the multistream field also results in softening of the sub-structures Ramachandra & Shandarin (2017). Since our framework of

Table 5.2: Number of  $\lambda_3$ -haloes identified at smoothing of  $n_{str}(\mathbf{x})$  at different scales.

$N_p$	$0.20h^{-1}\text{Mpc}$	$0.39h^{-1}\text{Mpc}$	$0.78h^{-1}\text{Mpc}$
$128^3$	5181	4492	2923
$256^3$	27929	18221	7897

detecting haloes isolates the multistream regions with local maxima, the closed curvatures are resolved separately. The halo or a sub-halo regions, that enclose the local maxima of the  $n_{str}$  field, vary with the smoothing scale of the multistream field. By increasing smoothing of the multistream field, the number of haloes are reduced as shown in Table 5.2. In the simulation with  $256^3$  particles, 27929  $\lambda_3$ -haloes are detected at smoothing scale equal to the diagnostic grid length,  $l_d = 0.20h^{-1}\text{Mpc}$ . The number of haloes decreases to 18221 and 7897 at softening scales of two- and four times the grid lengths respectively.

Moreover, since the spatial resolution is higher at the low softening, more small haloes are detected, as shown in lower mass regime of halo mass functions in Figure 5.6. The tail of halo mass functions reveal that large haloes are more massive for higher softening scales. For instance, the largest haloes for the same simulation with multistream softening length of  $0.20h^{-1}\text{Mpc}$ ,  $0.39h^{-1}\text{Mpc}$  and  $0.78h^{-1}\text{Mpc}$  have 30650, 38333 and 56257 particles respectively.

The sub-halo finder methods (see Onions et al. 2012 and references therein) identify substructures within a large host halo. The sub-haloes are resolved individually as  $\lambda_3$ -haloes at different scales from our algorithm if the local maxima of the smoothed multistream field is enclosed within the boundary.

#### 5.4.4 Effect of multistream thresholds

Environmental dependence of the haloes results in various multistream values for the halo core. Theoretical toy models of halo formation, such as the tetrahedral collapse model Neyrinck (2016) describes a three-dimensional halo with four filaments accreting mass into it, has 15 stream crossings. Ramachandra & Shandarin (2015) have previously showed that a high threshold of  $n_{str} \geq 90$  is equivalent of virial density of  $\rho_{vir} = 200$ , and filters most of the large haloes above  $10^{13}M_\odot$ .

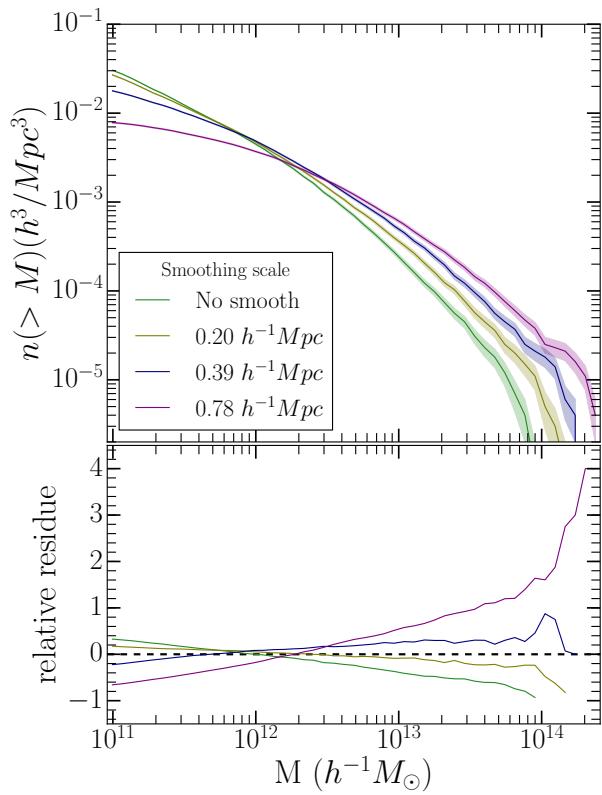


Figure 5.6: Top panel shows Halo Mass functions of the potential  $\lambda_3$ -haloes in the multistream field refinement factor  $l_l/l_d = 2$  with various smoothing scales. Simulation box has  $256^3$  particles. Lower panel shows the deviation of the each halo mass function with respect to their average.

The algorithm used for detecting multistream haloes initially detects all the closed regions in the multistream ( $n_{str} > 1$ ) regions of the cosmological simulation. In order to exclude some of the obvious non-halo sites, we impose a lower threshold of  $n_{str} \geq 7$  on the multistream maximum (these regions were also seen as parts of walls or filaments in Ramachandra & Shandarin 2015 ), so that all the sites with three or more foldings in the Lagrangian sub-manifold are chosen. Combining this with the conditions on local eigenvalues, number of particles in haloes etc, we got a pretty good correspondence with other halo finders as demonstrated in section 5.6 .

Although this condition is by no means strict, it is necessary to check the validity of the assumption. Figure 5.7 shows the halo mass functions for the haloes detected with changing thresholds on the multistream values of the halo cores. The figure demonstrates that increasing the cut-off from  $n_{str} \geq 3$  to  $n_{str} \geq 25$  systematically excludes small mass haloes while the number of haloes with  $M \gtrsim 2 \times 10^{12} M_\odot$  remains the same.

## 5.5 Halo properties

Multistream environment of haloes can be very diverse. Ramachandra & Shandarin (2015) demonstrated that a majority of the FOF-haloes are in contact with the single-streaming voids. Illustration in Ramachandra & Shandarin (2017) also shows that a large number of FOF-haloes have more than 10 per cent void on the spherical surfaces with virial radii. The  $\lambda_3$  haloes are significantly different: none of the  $\lambda_3$ -haloes are in contact with the regions where gravitational collapse has not occurred. This is guaranteed by the lower bound of  $n_{str} = 3$  on all potential halo candidates. Condition on the multistream field within the potential halo sites also ensures that there are collapses along more than one direction, which corresponds to  $n_{str} = 7$ . Hence by definition, for any multistream halo  $H_i$ , highest and the lowest multistream value are  $n_{str}^{high}(H_i) \geq 7$  and  $n_{str}^{low}(H_i) \geq 3$  respectively.

The potential haloes  $H_i$ s selected by eigenvalue condition  $\lambda_3 > 0$  have a local maxima of  $n_{str}^{high}(H_i)$  inside their boundaries. For a large number of these  $\lambda_3$ -halo candidates, the maximum  $n_{str}^{high}$  is higher than the bound of  $n_{str} \geq 7$ , as shown in Table 5.3 and Figure 5.8. For simulation with

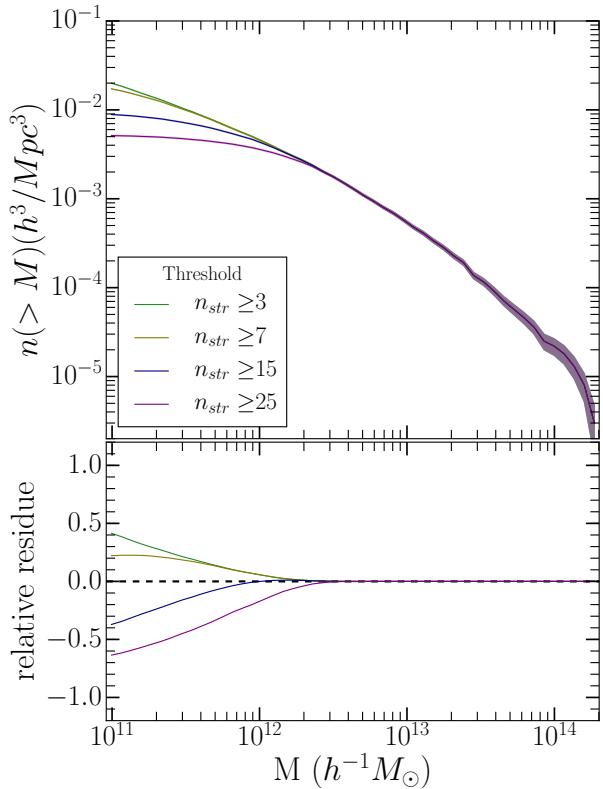


Figure 5.7: Halo Mass functions of the potential  $\lambda_3$ -haloes in the multistream field refinement factor  $l_l/l_d = 2$  (in Simulation box with  $N_p = 256^3$ ) with various thresholds on local maxima of  $n_{str}$  within the halo. Lower panel shows the deviation of the each halo mass function with respect to their average.

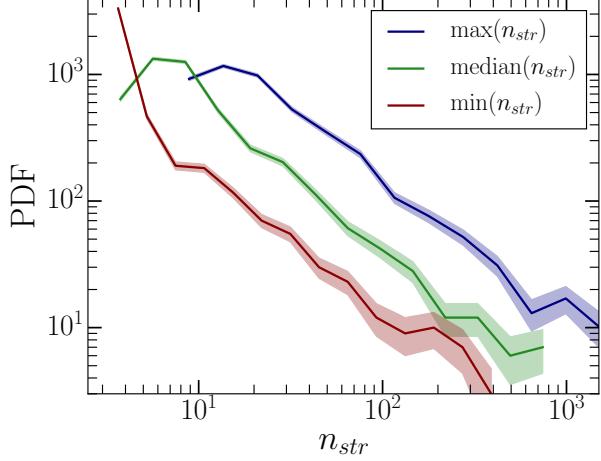


Figure 5.8: Maximum, minimum and median of  $n_{str}$  in each of 4492 halo candidates. The closed contours of haloes encompass a wide range of multistream values. None of the haloes are in contact with the void region, since lowest value of  $\min(n_{str})$  is 3. Shaded regions are the  $1\sigma$  absolute errors in the number of  $\lambda_3$ -haloes.

Table 5.3: Local maxima and minima of  $n_{str}$  in each of 4492 haloes. The highest  $n_{str}$  values in the interior of haloes span over a large range of values. Low values of  $n_{str}$  in haloes, which are generally near halo boundaries, have a median of 3.

Statistics	$n_{str}^{high}(H_i)$	$n_{str}^{low}(H_i)$
Minimum	7	3
Maximum	2831	459
Median	17	3

128<sup>3</sup> particles, the median of this peak  $n_{str}^{high}(H_i)$  value is 17. Unsurprisingly, the global maximum of the multistream field ( $n_{str} = 2831$ ) is also a local maximum for one of the haloes. On the other hand, the median of lowest multistream value  $n_{str}^{low}(H_i)$  in the haloes is 3 (Table 5.3), which is also the first stage of non-linearity.

An important feature of our halo detection method is that the detected  $\lambda_3$ -haloes do not have a global threshold on  $n_{str}$  or mass density values. The local conditions may be more suited in identifying haloes in multistream fields, since the multistream environments around haloes are very diverse. For instance, regions with  $n_{str} \geq 25$  are tubular around one of the large haloes in Figure 5.9. Even the region with more than 75 streams does not enclose a convex multistream region. Whereas, for  $n_{str} \geq 200$  the region is convex and the particles detected by our method reside mostly within. We detect closed regions in the multistream field as long as they are not in

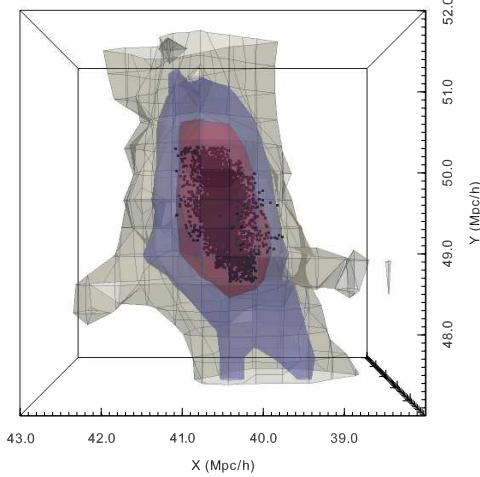


Figure 5.9: Multistream environment of a  $\lambda_3$ -halo. The contours represent regions with 3 different multistream values: Outermost  $n_{str} \geq 25$  (gray) is tubular, The blue region has  $n_{str} \geq 75$ . The inner region is highly non-linear with  $n_{str} \geq 200$ . The black dots represent the mass particles belonging to a  $\lambda_3$ -halo, as detected by our algorithm.

void, and have at least three foldings in the Lagrangian sub-manifold.

