Contents

List of Tables vi

List of Figures vii

Introduction 1

1 Literature review 6
1.1 Algorithms . . . 6
1.2 Theoretical basis . . . 6
1.3 Mock galaxy data . . . 8
1.4 Visualizations . . . 10

2 Background 11
2.1 The original Vocludet Algorithm . . . 11
2.1.1 Description . . . 11
2.2 The improved algorithm . . . 15
2.2.1 Problem identification . . . 15
2.2.2 GapperR200 Algorithm . . . 15

3 Analysis and Visualization software 20
3.1 Analysis . . . 20
3.1.1 Results module and graphing library . . . 20
3.1.2 Graphing library . . . 20
3.2 Visualization Software . . . 21
3.2.1 Software architecture . . . 21
3.2.2 Backend . . . 23
3.2.3 Server . . . 23
3.2.4 Frontend . . . 24
3.2.5 Potential uses of the tool . . . 29

4 Optimization 30
4.1 Mock galaxy data . . . 30
4.2 Parameter optimization . . . 33
4.3 Domain size . . . 33
4.4 Velocity gap . . . 35

5 Analysis and validation 40

5.1 Analysis . . . 40
5.1.1 VTMLE: the geometrical step . . . 40
5.1.2 Number of galaxies . . . 43
5.1.3 Cluster Mass . . . 43
5.1.4 Redshift dependency . . . 48
5.1.5 False positive rate . . . 50
5.1.6 The GapperR200 stage . . . 53
5.2 Validation . . . 54
5.2.1 Resulting catalog . . . 54
5.2.2 Completeness and purity . . . 55
5.2.3 Velocity dispersions . . . 60
5.2.4 Comparison with other works . . . 63

Conclusions **63**

Bibliography **66**

List of Tables

4.1 Summary of the Velocity-Gap optimization results 39

5.1 Summary of VTMLE recovery rates 53
5.2 Breakdown of Vocludet detections (1614) by ranges of multiplicity and redshift 55

List of Figures

1.1 Example of a 2D Voronoi tessellation applied on a previously detected VTMLE cluster. Each point represents a galaxy and the dashed lines the Voronoi tessellation. The polygon is the convex hull of the cluster. In this figure it can be seen the relation between cell size and local density, since the cells get smaller the closer to the denser central region they are. The actual algorithm uses a 3D tessellation. 9

3.1 Visualization software overall architecture diagram 22
3.2 Visualization software backend data structure 23
3.3 Visualization software screen capture. It shows the view from inside the 1 Mpc radius cone where a combination of radial and angular distribution can be seen. 25
3.4 Visualization software screen capture. It shows the view from outside the 1 Mpc radius cone where the radial distribution is better displayed. 26
3.5 Visualization software screen capture. It shows the view from the center of the cluster cones where the angular distribution is best appreciated. 26
3.6 Visualization software screen capture. It shows the view from outside the 1 Mpc radius cone using a higher field of view than figure 3.5, which helps separate the cluster from the nearby galaxies. 27
3.7 Lens flare texture used in the visualization software 27
3.8 Visualization software control panel 28
3.9 Sample histogram using bin width $\sim 50 kms^{-1}$ 29
3.10 Sample histogram using bin width $\sim 100 kms^{-1}$ 29

4.1 Wedge diagram showing the mock 2dF database, restricted to $z < 0.14$, based on the Millennium Simulation 31
4.2 Reference catalog cluster mass distribution. N indicates the number of clusters per bin. 32
4.3 VTMLE recovery rate by domain size. It can be seen that for clusters up to 30 galaxies, the best results are obtained with a domain size of $35 Mpc$ . . . 34
4.4 Velocity dispersion box plot of multiple values of velocity gap, including the Millennium reference catalog. The bottom and top of the boxes represent the first and third quartiles, and the band inside the boxes indicates the second quartile (the median). The bottom and top ends of the vertical lines indicate the minimum and maximum values, respectively. 35

4.5 Vocludet completeness rate (fraction of reference clusters recovered) by minimum number of galaxies in the Millennium cluster (Ng_{mil}). To consider a match, the galaxy overlap must be of at least 25%. 36
4.6 Vocludet purity rate (fraction of detected clusters which have a match in the reference catalog) by minimum number of galaxies in the Millennium cluster. To consider a match, the galaxy overlap must be of at least 25%. 37
4.7 Vocludet vs Millennium reference catalog velocity dispersion for different values of velocity gap. The red lines are lines of slope 1 passing through the origin. Figures to the left include all valid vocludet clusters while figure to the right include only the ones with 10 or more galaxies. 38

5.1 VTMLE cluster distance to closest Millennium cluster 41
5.2 Cumulative distribution function of the VTMLE intersecting clusters projected distance, with respect to Millennium reference clusters. Red vertical lines indicate values 0.5 and 0.75, respectively 42
5.3 VTMLE recovery rates by number of galaxies in Millennium clusters 44
5.4 VTMLE clusters recovered by number of galaxies in Millennium clusters. Dashed lines represent the mean recovery rates of the respective detection quality 45
5.5 VTMLE recovery rates by minimum mass of Millennium clusters 46
5.6 VTMLE recovery rates by mass value of Millennium clusters 47
5.7 VTMLE recovery rates by redshift interval of Millennium clusters 48
5.8 VTMLE clusters recovered by redshift interval of Millennium clusters 49
5.9 VTMLE false positive detections by seed ranking 51
5.10 Cumulative distribution function of distance between VTMLE detected clusters and the closest reference cluster, including executions with and without error introduced. 52
5.11 Millennium and Vocludet clusters. For Vocludet, solid lines are the total, dotted are Type I clusters and dashed lines are Type II clusters. 54
5.12 Distribution of number of galaxies per Vocludet cluster, up to 50 galaxies. There are 22 clusters with $N_{gal} > 50$ not included in the histogram. Solid lines are the total, dotted are Type I clusters and dashed are Type II clusters. 56
5.13 Distribution of projected radii of Vocludet clusters. Solid lines are the total, dotted are Type I clusters and dashed are Type II clusters. 57
5.14 $N \geq 10$ Vocludet-Millennium clusters, multiplicity comparison. Crosses are Type I clusters and open circles Type II. 58
5.15 $N \geq 10$ Vocludet-Millennium clusters, σ_v comparison. Crosses are Type I clusters and open circles Type II. 59
5.16 Vocludet clusters galaxy completeness histogram 60
5.17 Millennium clusters recovery (dotted) and purity (solid) rate by cluster multiplicity range. 61
5.18 Millennium clusters recovery (dotted) and purity (solid) rate by cluster mass range. 62