Programa de Doctorat d’Astronomia i Meteorologia

Bienni 2004–2006

Memòria presentada per M. Carmen Toribio Pérez per optar al grau de Doctora per la Universitat de Barcelona

Director de la tesi

Dr. Josep Maria Solanes Majúa
Agraïments-Agradecimientos-Acknowledgements

I would like to express my gratitude to all the people that have contributed, in one way or another, to the realization of this work. In particular, I would like to specially acknowledge the help of the following people:

En primer lloc, voldria agrair a en Josep M. Solanes haver dirigit aquesta tesi. Gràcies per la teva orientació, la intensa dedicació i els teus esforços per dur a bon port aquest treball.

I would like to thank Juan M. Uson for his unvaluable guiddance with the VLA observations and his hospitality while I was in Charlottesville. I appreciate your generosity and all the time you devoted to carefully answer my countless questions.

I would like to thank the ALFALFA team and their PIs, Riccardo Giovanelli and Martha P. Haynes, for allowing me to participate in the project, and also give special thanks to those members of the ExtraGalactic Group I met at Cornell University: Ann Martin, Sabrina Stierwalt, Brian Kent, and Betsey Adams. Thank you all for your help and for keeping me warm during Ithaca winter.

Additionally, I would like to take this opportunity to thank all the members of the IDILICO collaboration.

Quisiera agradecer a José M. Vílchez, Jorge Iglesias y Vasiliki Petropoulou el precioso tiempo que me habéis dedicado y los buenos momentos pasados en el Instituto de Astrofísica de Andalucía (IAA) y en Calar Alto. Quiero hacer extensivo este agradecimiento a todas las personas que he ido conociendo en el IAA en mis sucesivas visitas: Lourdes Verdes-Montenegro y su grupo, Jaime Perea, Gabriella Gilli y Silbia López, entre otros. Gracias a todos por la estupenda acogida que me habéis ofrecido.

Vull expressar el meu agraïment a tots els membres del Departament d’Astronomia i Meteorologia (DAM) de la Universitat de Barcelona, on he re-
alitzat aquesta tesi. Gràcies a l’Eduard Salvador, Francesca Figueras, Carme Jordi, Xavier Luri, José M. Torrelles, als membres de suport (Gaby, Jordi, Ruben i Joaquim), i als nostres secretaris (JR, Montse i Rosa). Vull agrair també a tots els doctorands i doctors, i d’altres enginyers, els bons moments compartits: Rosa, Pol, Gemma, Sinue, Jordi, Josep M., Adolfo, Iñigo, Javi(s), Álvaro, Víctor, Teresa, Óscar, Yago, Laura, Maria, Mercè, Albert, Pere, Pau, Aidan, Cristina, Andreu, Neus, Àngels, Aina, Josep Manel, Salvador, Guillem, Francesc... a tots els DAMeros en general, i als companys de l’IEEC: Moi, Carlo, Anna, Juan Carlos i Felipe.

Quiero dar las gracias a Guillermo González, por ayudarme en los primeros pasos. Y a Begoña Ascaso. Mil gracias, Bego. Es una suerte tenerte de hermana mayor.


Gracias a mis padres y a mis hermanas por quererme y cuidarme, i deixarme fer. Sin vuestro apoyo no podría haber llevado a cabo esta tesis. Ni esta tesis, ni ninguna otra.

Óscar, ya tú sabes.

En fi, sense l’ajuda i suport de tots i cadascun de vosaltres, no estaria escrivint ara aquestes línies. Per això, aquesta tesi és tan meva com vostra.
Contents

List of Figures vi

List of Tables ix

Overview xiii

Resum xvii

1 Introduction 1

1.1 The 21-cm transition line .................. 2

1.2 Extragalactic H I observations .............. 4

1.3 Tully-Fisher distance estimates .............. 5

1.4 The H I deficiency parameter ................. 8

1.5 Environmental effects on the H I of galaxies . 10

1.6 The Virgo Cluster Region .................. 15

1.6.1 Overall structure .......................... 16

1.6.2 Distribution of morphological types ........ 20

1.7 H I studies of the Virgo Cluster ............. 22
I VLA observations of gas-poor spirals in the Virgo Cluster region 31