However, the  $\lambda_3$ -halo boundary is different from any constant multistream contour. That is, from the function  $n_{str}(\mathbf{x})$ , convex regions in the four-dimensional function space  $(-n_{str}, x, y, z)$  are projected onto three-dimensional co-ordinate space using eigenvalues. This is different from projecting ‘iso-multistream’ slice onto three-dimensional co-ordinate space. Appendix C illustrates the difference in the two approaches for a one-dimensional function.

The multistream field usually has concentric shells in the regions of successive gravitational collapses (as explained in Section 5.2 and Appendix in Ramachandra & Shandarin 2017). In a specific scenario of Figure 5.9, regions of lower number-of-streams ( $n_{str} = 25$  and lower) is are tubular and have regions of higher  $n_{str}$  inside ( $n_{str} = 200$  and higher) that is closed. However, this transition from concavity to convexity of the multistream field does not occur at a constant value of  $n_{str}$  throughout the field. Instead, it is a local geometrical change that occurs at  $\lambda_3 = 0$ . For the  $\lambda_3$ -haloes in our simulation ( $N_p = 128^3$ ), minimum of multistream values  $n_{str}^{low}(H_i)$  within each halo has a range of values shown in Table 5.3 and Figure 5.8 – this varies between  $3 \leq n_{str}^{low}(H_i) \leq 459$ . Hence a global condition on  $n_{str}$  does not guarantee that the region is convex.

The particles in a massive  $\lambda_3$ -halo shown in Figure 5.9 form a spheroidal structure. The number of particles in similar massive haloes are in the order of  $10^3 - 10^4$  particles. For instance, the most massive halo in the simulation (with  $N_p = 128^3$ ) has 5593 particles. We have chosen a minimum threshold of 20 particles, which is an artificial parameter (may be cooked up or ad hoc) used by most halo finder methods. Majority of the  $\lambda_3$ -haloes have low number of particles; median of number of particles per halo is 49. Each particle in this simulation is approximately  $3.65 \times 10^{10} h^{-1} M_\odot$ . Hence the halo mass range varies in the order of  $10^{11} M_\odot$  to  $10^{14} M_\odot$ . Combined mass of all the  $\lambda_3$ -halo candidates is about 31 per cent of the total mass in the simulation. In contrast, the haloes occupy just 0.3 per cent of the total volume. Thus the  $\lambda_3$ -haloes are extremely dense structures. Further analysis of halo mass function of  $\lambda_3$ -haloes and comparison with AHF- and FOF-haloes is done in the Section 5.6.

## 5.6 Correspondence with other halo finders

Comparison of haloes obtained from AHF and FOF method, along with our geometric analysis of the multistream field reveals several interesting features. The number of haloes ( $N_H$ ) with at least 20 particles that were detected by all the algorithms is shown in Table 5.4. For both the simulations, FOF detects the highest number of haloes and AHF detects the least. By applying the Hessian algorithm on multistream fields smoothed at the scale of diagnostic grid size,  $l_d$ , we detected around 4500 and 28000 haloes in simulations with  $128^3$  and  $256^3$  particles respectively. The number of  $\lambda_3$ -halo is close to the mean of AHF- and FOF- haloes in each simulation – i.e.,  $N_H^{\lambda_3}$  is around 2 per cent of mean of  $N_H^{\text{AHF}}$  and  $N_H^{\text{FOF}}$  for the  $N_p = 128^3$  simulation and 8 per cent for the  $N_p = 256^3$  simulation. Multistream field both the simulations we calculated at a refinement factor of  $l_l/l_d = 2$ .

The halo mass functions from all three finders are shown in Figure 5.10. For smaller haloes of order of  $10^{13} M_\odot$ , our method predicts a slightly higher number of haloes than FOF and AHF. For the most massive haloes of mass around  $10^{14} M_\odot$ , number of  $\lambda_3$ -haloes is fewer than the other 2

Table 5.4: Number of haloes,  $N_H$  detected by the three halo finder algorithms in the two simulations of  $L = 100h^{-1}\text{Mpc}$  with different mass resolutions. Values shown for  $\lambda_3$ -haloes are done in the multistream fields with refinement factor of 2, and smoothing scale equal to the diagnostic grid size.

$N_p$	$N_H^{\text{AHF}}$	$N_H^{\lambda_3}$	$N_H^{\text{FOF}}$
$128^3$	3374	4492	5440
$256^3$	24710	27929	35765

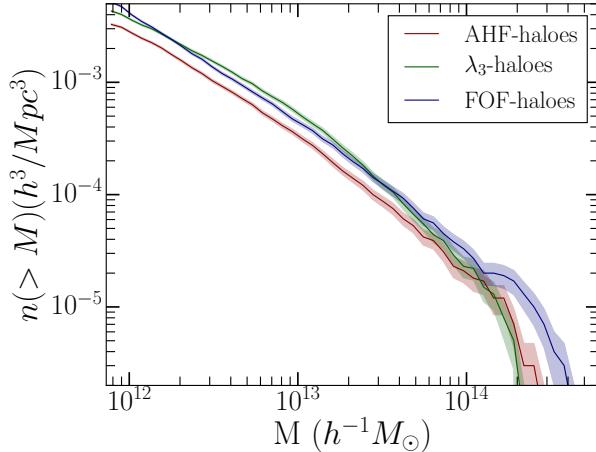


Figure 5.10: Halo Mass functions from AHF-, FOF- and  $\lambda_3$ -haloes. The AHF-haloes are fewer than FOF- or  $\lambda_3$ -haloes. The number of haloes above a mass threshold are binned and taken along vertical axis, normalized to simulation box volume. Error of  $1\sigma$  is shown in shaded region.

methods, albeit around the range of error of AHF-haloes.

By observing some of the massive haloes, like the one in Figure 5.11, we find that the  $\lambda_3$ -halo particles are within AHF- or FOF-halo region. This is generally observed in other massive haloes too: the large  $\lambda_3$ -haloes have fewer particles than corresponding AHF- or FOF-haloes. For haloes greater than  $10^{14}M_\odot$ ,  $\lambda_3$ -haloes have boundaries slightly within the AHF virial radius. Without unbinding, the FOF-haloes can be very large compared to other methods, as seen in Figure 5.11. This results in a deviation in the  $\lambda_3$ -halo mass function (Figure 5.10) from the other two methods over halo mass of  $10^{14}M_\odot$ . Further discussion of size of the detected  $\lambda_3$ -haloes in the context of smoothing of the multistream is done in Section 5.4.3.

The particles identified by the AHF as belonging to haloes form spherical structures due to a series of processes (including virialization) applied to unbind the particles. Inherently, the iso-

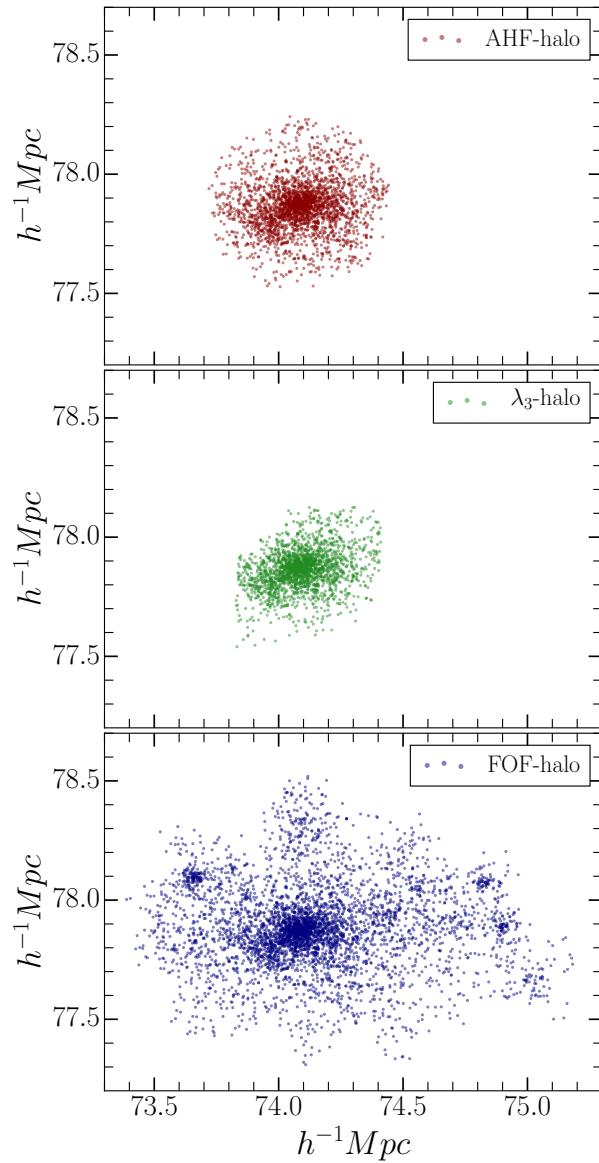


Figure 5.11: A large halo that was detected by AHF (top, red), our geometric analysis in the multistream field (centre, green) and FOF (bottom, blue). Halo boundary differs for each halo finder method. AHF detects particles within a sphere of virial radius. FOF-halo is irregularly shaped.  $\lambda_3$ -halo particles are in a non-spherical, yet compact structure.

density contours at virial levels are not spherical or spheroidal. The virialized AHF-haloes on the web are shown in the top panel of Figure 5.12. However, the spherical AHF-halo are fewer in number compared to the other methods.

The popular choice of linking length of  $b = 0.2$ , although corresponding to virial density, does not ensure that the haloes have positive curvature. Most algorithms based on the FOF method redefine the halo boundaries by unbinding the particles outside a truncation radius. This truncation radius maybe the distance from the centre of mass of the halo to the farthest particle, rms distance, or an inflection point in the density field (For details on these methods, see Knebe et al. 2011 and references therein). Some halo finders define the virial radius,  $r_{vir}$  at the distance from halo center where the density is  $\Delta_{vir}$  times the background density. In the middle panel of Figure 5.12, the FOF-haloes are shown without any of the above post-processing schemes. Without any unbinding, the FOF-haloes are generally larger in size than  $\lambda_3$ -haloes in the centre panel of Figure 5.12. For a specific case of a massive halo, Figure 5.11, FOF identifies more particles as bounded, than AHF or our algorithm.

In contrast to the AHF and FOF algorithms, our halo method directly detects a closed, convex surface (approximately the largest one, since  $\lambda_3 > 0$ ) for each of the haloes. There is no unbinding procedure on the particles identified within the halo site. The boundaries of  $\lambda_3$ -haloes themselves are not spherical or of any regular structure, but they are closed convex surfaces, as seen in the middle panel of Figure 5.12.

Haloes from the three finders in Figure 5.12 also show other differences in the halo boundaries. While all the AHF-haloes are spherical by definition, the FOF-haloes are irregular. The boundaries of the  $\lambda_3$  haloes are not spherical either, but are more compact than FOF-haloes and in addition they are convex surfaces by design. At some junctions of the filaments, FOF identifies a large region as belonging to single halo, whereas AHF and our algorithm detect multiple isolated haloes. Each isolated  $\lambda_3$ -halo region enclose one maximum of multistream field, thus guaranteeing that multiple haloes are always resolved. On the other hand, a linking length cut-off or a constant threshold on scalar fields may enclose regions with multiple local maxima ( For one-dimensional fields, an

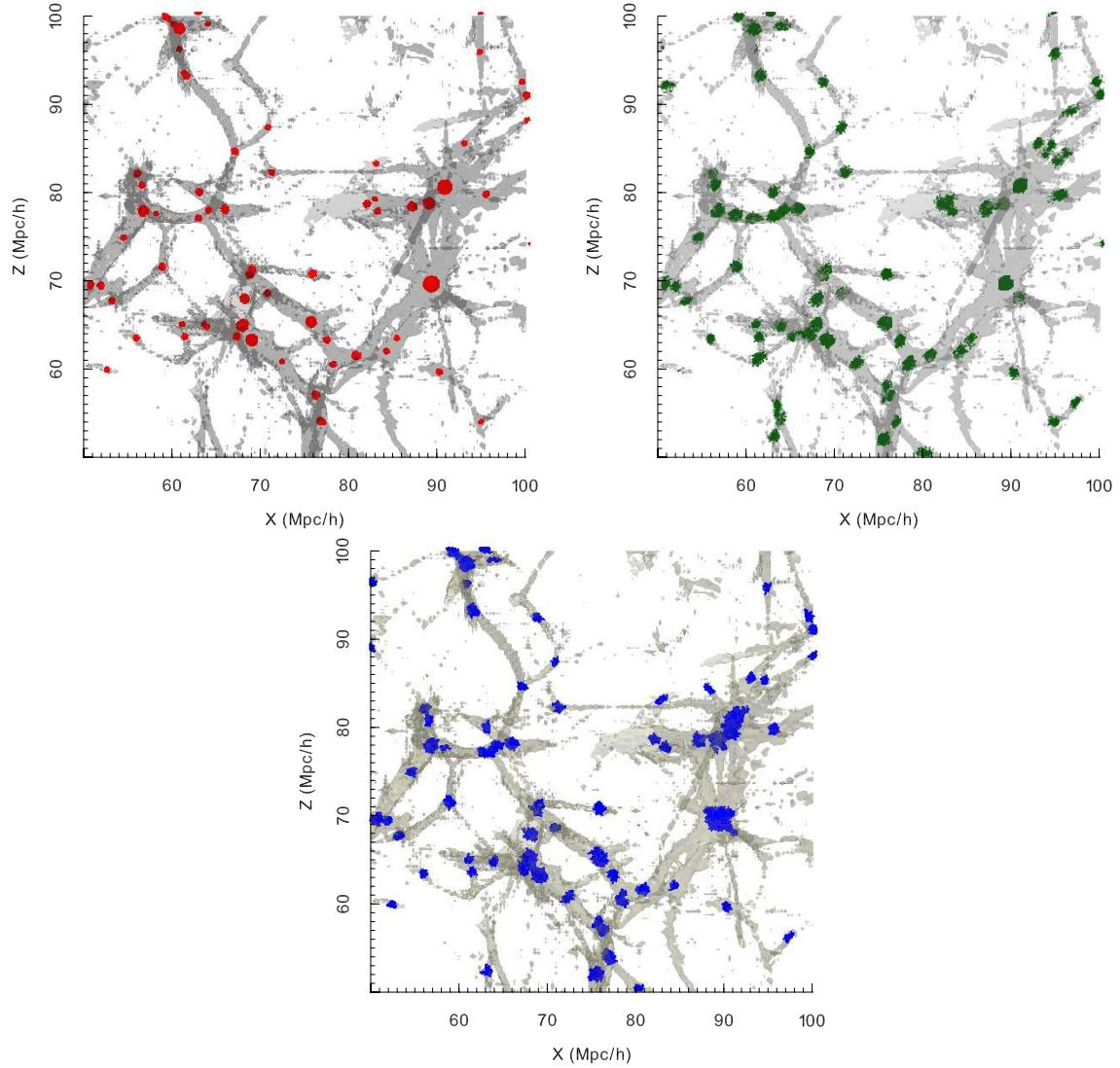


Figure 5.12: Potential haloes detected by AHF (top), our analysis (centre) and FOF (bottom). Most of the haloes are embedded in a percolating filament with  $n_{str} \geq 9$ . AHF-haloes are spherical by definition. FOF-haloes without any post processing are elongated along the filament.  $\lambda_3$ -halo candidates are neither spherical, nor elongated. Boundaries of  $\lambda_3$ -haloes are well resolved.

illustration of this is shown in Appendix C).

For a simulation box with  $N_p$  particles, each of mass  $m_p$ , the halo-mass fractions,  $f_h = (\sum_{i=1}^{N_H} m_H(i))/(m_p N_p)$  (where  $m_H$  is the mass of each halo and  $N_H$  is total number of haloes) are shown in a Venn diagram in Figure 5.13. For the simulation with  $N_p = 128^3$  particles, AHF-,  $\lambda_3$ -, and FOF- haloes occupy 22, 31 and 35 per cent of the total mass respectively. Nearly 19 per cent of the total mass are concurrently detected as belonging to haloes by all the three algorithms. FOF (with highest halo mass fraction) detects virtually all the haloes that AHF (with least halo mass fraction). About 3 per cent of the particles classified as belonging to haloes by both AHF and FOF are not classified as multistream halo particles. Our method also detected nearly 6 per cent of mass particles as haloes, which neither FOF nor AHF classify as haloes. For simulation with  $N_p = 256^3$  particles, the corresponding halo mass fractions  $f_h^{\text{AHF}} = 30$  per cent,  $f_h^{\lambda_3} = 32$  per cent, and  $f_h^{\text{FOF}} = 42$  per cent respectively. Thus the mass fraction  $f_h^{\lambda_3}$  remains fairly consistent over increasing mass resolution, as opposed to AHF and FOF. However, large fractions of these mass particles, nearly 23 per cent of the  $N_p = 256^3$  (increased from 19 per cent for low mass resolution simulation), are simultaneously detected as belonging to haloes by different methods, as shown in the right panel of Figure 5.13. For the simulation with  $N_p = 256^3$ , we also see increase in agreement between any two pairs for halo finders. That is, the mass fraction of haloes simultaneously detected by  $\lambda_3$  and FOF jumps from 25 per cent to 27 per cent. This correspondence increases from 19 to 23 per cent for  $\lambda_3$ -AHF, and 22 to 30 per cent for FOF-AHF pairs. For the same pair (in the  $N_p = 256^3$  simulation), 12.3 per cent of particles are detected by FOF but not AHF, whereas almost all the particles ( $> 99.9$ ) for the AHF particles were also detected by FOF. For the pair  $\lambda_3$ -AHF, 9.5 per cent of particles are detected by  $\lambda_3$  but not AHF, and 7.3 per cent of particles were detected by AHF but not by  $\lambda_3$ . Finally, for the  $\lambda_3$ -FOF pair, 5 per cent of particles are detected by  $\lambda_3$  but not FOF, and 15 per cent of particles were detected by FOF but not by  $\lambda_3$ .

On the other hand, looking at the mass particles that were only detected as haloes by one method, but not by other two, we see that only the multistream haloes improve (i.e., the disagreement reduces from 5.9 to 5 per cent) with mass resolution. FOF detects 6.3 and 7.8 per cent of

haloes in simulations of  $128^3$  and  $256^3$  particles respectively, that were not classified as haloes by the other two methods. AHF-halo particles, being sub-set of FOF-haloes for the most part, show less than 0.1 per cent disagreement with other finders.

The discrepancies may have to be addressed on a case-by-case basis. One of the major difference between the haloes detected by isolating local multistream maxima regions and AHF/FOF is shown for a large halo in Figure 5.11. Without any unbinding procedure, FOF may detect very large irregular sized haloes, often consisting of multiple sub-haloes as shown in the bottom panel. On the other hand, the corresponding AHF-halo (top panel) is smaller spherical subset of FOF-halo. Furthermore, the  $\lambda_3$ -halo in the middle panel is smaller than both. Our multistream field detection technique selects convex regions with strictly one  $n_{str}$  maxima within them. The sub-haloes detected by FOF (or AHF), may be detected as separate  $\lambda_3$ -haloes. Nevertheless, some of the mass particles between the two neighboring haloes (like ones along saddle regions of multistream fields) will not be included as belonging to the halo. This effect is seen in halo mass functions (Figure 5.10 for large haloes of mass more than  $10^{14} M_\odot$  – number density of large  $\lambda_3$ -haloes is smaller than FOF. Similarly it causes a few discrepancies in mass fractions of potential haloes as well.

Other cause for differences in mass fraction is also rooted in the definition of haloes. Single-streaming regions are excluded from our halo search completely. Whereas, FOF and AHF employ no such mechanism to check for number of gravitational collapses. Ramachandra & Shandarin (2015) showed that a significant fraction (nearly 35 per cent) of FOF-haloes are in contact with single-streaming voids. Particles within these regions would not be considered as potential  $\lambda_3$ -halo particles. This also contributes to the discrepancy in halo mass fraction by different halo finders.

### 5.6.1 Haloes in the percolating filament

The excursion set of multistreams above an  $n_{str}$  threshold hosts a varying number of haloes. We compare the multistream halo candidates from our geometric method with the AHF and the FOF method in Figure 5.14

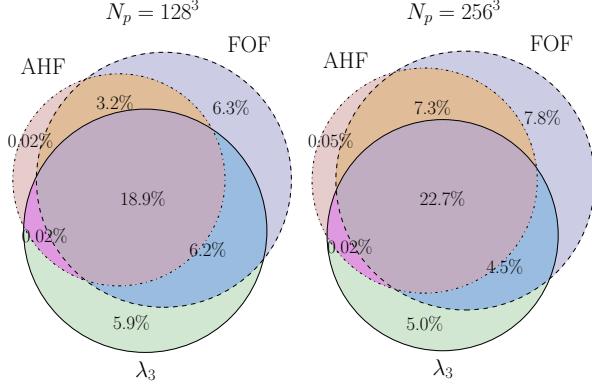


Figure 5.13: Mass fraction of haloes  $f_h$  (in per cent) as detected by the three finders. Each circle represents fractions of mass of all halo particles (AHF, FOF or  $\lambda_3$ ) in the total mass of the simulation box. The particles concurrently detected as belonging to haloes by different frameworks are shown in the intersections.

for the simulation with  $N_p = 256^3$ . The regions in the co-ordinate space are classified into excursion set and non-excursion set regions based on whether the multistream is over or under the  $n_{str}$  threshold. In the excursion set we also distinguish the largest structure from the rest of the structures because the largest region of the excursion set plays the crucial role in detecting the transition to percolation. Percolation takes place at thresholds  $n_{str} \leq 17$  (Ramachandra & Shandarin 2017) to the right from the vertical dashed dashed line. Based on the coordinates of the halo particles, we check if a halo is in contact with the largest region of the excursion set or with rest of the excursion set.

The fraction of haloes in the non-excursion set are shown at various  $n_{str}$  thresholds in the top panel of Figure 5.14. At thresholds greater than 17 streams (i.e. in non-percolating regime), a large fractions of the AHF-, FOF- and  $\lambda_3$ -haloes are in the non-excursion set, as shown in the top panel of Figure 5.14. The fraction of  $\lambda_3$ -haloes is slightly higher than FOF or AHF in this regime. At relatively high threshold of, say,  $n_{str} = 35$ , about 65% of the AHF-haloes, about 75% of the FOF-haloes and about 80% of the  $\lambda_3$ -haloes are in the non-excursion set.