2 H I Distribution and Tully–Fisher Distances 33

2.1 Introduction .............................................. 34
2.2 Galaxy selection ........................................... 36
2.3 Data acquisition and processing .......................... 39
    2.3.1 Observations ........................................ 39
    2.3.2 Data calibration ...................................... 41
    2.3.3 Continuum emission subtraction ..................... 43
    2.3.4 H I synthesis results ................................. 44
        2.3.4.1 Channel maps .................................. 44
        2.3.4.2 Moments ...................................... 44
        2.3.4.3 Global H I profiles ............................ 53
        2.3.4.4 Position-velocity diagrams and rotation curves 55
2.4 Case by case analysis of the H I synthesis data ........... 56
    2.4.1 NGC 4307 .............................................. 66
    2.4.2 NGC 4356 and VCC 740 ............................. 67
    2.4.3 NGC 4411A/B, VCC 933 and VCC 976 ............... 68
    2.4.4 NGC 4492 .............................................. 70
2.5 TF-based distance estimates to our galaxies ............... 71
2.6 Summary and conclusions ................................... 76
II Standards of HI content from the ALFALFA Survey

3 Selection of a control sample

3.1 Introduction ........................................... 82
3.2 Main 21-cm and optical catalogs ......................... 84
  3.2.1 Statistical Properties of Catalogs and Cross-identification of Sources .................. 84
  3.2.2 Cosmic velocity and radial distance of sources ...... 91
3.3 HI control samples ...................................... 93
  3.3.1 Local number density measurements for ALFALFA and SDSS galaxies .................. 94
  3.3.2 Low Density Environment HI galaxy sample ......... 97
  3.3.3 Isolated galaxy samples ............................. 101
3.4 Properties of the HI control samples ...................... 104
  3.4.1 Structure, colors, and morphology of the galaxies in LDEs ................................. 104
  3.4.2 Structure, colors, and morphology of the isolated galaxies .............................. 111
3.5 Summary and conclusions .................................. 112

4 Multivariate analysis of a flux-limited database

4.1 Introduction ........................................... 116
4.2 Selection of galaxy properties ........................... 117
4.3 Relationships among HI and optical properties ............ 122
  4.3.1 PCA results ....................................... 123
  4.3.2 Standards of HI content ............................ 128
4.3.3 Planar correlation diagrams . . . . . . . . . . . . . . 129

4.4 Summary and conclusions . . . . . . . . . . . . . . . . . . 136

5 Conclusions and future perspectives 139

A In search of red, concentrated, gas-healthy galaxies in Low Density Environments 145

B Hubble test for the ALFALFA catalog 149
List of Figures

1.1 HI deficiency vs. projected clustercentric distance for HI-deficient clusters. .................................. 12
1.2 Galaxies in the Virgo Cluster Catalog .............................. 17
1.3 Comparison between the X-ray image and the optical appearance of the Virgo Cluster. ....................... 19
1.4 Distribution of morphological types in the Virgo Cluster .... 21
1.5 2D distribution of HI deficiency in the Virgo Cluster region. 23
1.6 3D distribution of HI deficiency in the Virgo Cluster region. 24
1.7 3D distribution of HI deficiency in the Virgo Cluster region: voxel projection. .................................. 26
1.8 Velocity-distance diagrams of simulations and observations in the Virgo Cluster region. ......................... 27