For multistream values lower than the percolation threshold of  $n_{str} \leq 17$  (i.e. in percolating regime) the fractions of AHF-,  $\lambda_3$ - and FOF-haloes in the largest (i.e. percolating) region quickly grow with decreasing values of the threshold and surpass both the corresponding fractions in the

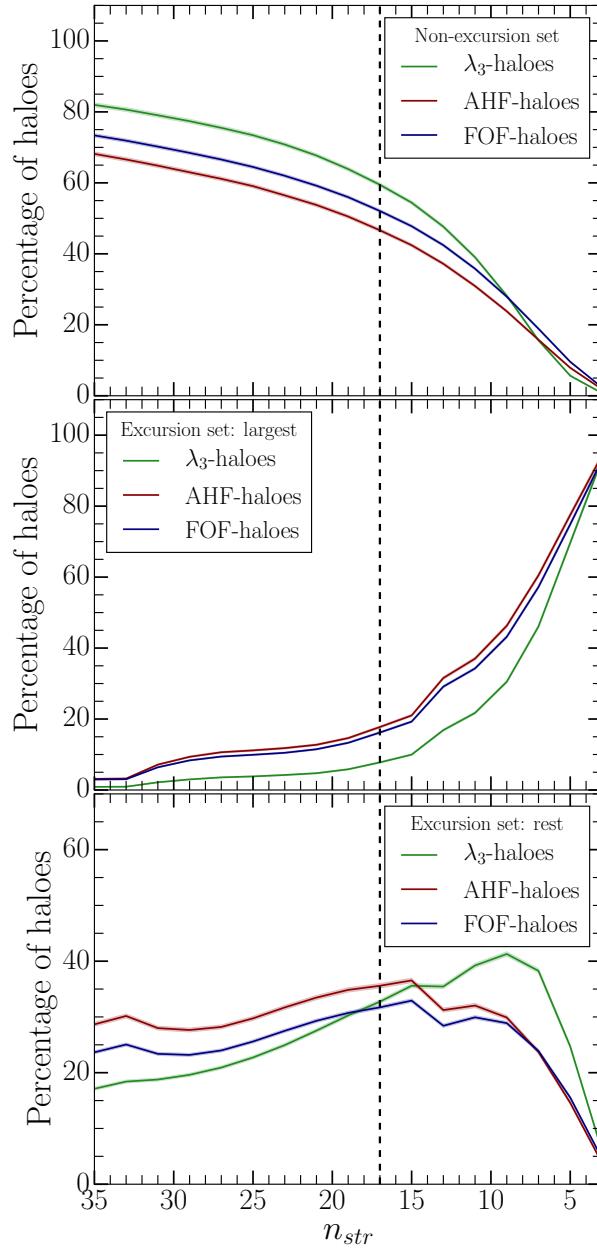


Figure 5.14: Percentage of haloes detected (in the simulation with  $N_p = 256^3$ ) that are embedded in the non-excision set (top panel), largest excursion set segment (center panel) and the rest of excursion set regions (bottom panel). Multistream haloes, AHF-haloes and FOF-haloes are shown in green, red and blue respectively. Vertical dashed line at  $n_{str} = 17$  is where percolation transition occurs.

non-percolating regions of the excursion set and that in the non-excursion set at  $n_{str} \approx 10$ . The majority of the haloes belong to the single percolating structure (shown for the simulation with  $N_p = 128^3$  in Figure 5.12. Similar spatial distribution of SUBFIND haloes Springel et al. 2001a in the multistream regions is shown by Aragon-Calvo et al. 2016) and at  $n_{str} = 3$ , all the haloes are attached to the web.

At  $n_{str} = 3$ , the filling fraction  $f_1/f_{ES}$  is almost close to unity Ramachandra & Shandarin (2017). Most halo candidates from all three algorithms are at least in contact with the percolating cosmic web. Due to the threshold on  $n_{str}$  in our halo detection method, the  $\lambda_3$ -haloes are not only in contact, but completely within this structure.

## 5.7 Discussion

The multistream field can have only integral values, moreover these values must be odd numbers practically at every point, because the number of streams can be even only on a set of measure zero. It means that in numerical simulations even values may occur on extremely rare occasions. We have analysed functional variation of the scalar field  $-n_{str}(\mathbf{x})$  using Hessian eigenvalues. The Hessian analysis is generally done for inherently continuous fields since it requires evaluation of the second derivatives. Geometries of web structures unveiled by Hessian signatures of smoothed density fields (such as Sousbie et al. 2008, Aragon-Calvo et al. 2007, Aragon-Calvo et al. 2010, Cautun et al. 2014, Bond et al. 2010a and many more), tidal shear or velocity shear tensor (Hahn et al. 2007, Forero-Romero et al. 2009, Hoffman et al. 2012, Hoffman et al. 2012, Libeskind et al. 2013, Cautun et al. 2014 etc.) and observational data from surveys (Sousbie et al. 2008, Bond et al. 2010a, Bond et al. 2010b, Pahwa et al. 2016 etc.).

Although the multistream field has discrete values by definition, it may be smoothed for numerical analysis at some scale, typically the scale of grid length of the field. The resulting Hessian eigenvalues approximately characterize the geometry in a four-dimensional hyper-space of  $(-n_{str}, x, y, z)$ . Our only assumption about the shape of the boundary of a virialized halo is that it is

a convex surface. Therefore the boundary of a halo can be defined as a region with  $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$  since it is a closed convex contour in the  $(-n_{str}, x, y, z)$  hyper-space, and thus its projection onto the three-dimensional  $(x, y, z)$  space is also closed and convex.

Differentiating a smoothed  $n_{str}(\mathbf{x})$ -field may still pose a problem in the regions where  $n_{str}(\mathbf{x})$  is close to a constant and therefore the eigenvalues represent noise about the zero level. Fortunately in the most of such regions the unsmoothed  $n_{str}(\mathbf{x}) = 1$  therefore they can be easily eliminated.

Our algorithm for detecting potential dark matter haloes is unique due to two important factors: the geometrical attribute and the choice of field. Local geometrical analysis on the multistream field conveniently delineates the non-void structures without any free parameters. The dark halo candidates have compact surfaces that enclose local maxima of the multistream field. We do not employ non-local thresholds that several halo finders use (see Knebe et al. 2011, Knebe et al. 2013 and Onions et al. 2012 for comparisons of various halo finders). Global thresholds (like a constant  $n_{str}$  cut-off) might be unsuitable for detecting halo candidates since the halo multistream environments are generally not categorical. Secondly, the  $n_{str}$  field enables us to mask out the regions belonging to mono-stream regions without a heuristic criteria. Our method guarantees that none of the  $\lambda_3$ -halo particles belong to void region.

We note that present halo finders employ a variety of physical and numerical processes to identify dark matter halo candidates. Furthermore, there is no consensus in the definition of haloes itself (see discussion in Knebe et al. 2011). This is also the cause for the few differences between FOF-, AHF- and multistream haloes: FOF and AHF haloes only use Eulerian co-ordinates  $\mathbf{x}(z)$  – either raw positions or in the form of mass density. On the other hand, we utilize a mapping on the Lagrangian sub-manifold  $\mathbf{x}(\mathbf{q}, z)$  to define the multistream field  $n_{str}(\mathbf{x})$ . The boundaries of haloes in FOF and AHF are defined by the free-parameter thresholds of linking length and density, and the halo-center is usually defined as the center of mass of the particles. Conceptually, the center of a  $\lambda_3$ -halo is the location of the local multistream maximum, and the boundary of the halo is the convex region surrounding it. At least for large haloes like the one in Figure 5.11, this convex region is well within the FOF boundary.

The applicability of non-local thresholds in detecting haloes deserves deeper investigation. Lower bounds on over-density or linking-length thresholds traditionally define halo regions in several halo finders. Values such as  $\Delta_{vir} \approx 180$  or  $b \approx 0.2$  correspond to virial theorem applied to isolated spherical collapse models. Recently More et al. (2011) demonstrated that depending on halo environment, cosmology and redshift the over-densities corresponding to  $b = 0.2$  have different values. The virial theorem itself is a good measure of equilibrium of a system. However, the global thresholds empirically derived from it may not be pertinent to diverse environment of dark matter haloes.

The algorithm prescribed in Section 5.4 lists out a set of physically motivated steps that filter out the noisy  $\lambda_3 > 0$  regions that cannot be identified as haloes. The analysis in the simulation of  $100h^{-1}$  Mpc side length and  $128^3$  particles, with the multistream calculated on  $256^3$  diagnosis box, approximately 40000 labelled segments satisfying  $\lambda_3 > 0$  criterion in the non-voids were filtered out by a lower limit on multistreaming regions. One of the possible improvements in our algorithm would be to use information of number of flip-flops of each particle (For instance, using methods prescribed by Shandarin & Medvedev 2014, 2017). Such methods involving the Lagrangian sub-manifold may reveal rich sub-structure in the haloes.

The requirement that each halo should have closed convex surfaces with a multistream maximum inside may identify sub-haloes in large haloes but is too demanding because a halo with sub-haloes must have saddle points in the  $n_{str}$  field. This may explain the shortage of massive haloes shown in Figure 5.10. Applying a filter for smoothing the  $n_{str}$  field increases the number of massive haloes but reduces the number of low mass haloes. Although our present method does not currently perform an analysis simultaneously on multiple smoothing scales, such approaches done in density, lognormal density, tidal, velocity divergence or velocity shear fields (see MMF by Aragon-Calvo et al. 2007 and NEXUS+ by Cautun et al. 2013) have shown interesting multi-scale features of the cosmic web. Applying a more sophisticated procedure for linking  $\lambda_3$ -sub-haloes in a more massive halo will be done in the follow-up paper.

Dark matter haloes, being localised structures, are uniquely convenient for our local Hessian

analysis. Conditions of  $\lambda_1 > 0 > \lambda_2 \geq \lambda_3$  and  $\lambda_1 \geq \lambda_2 > 0 > \lambda_3$  also give information about curvature. Hessian eigenvalue analysis at high resolution of multistream fields may be very interesting in understanding the tubular edges of filaments and surfaces of walls at smaller scales. However, in this study, Hessian analysis is only applied to haloes. Walls and filaments span large volumes in the dark matter simulations, and we employ topological tools to investigate them.

## 5.8 Summary

We studied certain geometrical of the multistream field in the context of halo formation. Findings from our analysis are summarized as follows:

1. Several aspects of halo formation in the Lagrangian sub-manifold are considerably different than that of reference models of spherical top-hat collapse and ellipsoidal collapse. Successive formations of caustics (and consequently multiple velocity streams) play a crucial role in the process of clustering.
2. We present a novel halo detection algorithm for identifying dark matter halo candidates in the multistream field. Conditions on the local geometric indicators of the field are used to ensure that each closed halo boundary hosts a local multistream maxima. The positive signs of all principal curvatures (please note that we use curvatures of  $-n_{str}(\mathbf{x})$  field) inside the boundary also guarantee that the boundary is convex. Bounds on  $n_{str}$  guarantee that all the halo particles are in the non-void structure. We also ensure that the halo regions have foldings in the Lagrangian sub-manifold in more than one direction.
3. The multistream field within the halo boundaries may be very diverse. We do not detect halo candidates from a global lower bound on  $n_{str}$ . Instead, we look for closed convex regions in the multistream field. For the simulation with  $128^3$  particles, minima of  $n_{str}$  in each halo vary from 3 to nearly 450. Maxima of  $n_{str}$  in the halo vary from 7 to about 2800.
4. Our multistream halo candidates had a reasonably good correspondence with haloes from

AHF and FOF catalogues. One notable difference was found with massive haloes. Our algorithm predicted fewer particles than the FOF method. This is likely to be caused by the requirement that the multistream field in the regions of the  $\lambda_3$ -halo candidates is convex which may be a reasonable approximation for simple haloes (i.e. having no sub-haloes) but massive haloes are more likely to have sub-haloes and therefore the  $n_{str}$  field in the corresponding regions must have saddle points and therefore cannot be entirely convex. Our study of the smoothing effects has shown that the number of massive sub-haloes tend to increase with growing smoothing scale which seems to agree with the above explanation. We will address this problem in the following study.

5. Halo candidates were mostly embedded on the excursion set of the multistream field after percolation transition ( $n_{str} = 17$  in the simulation with  $256^3$  particles). At lower thresholds (around  $n_{str} = 5$  to 11), the largest percolating structure in the excursion set hosts most of the haloes.

In conclusion, the Lagrangian sub-manifold contains dynamical information of structure formation. We analysed the multistream field that contains the information of foldings in the sub-manifold. In addition, we demonstrated the use of geometrical features of the multistream field in identifying potential dark matter halo candidates in cosmological N-body simulations.

## Acknowledgements

This work has been funded in part by the University of Kansas FY 2017 Competition General Research Fund, GRF Award 2301155. This research used resources of the Advanced Computing Facility, which is part of the Center for Research Computing at the University of Kansas. We thank Mark Neyrinck and Mikhail Medvedev for discussions and suggestions. We also thank the anonymous referee for insightful comments on improving this manuscript and the editor for help with the submission.

# References

- Abel, T., Hahn, O., & Kaehler, R. (2012). Tracing the dark matter sheet in phase space. *MNRAS*, 427(1), 61–76.
- Achitouv, I., Neyrinck, M., & Paranjape, A. (2015). Testing spherical evolution for modelling void abundances. *MNRAS*, 451(4), 3964–3974.
- Albareti, F. D., Allende Prieto, C., Almeida, A., Anders, F., Anderson, S., Andrews, B. H., Aragon-Salamanca, A., Argudo-Fernandez, M., Armengaud, E., Aubourg, E., Avila-Reese, V., Badenes, C., Bailey, S., Barbuy, B., Barger, K., Barrera-Ballesteros, J., Bartosz, C., Basu, S., Bates, D., Battaglia, G., Baumgarten, F., Baur, J., Bautista, J., Beers, T. C., Belfiore, F., Bershadsky, M., de Lis, S. B., Bird, J. C., Bizyaev, D., Blanc, G. A., Blanton, M., Blomqvist, M., Bolton, A. S., Borissova, J., Bovy, J., Brandt, W. N., Brinkmann, J., Brownstein, J. R., Bundy, K., Burtin, E., Busca, N. G., Chavez, H. O. C., Diaz, M. C., Cappellari, M., Carrera, R., Chen, Y., Cherinka, B., Cheung, E., Chiappini, C., Chojnowski, D., Chuang, C.-H., Chung, H., Cirolini, R. F., Clerc, N., Cohen, R. E., Comerford, J. M., Comparat, J., Cousinou, M.-C., Covey, K., Crane, J. D., Croft, R., Cunha, K., da Costa, L., Ilha, G. d. S., Darling, J., Davidson, J. W., Dawson, K., De Lee, N., de la Macorra, A., de la Torre, S., Machado, A. D., Delubac, T., Diamond-Stanic, A. M., Donor, J., Downes, J. J., Drory, N., des Bourboux, H. d. M., Du, C., Dwelly, T., Ebelke, G., Eigenbrot, A., Eisenstein, D. J., Elsworth, Y. P., Emsellem, E., Eracleous, M., Escoffier, S., Evans, M. L., Falcon-Barroso, J., Fan, X., Favole, G., Fernandez-Alvar, E., Fernandez-Trincado, J. G., Feuillet, D., Fleming, S. W., Font-Ribera, A., Freischlad, G., Frinchaboy, P., Fu, H., Gao, Y., Garcia-Hernandez, D. A., Perez, A. E. G., Garcia, R. A., Garcia-Dias, R., Gaulme, P., Ge,

J., Geisler, D., Marin, H. G., Gillespie, B., Girardi, L., Goddard, D., Chew, Y. G. M., Gonzalez-Perez, V., Grabowski, K., Green, P., Grier, C. J., Grier, T., Guo, H., Guy, J., Hagen, A., Hall, M., Harding, P., Harley, R. E., Hasselquist, S., Hawley, S., Hayes, C. R., Hearty, F., Hekker, S., Toledo, H. H., Ho, S., Hogg, D. W., Holley-Bockelmann, K., Holtzman, J. A., Holzer, P. H., Hu, J., Huber, D., Hutchinson, T. A., Hwang, H. S., Ibarra-Medel, H. J., Ivans, I. I., Ivory, K., Jaehnig, K., Jensen, T. W., Johnson, J. A., Jones, A., Jullo, E., Kallinger, T., Kinemuchi, K., Kirkby, D., Klaene, M., Kneib, J.-P., Kollmeier, J. A., Lacerna, I., Lane, R. R., Lang, D., Laurent, P., Law, D. R., Goff, J.-M. L., Leauthaud, A., Li, C., Li, R., Li, C., Li, N., Liang, F.-H., Liang, Y., Lima, M., Lin, L., Lin, L., Lin, Y.-T., Long, D., Lucatello, S., MacDonald, N., MacLeod, C. L., Mackereth, J. T., Mahadevan, S., Maia, M. A.-G., Maiolino, R., Majewski, S. R., Malanushenko, O., Mallmann, N. D., Manchado, A., Maraston, C., Marques-Chaves, R., Valpuesta, I. M., Masters, K. L., Mathur, S., McGreer, I. D., Merloni, A., Merrifield, M. R., Meszaros, S., Meza, A., Miglio, A., Minchev, I., Molaverdikhani, K., Montero-Dorta, A. D., Mosser, B., Muna, D., Myers, A., Nair, P., Nandra, K., Ness, M., Newman, J. A., Nichol, R. C., Nidever, D. L., Nitschelm, C., O'Connell, J., Oravetz, A., Padilla, N., Palanque-Delabrouille, N., Pan, K., Parejko, J., Paris, I., Peacock, J. A., Peirani, S., Pellejero-Ibanez, M., Penny, S., Percival, W. J., Percival, J. W., Perez-Fournon, I., Petitjean, P., Pieri, M., Pinsonneault, M. H., Pisani, A., Prada, F., Prakash, A., Price-Jones, N., Raddick, M. J., Rahman, M., Raichoor, A., Rembold, S. B., Reyna, A. M., Rich, J., Richstein, H., Ridl, J., Riffel, R., Riffel, R. A., Rix, H.-W., Robin, A. C., Rockosi, C. M., Rodriguez-Torres, S., Rodrigues, T. S., Roe, N., Lopes, A. R., Roman-Zuniga, C., Ross, A. J., Rossi, G., Ruan, J., Ruggeri, R., Runnoe, J. C., Salazar-Albornoz, S., Salvato, M., Sanchez, A. G., Sanchez, S. F., Sanchez-Gallego, J. R., Santiago, B. X., Schiavon, R., Schimoia, J. S., Schlafly, E., Schlegel, D. J., Schneider, D. P., Schoenrich, R., Schultheis, M., Schwope, A., Seo, H.-J., Serenelli, A., Sesar, B., Shao, Z., Shetrone, M., Shull, M., Aguirre, V. S., Skrutskie, M. F., Slosar, A., Smith, M., Smith, V. V., Sobeck, J., Somers, G., Souto, D., Stark, D. V., Stassun, K. G., Steinmetz, M., Stello, D., Bergmann, T. S., Strauss, M. A., Streblyanska, A., Stringfellow, G. S., Suarez, G., Sun, J., Taghizadeh-Popp, M.,

Tang, B., Tao, C., Tayar, J., Tembe, M., Thomas, D., Tinker, J., Tojeiro, R., Tremonti, C., Troup, N., Trump, J. R., Unda-Sanzana, E., Valenzuela, O., van den Bosch, R., Vargas-Magana, M., Vazquez, J. A., Villanova, S., Vivek, M., Vogt, N., Wake, D., Walterbos, R., Wang, Y., Wang, E., Weaver, B. A., Weijmans, A.-M., Weinberg, D. H., Westfall, K. B., Whelan, D. G., Wilcots, E., Wild, V., Williams, R. A., Wilson, J., Wood-Vasey, W. M., Wylezalek, D., Xiao, T., Yan, R., Yang, M., Ybarra, J. E., Yeche, C., Yuan, F.-T., Zakamska, N., Zamora, O., Zasowski, G., Zhang, K., Zhao, C., Zhao, G.-B., Zheng, Z., Zheng, Z., Zhou, Z.-M., Zhu, G., Zinn, J. C., & Zou, H. (2016). The Thirteenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-IV Survey Mapping Nearby Galaxies at Apache Point Observatory.

Angulo, R. E., Chen, R., Hilbert, S., & Abel, T. (2013a). Noiseless Gravitational Lensing Simulations.

Angulo, R. E., Hahn, O., & Abel, T. (2013b). How closely do baryons follow dark matter on large scales? *MNRAS*, 434(2), 1756–1764.

Angulo, R. E., Hahn, O., & Abel, T. (2013c). The warm dark matter halo mass function below the cut-off scale. *MNRAS*, 434(4), 3337–3347.

Angulo, R. E., Hahn, O., Ludlow, A., & Bonoli, S. (2016). Earth-mass haloes and the emergence of NFW density profiles. *MNRAS*, 15(April), 1–15.

Aragon-Calvo, M. A., Jones, B. J. T., Van De Weygaert, R., M, J., & Der Hulst, V. (2007). The Multiscale Morphology Filter: Identifying and Extracting Spatial Patterns in the Galaxy Distribution. *Astron. Astrophys.*, 474(1), 28.

Aragon-Calvo, M. A., Platen, E., Van De Weygaert, R., & Szalay, A. S. (2008). The Spine of the Cosmic Web. *ApJ*, 723(1), 364–382.

Aragon-Calvo, M. A., Van De Weygaert, R., Jones, B. J. T., Arag??n-Calvo, M. A., Van De Weygaert, R., Jones, B. J. T., Aragon-Calvo, M. A., Van De Weygaert, R., Jones, B. J. T., Aragon-

- Calvo, M. A., Van De Weygaert, R., & Jones, B. J. T. (2010). Multiscale phenomenology of the cosmic web. *MNRAS*, 408(4), 2163–2187.
- Aragon-Calvo, M. A. M., Neyrinck, M. M. C., & Silk, J. (2016). How Cosmic Web Detachment Drives Galaxy Quenching. *ArXiv e-prints*, (July).
- Arnold, V., Shandarin, S., & Zeldovich, I. (1982). The large scale structure of the universe. I - General properties One- and two-dimensional models. *Geophys. Astrophys. Fluid Dyn.*, 20, 111–130.
- Ascasibar, Y. (2010). Estimating multidimensional probability fields using the Field Estimator for Arbitrary Spaces (FiEstAS) with applications to astrophysics. *Comput. Phys. Commun.*, 181(8), 1438–1443.
- Bagla, J. S. & Padmanabhan, T. (1997). Cosmological N-Body Simulations. *Pramana*, V, 161.
- Beacom, J. F., Dominik, K. G., Melott, A. L., Perkins, S. P., & Shandarin, S. F. (1991). Gravitational clustering in the expanding Universe. *ApJ*, 372, 351–363.
- Behroozi, P. S., Wechsler, R. H., & Wu, H.-y. (2013). the Rockstar Phase-Space Temporal Halo Finder and the Velocity. *ApJ*, 762(109), 20pp.
- Bertschinger, E. (1985). Self-Similar Secondary Infall and Accretion in an Einstein-de Sitter Universe. *ApJS*, 58, 39–66.
- Bertschinger, E. (1998). Simulations of Structure Formation in the Universe. *Annu. Rev. Astron. Astrophys.*, 36(1), 599–654.
- Bertschinger, E. & Gelb, J. (1991). Cosmological N-Body Simulations. *Comput. Phys.*, 164(5), 164.
- Blumenthal, G., Da Costa, N., Goldwirth, D., Lecar, M., & Piran, T. (1992). The largest possible voids. *ApJ*, 388, 234–241.

- Bond, J. R., Kofman, L., & Pogosyan, D. (1996). How filaments are woven into the Cosmic web. *Nature*, 380, 603–606.
- Bond, N. A., Strauss, M. A., & Cen, R. (2010a). Crawling the cosmic network: Exploring the morphology of structure in the galaxy distribution. *MNRAS*, 406(3), 1609–1628.
- Bond, N. A., Strauss, M. A., & Cen, R. (2010b). Crawling the cosmic network: Identifying and quantifying filamentary structure. *MNRAS*, 409(1), 156–168.
- Cautun, M., van de Weygaert, R., & Jones, B. J. T. (2013). Nexus: Tracing the cosmic web connection. *MNRAS*, 429(2), 1286–1308.
- Cautun, M., Van De Weygaert, R., Jones, B. J. T., & Frenk, C. S. (2014). Evolution of the cosmic web. *MNRAS*, 441(4), 2923–2973.
- Cautun, M. C. M. & van de Weygaert, R. (2011). The DTFE public software: The Delaunay Tessellation Field Estimator code. *Astrophysics Source Code Library*.
- Centrella, J. & Melott, A. (1983). Three-dimensional simulation of large-scale structure in the universe. *Nature*, 305, 196–198.
- Colberg, J. M., Pearce, F., Foster, C., Platen, E., Brunino, R., Neyrinck, M., Basilakos, S., Fairall, A., Feldman, H., Gottl??ber, S., Hahn, O., Hoyle, F., M??ller, V., Nelson, L., Plionis, M., Porciani, C., Shandarin, S., Vogeley, M. S., & Van De Weygaert, R. (2008). The Aspen-Amsterdam void finder comparison project. *MNRAS*, 387(2), 933–944.
- Colberg, J. M., Sheth, R. K., Diaferio, A., Gao, L., & Yoshida, N. (2005). Voids in  $\Lambda$ CDM universe. *MNRAS*, 360, 216–226.
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. (1985). The evolution of large-scale structure in a universe dominated by cold dark matter. *ApJ*, 292, 371.
- de Lapparent, V., Geller, M. J., & Huchra, J. P. (1986). A slice of the universe. *ApJ*, 302, L1–L5.

- Diemand, J., Kuhlen, M., & Madau, P. (2006). Early supersymmetric cold dark matter substructure . *ApJ*, 649, 1–13.
- Doroshkevich, A. G. (1973). Spatial structure of perturbations and origin of galactic rotation in fluctuation theory. *Astrophysics*, 6(4), 320–330.
- Einasto, J., Klypin, A., Saar, E., & Shandarin, S. F. (1984). Structure of cuperclusters. *MNRAS*, 206, 529–558.
- Eisenstein, D. J. & Hu, W. (1998). Baryonic Features in the Matter Transfer Function. *ApJ*, 496, 605.
- Eisenstein, D. J. & Hut, P. (1998). HOP: A New Group-finding Algorithm for N-Body Simulations. *ApJ*, 498(1), 137–142.
- Eisenstein, D. J. & Loeb, A. (1995). An analytical model for the triaxial collapse of cosmological perturbations. *ApJ*, 439, 520.
- Elahi, P. J., Han, J., Lux, H., Ascasibar, Y., Behroozi, P., Knebe, A., Muldrew, S. I., Onions, J., & Pearce, F. (2013). Streams going Notts: The tidal debris finder comparison project. *MNRAS*, 433(2), 1537–1555.
- Evrard, A. E., MacFarland, T. J., Couchman, H. M. P., Colberg, J. M., Yoshida, N., White, S. D. M., Jenkins, A., Frenk, C. S., Pearce, F. R., Peacock, J. A., & Thomas (The Virgo Consortium), P. A. (2002). Galaxy Clusters in Hubble Volume Simulations: Cosmological Constraints from Sky Survey Populations. *ApJ*, 573(1), 7–36.
- Falck, B. & Neyrinck, M. C. (2015). The persistent percolation of single-stream voids. *MNRAS*, 450(3), 3239–3253.
- Falck, B. L., Neyrinck, M. C., & Szalay, A. S. (2012). Origami: Delineating Halos Using Phase-Space Folds. *ApJ*, 754(2), 126.

Fillmore, J. A., Goldreich, P. (1984). Self-similar gravitational collapse in an expanding universe. *ApJ*, 281, 1.

Forero-Romero, J. E., Hoffman, Y., Gottl??ber, S., Klypin, A., Yepes, G., Gottlöber, S., Klypin, A., Yepes, G., Gottl??ber, S., Klypin, A., & Yepes, G. (2009). A dynamical classification of the cosmic web. *MNRAS*, 396(3), 1815–1824.

Frenk, C., White, S., & Davis, M. (1983). Nonlinear evolution of large-scale structure in the universe. *ApJ*, 271(Part 1), 417–430.

Frenk, C. S., White, S. D. M., Davis, M., & Efstathiou, G. (1988). The formation of dark halos in a universe dominated by cold dark matter. *ApJ*, 327, 507–525.

Gill, S. P. D., Knebe, A., & Gibson, B. K. (2004). The evolution of substructure -I. A new identification method. *MNRAS*, 351(2), 399–409.

Giocoli, C., Tormen, G., Sheth, R. K., & van den Bosch, F. C. (2010). The substructure hierarchy in dark matter haloes. *MNRAS*, 404(1), 502–517.

Gott, J. R. I., Melott, A. L., & Dickinson, M. (1986). Sponge-like topology. *ApJ*, 306, 341–357.

Gottlober, S., Klypin, A., & Kravtsov, A. V. (1999). Halo Evolution. *APS Conf. Ser.*, 176.

Gregory, S. A. & Thompson, L. A. (1978). The Coma A1367 supercluster and its environs. *ApJ*, 222, 784.

Gurbatov, S.N., Saichev, A.I., Shandarin, S. F., Gurbatov, S., Saichev, A., & Shandarin, S. (1989). The large-scale structure of the universe in the frame of the model equation of non-linear diffusion. *MNRAS*, 236, 385–402.

Habib, S., Pope, A., Lukić, Z., Daniel, D., Fasel, P., Desai, N., Heitmann, K., Hsu, C.-H., Ankeny, L., Mark, G., Bhattacharya, S., & Ahrens, J. (2009). Hybrid petacomputing meets cosmology: The Roadrunner Universe project. *J. Phys. Conf. Ser.*, 180, 2019.

- Hahn, O. & Abel, T. (2011). Multi-scale initial conditions for cosmological simulations. *MNRAS*, 415(3), 2101–2121.
- Hahn, O., Abel, T., & Kaehler, R. (2013). A new approach to simulating collisionless dark matter fluids. *MNRAS*, 434(2), 1171–1191.
- Hahn, O. & Angulo, R. E. (2016). An adaptively refined phase-space element method for cosmological simulations and collisionless dynamics. *MNRAS*, 455(1), 1115–1133.
- Hahn, O., Angulo, R. E., & Abel, T. (2015). The properties of cosmic velocity fields. *MNRAS*, 454(4), 3920–3937.
- Hahn, O., Carollo, C. M., Porciani, C., & Dekel, A. (2007). The evolution of dark matter halo properties in clusters, filaments, sheets and voids. *MNRAS*, 381(1), 41–51.
- Han, J., Jing, Y. P., Wang, H., & Wang, W. (2012). Resolving subhaloes' lives with the Hierarchical Bound-Tracing algorithm. *MNRAS*, 427(3), 2437–2449.
- Heitmann, K., Frontiere, N., Sewell, C., Habib, S., Pope, A., Finkel, H., Rizzi, S., Insley, J., & Bhattacharya, S. (2015). the Q Continuum Simulation: Harnessing the Power of Gpu Accelerated Supercomputers. *ApJS*, 219(2), 34.
- Hidding, J., Shandarin, S. F., & van de Weygaert, R. (2014). The zel'dovich approximation: Key to understanding cosmic web complexity. *MNRAS*, 437(4), 3442–3472.
- Hockney, R. W. & Eastwood, J. W. (1988). *Computer simulation using particles*. CRC Press.
- Hoffman, Y., Metuki, O., Yepes, G., Gottl, S., Forero, J. E., Libeskind, N. I., & Knebe, A. (2012). A kinematic classification of the cosmic web. *MNRAS*, 425, 2049–2057.
- Hoffmann, K., Planelles, S., Gaztañaga, E., Knebe, A., Pearce, F. R., Lux, H., Onions, J., Muldrew, S. I., Elahi, P., Behroozi, P., Ascasibar, Y., Han, J., Maciejewski, M., Merchan, M. E., Neyrinck, M., Ruiz, A. N., & Sgro, M. A. (2014). Subhaloes gone Notts: Subhaloes as tracers of the dark matter halo shape. *MNRAS*, 442(2), 1197–1210.

Huchra, J. P., Macri, L. M., Masters, K. L., Jarrett, T. H., Berlind, P., Calkins, M., Crook, A. C., Cutri, R., Erdođdu, P., Falco, E., George, T., Hutcheson, C. M., Lahav, O., Mader, J., Mink, J. D., Martimbeau, N., Schneider, S., Skrutskie, M., Tokarz, S., & Westover, M. (2012). the 2Mass Redshift Survey—Description and Data Release. *ApJS*, 199(2), 26.

Icke, V. (1973). Formation of Galaxies inside Clusters. *A&A*, 27, 1.

Icke, V. & van de Weygaert, R. (1991). Voronoi. *QJRAS*, 32.

Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., & Yoshida, N. (2001). The mass function of dark matter haloes. *MNRAS*, 321(2), 372–384.

Jennings, E., Li, Y., & Hu, W. (2013). The abundance of voids and the excursion set formalism. *MNRAS*, 434(3), 2167–2181.

Klypin, A., Gottlober, S., Kravtsov, A. V., & Khokhlov, A. M. (1999). Galaxies in N-Body Simulations: Overcoming the Overmerging Problem. *ApJ*, 516(2), 530–551.

Klypin, A. & Shandarin, S. (1983). Three-dimensional numerical model of the formation of large-scale structure in the Universe. *MNRAS*, 204(Sept. 1983), 891–907.

Knebe, A., Knollmann, S. R., Muldrew, S. I., Pearce, F. R., Aragon-Calvo, M. A., Ascasibar, Y., Behroozi, P. S., Ceverino, D., Colombi, S., Diemand, J., Dolag, K., Falck, B. L., Fasel, P., Gardner, J., Gottlöber, S., Hsu, C. H., Iannuzzi, F., Klypin, A., Lukic, Z., Maciejewski, M., McBride, C., Neyrinck, M. C., Planelles, S., Potter, D., Quilis, V., Rasera, Y., Read, J. I., Ricker, P. M., Roy, F., Springel, V., Stadel, J., Stinson, G., Sutter, P. M., Turchaninov, V., Tweed, D., Yepes, G., & Zemp, M. (2011). Haloes gone MAD: The Halo-Finder Comparison Project. *MNRAS*, 415(3), 2293–2318.

Knebe, A., Pearce, F. R., Lux, H., Ascasibar, Y., Behroozi, P., Casado, J., Moran, C. C., Diemand, J., Dolag, K., Dominguez-Tenreiro, R., Elahi, P., Falck, B., Gottlöber, S., Han, J., Klypin, A.,

- Lukić, Z., Maciejewski, M., McBride, C. K., Merchán, M. E., Muldrew, S. I., Neyrinck, M., Onions, J., Planelles, S., Potter, D., Quilis, V., Rasera, Y., Ricker, P. M., Roy, F., Ruiz, A. N., Sgró, M. A., Springel, V., Stadel, J., Sutter, P. M., Tweed, D., & Zemp, M. (2013). Structure finding in cosmological simulations: The state of affairs. *MNRAS*, 435(2), 1618–1658.
- Knollmann, S. R. & Knebe, A. (2009). Ahf: Amiga'S Halo Finder. *ApJS*, 182(2), 608–624.
- Kofman, L., Pogosyan, D., Shandarin, S. F. S., & Melott, A. A. L. (1992). Coherent structures in the Universe and the Adhesion Model. *ApJ*, 393, 437–449.
- Lacey, C. & Cole, S. (1994). Merger rate in hierarchical models. *MNRAS*, 271, 676–692.
- Lee, J. & Shandarin, S. F. (1998). The cosmological mass function in the Zel'dovich approximation. *ApJ*, 500(1984), 14–27.
- Lee, J. & Yepes, G. (2016). Turning Around along the Cosmic Web. *ApJ*, 832(2), 185.
- Libeskind, N. I., Hoffman, Y., Forero-Romero, J., Gottlöber, S., Knebe, A., Steinmetz, M., & Klypin, A. (2013). The velocity shear tensor: Tracer of halo alignment. *MNRAS*, 428(3), 2489–2499.
- Lin, C., Mestel, L., & Shu, F. (1965). The Gravitational Collapse of a Uniform Spheroid. *ApJ*, 142, 1431.
- LSST Science Collaboration, Abell, P., Allison, J., Anderson, S., Andrew, J., Angel, J., Armus, L., Arnett, D., Asztalos, S., Axelrod, T., & al., E. (2009). LSST Science Book, Version 2.0.
- Maciejewski, M., Colombi, S., Springel, V., Alard, C., & Bouchet, F. R. (2009). Phase-space structures - II. Hierarchical Structure Finder. *MNRAS*, 396(3), 1329–1348.
- Melott, A., Einasto, J., Saar, E., Suisalu, I., Klypin, A., & Shandarin, S. (1983). Cluster Analysis of the Nonlinear Evolution of Large-Scale Structure in an Axion/Gravitino/Photino-Dominated Universe. *Phys. Rev. Lett.*, 51(10), 935–938.

- Melott, A. L. & Shandarin, S. F. (1989). Gravitational instability with high resolution. *ApJ*, 342, 26–30.
- Melott, A. L. & Shandarin, S. F. (1990). Generation of large-scale cosmological structures by gravitational clustering. *Nature*, 346, 633–635.
- Melott, A. L. & Shandarin, S. F. (1993). Controlled Experiments in cosmological gravitational clustering. *ApJ*, 410, 469–481.
- More, S., Diemer, B., & Kravtsov, A. V. (2015). the Splashback Radius As a Physical Halo Boundary and the Growth of Halo Mass. *ApJ*, 810(1), 36.
- More, S., Kravtsov, A. V., Dalal, N., & Gottlöber, S. (2011). the Overdensity and Masses of the Friends-of-Friends Halos and Universality of Halo Mass Function. *ApJS*, 195(1), 4.
- Neyrinck, M. C. (2008). Zobov: A parameter-free void-finding algorithm. *MNRAS*, 386(4), 2101–2109.
- Neyrinck, M. C. (2016). Tetrahedral collapse: A rotational toy model of simultaneous dark-matter halo, filament and wall formation. *MNRAS*, 460(1), 816–826.
- Neyrinck, M. C., Gnedin, N. Y., & Hamilton, A. J. S. (2005). VOBOZ: an almost-parameter-free halo-finding algorithm. *MNRAS*, 53, 1222–1232.
- Onions, J., Knebe, A., Pearce, F. R., Muldrew, S. I., Lux, H., Knollmann, S. R., Ascasibar, Y., Behroozi, P., Elahi, P., Han, J., Maciejewski, M., Merch??n, M. E., Neyrinck, M., Ruiz, A. N., Sgr??, M. A., Springel, V., & Tweed, D. (2012). Subhaloes going Notts: The subhalo-finder comparison project. *MNRAS*, 423(2), 1200–1214.
- Oort, J. (1983). Superclusters. *Ann. Rev. Astron. Astrophys.*, 21, 373–428.
- Pahwa, I., Libeskind, N. I., Tempel, E., Hoffman, Y., Brent Tully, R., Courtois, H. M., Gottl??ber, S., Steinmetz, M., & Sorce, J. G. (2016). The alignment of galaxy spin with the shear field in observations. *MNRAS*, 457(1), 695–703.

- Paranjape, A., Lam, T. Y., Sheth, R. K., & van de Weygaert, R. (2012). A hierarchy of voids: More ado about nothing. *MNRAS*, 420(2), 1648–1655.
- Parihar, P., Vogeley, M. S., Gott, J. R., Choi, Y.-Y., Kim, J., Kim, S. S., Speare, R., Brownstein, J. R., & Brinkmann, J. (2014). a Topological Analysis of Large-Scale Structure, Studied Using the Cmass Sample of Sdss-Iii. *ApJ*, 796(2), 86.
- Park, C., Pranav, P., Chingangbam, P., van de Weygaert, R., Jones, B., Vegter, G., Kim, I., Hidding, J., & Hellwing, W. A. (2013). Betti numbers of Gaussian fields. *J. Korean Astron. Soc.*, 46(3), 125–131.
- Planelles, S. & Quilis, V. (2010). ASOHF: a new adaptive spherical overdensity halo finder. *Astron. Astrophys.*, A94, 16.
- Platen, E., Van De Weygaert, R., & Jones, B. J. T. (2007). A cosmic watershed: The WVF void detection technique. *MNRAS*, 380(2), 551–570.
- Press, W. H. & Schechter, P. (1974). Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187, 425.
- Ramachandra, N. S. & Shandarin, S. F. (2015). Multi-stream portrait of the cosmic web. *MNRAS*, 452(2), 1643–1653.
- Ramachandra, N. S. & Shandarin, S. F. (2016). Topology and geometry of the dark matter web.
- Ramachandra, N. S. & Shandarin, S. F. (2017). Topology and geometry of the dark matter web : a multistream view. *MNRAS*, 1762, 1748–1762.
- Rasera, Y., Alimi, J. M., Courtin, J., Roy, F., Corasaniti, P. S., Fuzfa, A., & Boucher, V. (2010). Introducing the Dark Energy Universe Simulation Series (DEUSS). *AIP Conf. Proc.*, 1241, 1134–1139.
- Sahni, V., Sathyaprakash, B. S., & Shandarin, S. F. (1998). Shapefinders: A New Shape Diagnostic for Large-Scale Structure. *ApJ*, 495(1), L5–L8.

- Sathyaprakash, B. S., Sahni, V., & Shandarin, S. F. (1998). Cosmological Gravitational Clustering. *ApJ*, 508, 551–569.
- Schaap, W. E. (2007). *DTFE: the Delaunay Tessellation Field Estimator*. PhD thesis.
- Schaap, W. E. & van de Weygaert, R. (2000). Continuous Fields and Discrete Samples: Reconstruction through Delaunay Tessellations. *Astron. Astrophys.*, 363, L39–L32.
- Schaye, J., Crain, R. A., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J. C., Jenkins, A., Rosas-Guevara, Y. M., White, S. D., Baes, M., Booth, C. M., Camps, P., Navarro, J. F., Qu, Y., Rahmati, A., Sawala, T., Thomas, P. A., & Trayford, J. (2015). The EAGLE project: Simulating the evolution and assembly of galaxies and their environments. *MNRAS*, 446(1), 521–554.
- Shandarin, S. (2011). The multi-stream flows and the dynamics of the cosmic web. *J. Cosmol. Astropart. Phys.*, 5, 15.
- Shandarin, S., Feldman, H. A., Heitmann, K., & Habib, S. (2006). Shapes and sizes of voids in the Lambda cold dark matter universe: Excursion set approach. *MNRAS*, 367(4), 1629–1640.
- Shandarin, S., Habib, S., & Heitmann, K. (2010). Origin of the cosmic network in LambdaCDM: Nature vs nurture. *Phys. Rev. D*, 81(10), 1–11.
- Shandarin, S., Habib, S., & Heitmann, K. (2012). Cosmic web, multistream flows, and tessellations. *Phys. Rev. D*, 85(8), 1–15.
- Shandarin, S. & Zeldovich, I. (1983). Topology of the large-scale structure of the universe. *Comments Astrophys.*, 10, 33–45.
- Shandarin, S. F. (1983a). Evolution of Potential Perturbations after Decoupling (The Adiabatic Scenario). In *Orig. Evol. Galaxies Proc. NATO Adv. Study Inst. held Erice, Italy, May 11–23, 1981. VIIth Course Int. Sch. Cosmol. Gravit.* (pp. 171–178). Dordrecht: Springer Netherlands.

- Shandarin, S. F. (1983b). Percolation theory. *Pis'ma Astron Zh.*, 53(9), 195–199.
- Shandarin, S. F. & Medvedev, M. V. (2017). The features of the Cosmic Web unveiled by the flip-flop field. *MNRAS*, 468(4), 4056–4076.
- Shandarin, S. F., Sheth, J. V., & Sahni, V. (2004). Morphology of the supercluster-void network in ??CDM cosmology. *MNRAS*, 353(1), 162–178.
- Shandarin, S. F. & Zeldovich, Y. B. (1984). topological mapping. *Phys. Rev. Lett.*, 52(17), 1488–1491.
- Shandarin, S. F. & Zeldovich, Y. B. (1989). Large scale structure of the Universe:. *Rev. Mod. Phys.*, 61(2), 185–220.
- Shandarin, S. F. S. & Medvedev, M. M. V. (2014). Tracing the Cosmic Web substructure with Lagrangian submanifold.
- Sheth, J. V., Sahni, V., Shandarin, S. F., & Sathyaprakash, B. S. (2003). SURFGEN. *MNRAS*, 343, 22–46.
- Shivshankar, N., Pranav, P., Natarajan, V., van de Weygaert, R., Bos, E. G. P., & Rieder, S. (2015). Felix: A Topology based Framework for Visual Exploration of Cosmic Filaments. *IEEE Trans. Vis. Comput. Graph.*, X(X), 1.
- Skory, S., Turk, M. J., Norman, M. L., & Coil, A. L. (2010). Parallel Hop: a Scalable Halo Finder for Massive Cosmological Data Sets. *ApJS*, 191(1), 43–57.
- Soneira, R. M. & Peebles, P. J. E. (1978). A computer model universe - Simulation of the nature of the galaxy distribution in the Lick catalog. *Astron. J.*, 83, 845.
- Sousbie, T. (2011). The persistent cosmic web and its filamentary structure - I. Theory and implementation. *MNRAS*, 414(1), 350–383.

- Sousbie, T. & Colombi, S. (2015). ColDICE: a parallel Vlasov-Poisson solver using moving adaptive simplicial tessellation.
- Sousbie, T., Pichon, C., Colombi, S., Novikov, D., & Pogosyan, D. (2008). The 3D skeleton: Tracing the filamentary structure of the Universe. *MNRAS*, 383(4), 1655–1670.
- Sousbie, T., Pichon, C., & Kawahara, H. (2011). The persistent cosmic web and its filamentary structure - II. Illustrations. *MNRAS*, 414(1), 384–403.
- Springel, V. (2005). The cosmological simulation code GADGET-2. *MNRAS*, 364(4), 1105–1134.
- Springel, V., White, S., Tormen, G., & Kauffmann, G. (2001a). Populating a cluster of galaxies - I. Results at z=0. *MNRAS*, 328(February), 726–750.
- Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., & Pearce, F. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435(7042), 629–636.
- Springel, V., Yoshida, N., & White, S. D. (2001b). GADGET: a code for collisionless and gasdynamical cosmological simulations. *New Astron.*, 6(2), 79–117.
- Sutter, P. M., Lavaux, G., Hamaus, N., Pisani, A., Wandelt, B. D., Warren, M., Villaescusa-Navarro, F., Zivick, P., Mao, Q., & Thompson, B. B. (2015). VIDE: The Void IDentification and Examination toolkit. *Astron. Comput.*, 9, 1–9.
- Sutter, P. M. & Ricker, P. M. (2010). Examining Subgrid Models of Supermassive Black Holes in Cosmological Simulation. *ApJ*, 723(2), 1308–1318.
- Tegmark, M., Strauss, M., Blanton, M., Abazajian, K., Dodelson, S., Sandvik, H., Wang, X., Weinberg, D., Zehavi, I., Bahcall, N., Hoyle, F., Schlegel, D., Scoccimarro, R., Vogeley, M., Berlind, a., Budavari, T., Connolly, a., Eisenstein, D., Finkbeiner, D., Frieman, J., Gunn, J., Hui, L., Jain, B., Johnston, D., Kent, S., Lin, H., Nakajima, R., Nichol, R., Ostriker, J., Pope,

- a., Scranton, R., Seljak, U., Sheth, R., Stebbins, a., Szalay, a., Szapudi, I., Xu, Y., & Others, . (2003). Cosmological parameters from SDSS and WMAP. *Phys. Rev. D*, 69(10), 1–27.
- Tweed, D., Devriendt, J., Blaizot, J., Colombi, S., & Slyz, A. (2009). Building merger trees from cosmological N -body simulations. *Astron. Astrophys.*, 506(2), 647–660.
- van de Weygaert, R., Aragon-Calvo, M. A., Jones, B. J. T., & Platen, E. (2009). Geometry and Morphology of the Cosmic Web: Analyzing Spatial Patterns in the Universe.
- van de Weygaert, R. & Bond, J. (2008a). Observations and Morphology of the Cosmic Web. In M. Plionis, O. López-Cruz, & D. Hughes (Eds.), *A Pan-Chromatic View Clust. Galaxies Large-Scale Struct.*, volume 740 of *Lecture Notes in Physics*, Berlin Springer Verlag (pp. 24).
- van de Weygaert, R. & Bond, J. R. (2008b). Clusters and the Theory of the Cosmic Web. In *A Pan-Chromatic View Clust. Galaxies Large-Scale Struct.* (pp. 335–408). Dordrecht: Springer Netherlands.
- van de Weygaert, R. & Schaap, W. (2009). Data Analysis in Cosmology: The Cosmic Web: Geometric Analysis. *Lect. Notes Phys.*, 665(September), 291–413.
- van Kampen, E. (1995). Improved numerical model. *MNRAS*, 273, 295–327.
- Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Nelson, D., & Hernquist, L. (2014). Introducing the Illustris Project : Simulating the coevolution of dark and visible matter in the Universe. *MNRAS*, 444(2), 1518–1547.
- Vogelsberger, M. & White, S. D. M. (2011). Streams and caustics: The fine-grained structure of Lambda cold dark matter haloes. *MNRAS*, 413(2), 1419–1438.
- Weinberg, D. H., Hernquist, L., & Katz, N. (1997). Photoionization, Numerical Resolution, and Galaxy Formation. *ApJ*, 477, 8–20.
- Wojtak, R., Powell, D., & Abel, T. (2016). Voids in cosmological simulations over cosmic time. *MNRAS*, 458, 4431–4442.

Zeldovich, Y. B. (1970). Gravitational Instability: An approximate theory for Large Density Perturbations. *Astron. Astrophys.*, 5(1), 84–89.

Zeldovich, Y. B., Einasto, J., & Shandarin, S. F. (1982). Giant voids in the Universe. *Nature*, 300(5891), 407–413.

## **Appendix A**

### **My Appendix, Next to my Spleen**

There could be lots of stuff here

## Appendix B

### The multistream flow field in one-dimension

The top panel in Figure B.1 shows the velocity multistreaming phenomenon in a one-dimensional collapse. The phase-space  $(\mathbf{p}, \mathbf{x})$  (where  $p$  is the momentum and  $x$  is the co-moving Eulerian coordinate) is single-valued in the linear stage of evolution (at redshift  $z_{ini}$ ). Non-linear stage of gravitational evolution of the collision-less dark matter particles then results in multi-valued  $\mathbf{p}(\mathbf{x}, z)$  at  $z = 0$ . The mass particles are sparsely distributed outside the region of gravitational collapse, and are denser in the inner streams.

A dynamically equivalent transformation  $(\mathbf{p}, \mathbf{x}) \mapsto (\mathbf{q}, \mathbf{x})$  (where  $\mathbf{q}$  is the Lagrangian coordinate) shows the Lagrangian sub-manifold in the middle panel of Figure B.1. This two-dimensional phase-space has foldings that correspond to multiple velocity streams, although the sub-manifold itself remains continuous. A projection of the Lagrangian sub-manifold at each point in the configuration space quantifies the number-of-streams. Folding in the sub-manifold are checked for points in configuration space using tessellations. The tessellating simplices in one-dimensional model are just the line-segments whose nodes are the dark matter particles in the Lagrangian space. Dynamical property is accounted for in this phase-space tessellation since labels of the nodes remain intact throughout the evolution; the line segments may shorten, extend or change orientation. Each folding in the Lagrangian sub-manifold increases the number of streams by a factor of two. In three-dimensional simulations, the sub-manifold twists in complicated ways in a six-dimensional

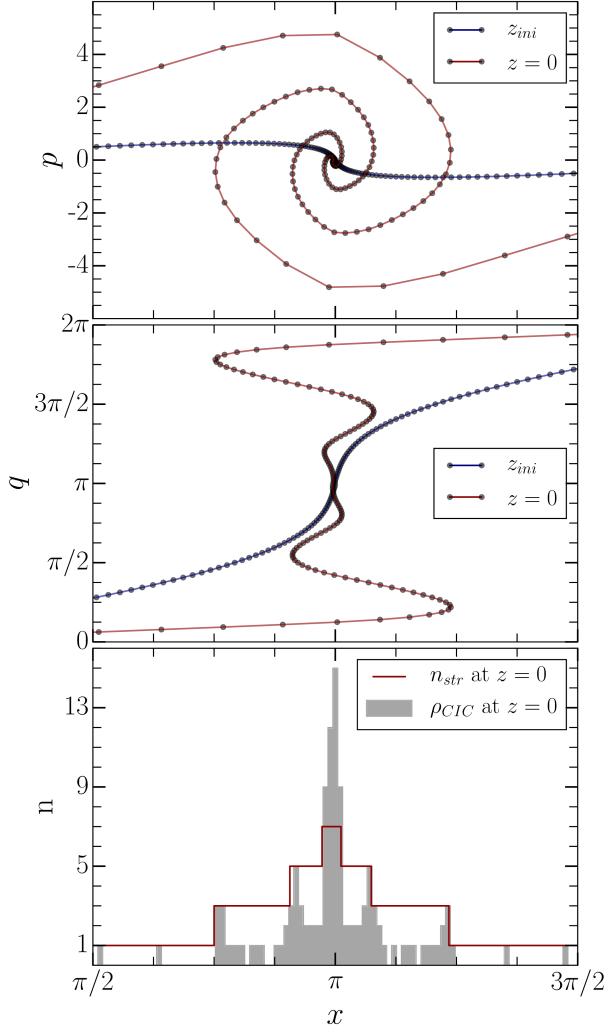


Figure B.1: Multi-streaming in one-dimension gravitational collapse. Top panel:  $(\mathbf{p}, \mathbf{x})$  phase-space representation redshift  $z_{ini}$  and  $z = 0$ . Dots represent the dark matter particles. Initially the mass particles are in the linear stage of evolution. At  $z = 0$ , multiple values of  $\mathbf{p}(\mathbf{x})$  is seen in the collapsed regions. Middle panel: Equivalent Lagrangian sub-manifold  $\mathbf{q}(\mathbf{x})$ . At  $z_{ini}$ , the dashed line represents  $\mathbf{q} = \mathbf{x}$ . Number of streams are parametrized from this sub-manifold. Bottom panel: The multistream field  $n_{str}$  and the number-density using CIC algorithm,  $n_{CIC}$  at  $z = 0$ .

phase space. The number-of-streams in N-body simulations (Shandarin et al. 2012 and ?) is calculated using Lagrangian/phase-space tessellations. This triangulation is conceptually different from the Voronoi (See Schaap & van de Weygaert 2000 and references therein) or Delaunay (Icke & van de Weygaert, 1991) tessellation schemes.

The bottom panel Figure B.1 shows the multistream field  $n_{str}(\mathbf{x})$  at  $z = 0$ . The field only takes the values of 1, 3, 5 and 7 in this scenario. Caustics occur at the folds in Lagrangian submanifold, and have a measure zero (study of caustics in one- and two-dimensional evolution is done in Hidding et al. (2014), three-dimensional caustic surface in a cosmological simulation is shown in Figure 4.12). Several properties of the multistream field are significantly different from mass density. The bottom panel also shows an illustration of CIC algorithm (cf. Hockney & Eastwood 1988) in calculating density, which is numerically equivalent to counting the number of particles on each cell of a regular grid. One major difference is in the regions before gravitational collapse:  $n_{str}$  is universally equal to unity, whereas number density fluctuates. It should also be noted that density by definition is a continuous field; numerical approximations like CIC discretise the field. Alternatively, multistream field is intrinsically a discrete-data field.

## B.1 Variations in the multistream field

A second-order local variations of a scalar field  $f$  is described by a Hessian. In a three-dimensional domain, the Hessian is given by Equation 5.1. The geometry of the scalar field is classified by the Eigenvalues of the Hessian. The convex regions have at-most one maxima within the (3+1)-dimensional functional space. Projection of this closed region onto three-dimensional coordinate space also gives a closed surface in coordinate space.

We treat  $n_{str}$  approximately continuous, for which the Hessian is always symmetric. In this study we use the scalar field  $n_{str}(\mathbf{x})$  inherently has discrete values like 1, 3, 5, and so on. The equation for numerical differentiation in the off-diagonal terms using Forward-difference method (using step-sizes of  $\Delta x_i$  and  $\Delta x_j$  along  $i$  and  $j$  respectively) is given in Equation B.1. Notice that

$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ , since RHS in Equation B.1 remains same. Hence the Hessian matrix in Equation 5.1 for the discrete scalar field  $n_{str}$  is always numerically symmetric. Backward or central difference give similar results too. Smoothing of the multistream field further reduces any numerical noise in the Hessian eigenvalues.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{1}{\Delta x_i \Delta x_j} [f_{i+1,j+1,k} - f_{i,j+1,k} - f_{i+1,j,k} + f_{i,j,k}] \quad (\text{B.1})$$

An integer-valued function, like the multistream field, is either constant or changes by a constant value in its real domain. In addition, the transitions in the multistream field are of multiples of 2, unless caustic surfaces are detected at the exact grid location. Consider  $f_{i,j,k} = n$  at any grid point. Due to the property of multistream field, the values in the neighbourhood differ by a multiple of 2. That is,  $f_{i+1,j,k} = n + 2p$ ,  $f_{i,j+1,k} = n + 2q$ ,  $f_{i+1,j+1,k} = n + 2r$ , for some integers  $p, q$  and  $r$ . Thus the second order variation of the multistream field reduces to Equation B.2.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{1}{\Delta x_i \Delta x_j} [2r - 2p + 2q] \quad (\text{B.2})$$

Thus the numerical differentiation is independent of  $n_{str}$  itself. It's important to note that this behaviour of the multistream field is independent of grid size. Also, the second order variation is a ratio of an even-number and the face area of the grid cube. The Equation B.2 becomes zero in a trivial case of  $r = p = q = 0$ , which corresponds to regions where  $n_{str}$  is constant, including voids. In the non-trivial case,  $r = (p + q)$ , for non-zero  $r, p$  and  $q$ . In the multistream grid,  $2(p + q)$  could be considered as sum of variations in  $n_{str}$  in the immediate neighbouring grid points. And  $2r$  is the variation between next closest grid point, which is along the face-diagonal.

On the other hand, mass density fields have sharp peaks at the multistream transitions. These peaks in the at the location of caustic are far less predictable, since the density fields become extremely noisy. For instance,? show noisy peaks of varying magnitude at the at high resolutions of mean density near halo locations. At lower resolutions, these sharp peaks are smoothed out, hence giving the impression of a smooth field. Hahn et al. (2015) show similar ‘ill-behaved’

derivatives in velocity fields at the caustic locations, where the derivatives are infinite.

# Appendix C

## Hessian signatures of the multistream field

Second-order local variations of a scalar field  $f$  is described by a Hessian matrix, whose element in a three-dimensional domain is given by Equation 5.1. The geometry of the scalar field is classified by the Eigenvalues of the Hessian. The convex regions have at-most one maxima within the (3+1)-dimensional functional space. Projection of this closed region onto three-dimensional coordinate space also gives a closed surface in coordinate space. An illustration of the projection is shown in Figure C.1 for a simpler function  $f(x)$  in one-dimensional domain. The eigenvalue criteria for regions are simplified: for instance,  $\frac{\partial^2 f}{\partial x^2} < 0$  for convex region. Projection of these regions onto coordinate space is shown in the shaded regions. This is different from regions within a contour, which is the projection of the curve along which the function has a constant value. Boundaries of these two regions may, but not necessarily, intersect.

In the case of cosmic fields, thresholds like  $\Delta_{vir}$  are equivalent to the green dotted line in Figure C.1. The over-dense regions (green shaded regions) are not constrained to be convex. Similarly structures selected based on  $n_{str}$  thresholds do not universally result in convex structures either. Local geometry can be probed from the eigenvalue criteria instead, as shown by the red line on the curve and corresponding shaded area. The projected structures, albeit convex, may have very small values of  $f(x)$  (like the red shaded area around  $x = 5$ ). In the framework of identifying potential haloes in multistream field, multistream thresholds are devised in so that some of these

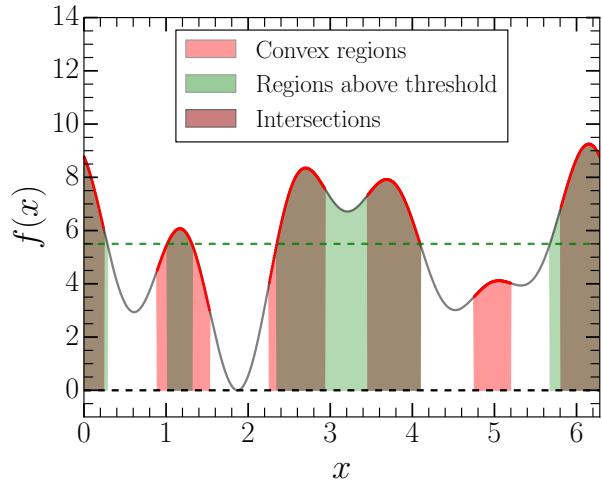


Figure C.1: Projections of regions of  $f(x)$  from (1+1)-dimensional function space onto one-dimensional coordinate space. Convex regions and regions above a threshold of an arbitrary function  $f(x)$  are shown. Both the regions intersect around a few maxima, but not universally.

small peaks detected by the Hessian are not considered as potential halo sites.