2.1 Sky distribution of our 4 main VLA targets ....................... 36
2.2 NGC 4307 channelmaps .............................................. 46
2.2 Continued .............................................................. 47
2.3 NGC 4356 channelmaps .............................................. 48
2.4 VCC 740 channelmaps ................................................. 49
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>NGC 4411 B/A channelmaps.</td>
<td>50</td>
</tr>
<tr>
<td>2.5</td>
<td>Continued.</td>
<td>51</td>
</tr>
<tr>
<td>2.6</td>
<td>NGC 4492 channelmaps.</td>
<td>52</td>
</tr>
<tr>
<td>2.7</td>
<td>H I synthesis results for NGC 4307.</td>
<td>57</td>
</tr>
<tr>
<td>2.7</td>
<td>Continued.</td>
<td>58</td>
</tr>
<tr>
<td>2.8</td>
<td>H I synthesis results for NGC 4356.</td>
<td>59</td>
</tr>
<tr>
<td>2.8</td>
<td>Continued.</td>
<td>60</td>
</tr>
<tr>
<td>2.8</td>
<td>Continued.</td>
<td>61</td>
</tr>
<tr>
<td>2.9</td>
<td>H I synthesis results for NGC 4411B and NGC 4411A.</td>
<td>62</td>
</tr>
<tr>
<td>2.9</td>
<td>Continued.</td>
<td>63</td>
</tr>
<tr>
<td>2.9</td>
<td>Continued.</td>
<td>64</td>
</tr>
<tr>
<td>2.10</td>
<td>H I synthesis results for NGC 4492.</td>
<td>65</td>
</tr>
<tr>
<td>2.11</td>
<td>Continued.</td>
<td>66</td>
</tr>
<tr>
<td>2.12</td>
<td>TF distances to the disk galaxies NGC 4411A (a) and NGC 4411B (b).</td>
<td>72</td>
</tr>
<tr>
<td>2.13</td>
<td>B-band TF templates considered in Sol02.</td>
<td>74</td>
</tr>
<tr>
<td>3.1</td>
<td>Wedge diagrams of ALFALFA and SDSS surveys.</td>
<td>86</td>
</tr>
<tr>
<td>3.2</td>
<td>Properties of the ALFALFA catalog.</td>
<td>88</td>
</tr>
<tr>
<td>3.3</td>
<td>Differences in position and velocity between ALFALFA detections and their SDSS counterparts.</td>
<td>89</td>
</tr>
<tr>
<td>3.4</td>
<td>Inverse light concentration index in the r band, $C_{59,r}$, versus $(u - r)$ for ALFALFA detections and the SDSS-spec sample.</td>
<td>92</td>
</tr>
<tr>
<td>3.5</td>
<td>SDSS 3D local density, $\rho_6$, vs. the projected density estimator calculated from the 2MASS galaxy distribution, $\mu_6$).</td>
<td>95</td>
</tr>
</tbody>
</table>
3.6 HI deficiency vs. local density in different velocity bins. . . . 98
3.7 Wedge diagrams of the ALFALFA LDE sample and detections excluded for lying in environments of higher density. . . . . . 100
3.8 HI deficiency vs. local density for the IG1 sample. . . . . . 103
3.9 \((u-r)\) color versus absolute magnitude in the r-band for ALFALFA galaxies in LDEs. . . . . . . . . . . . . . . . . . . . . . 107
3.10 \((u-r)\) color distribution for SDSS-spec galaxies in LDEs. . 108
3.11 \((u-r)\) color distribution for SDSS-spec galaxies obeying IG1 and IG2 criteria. . . . . . . . . . . . . . . . . . . . . . . . . . . 111
4.1 Planar correlation diagrams . . . . . . . . . . . . . . . . . . . . . 131
A.1 SDSS images of red \(((u-r) > 2.22)\), concentrated \((n > 4)\), and face-on \((b/a > 0.85)\) objects from our LDE HI galaxy sample. . . . . . . . . . . . . . . . . . . . . . . . . . . . 147
B.1 Hubble test for the ALFALFA catalog . . . . . . . . . . . . . . 150
List of Tables

1.1 Some popular gas sweeping mechanisms .......................... 13

2.1 Properties of the target galaxies ................................. 37

2.2 Summary of the observations ..................................... 42

2.3 Characteristics of the deconvolved images ...................... 45

2.4 Results from our VLA data and recent Arecibo measurements 54

3.1 Distributions of ALFALFA and SDSS Galaxies in Radial Velocity Bins ............................................ 90

3.2 ALFALFA Detections and SDSS Galaxies in LDEs Obeying Maller et al.’s Criterion ................................. 105

4.1 Principal Component Analysis of HI and Optical Properties for Weighted Data and Disattenuated Correlations .... 125

4.2 Principal Component Analysis of HI and Optical Properties for Weighted Data ....................................... 127

4.3 Coefficients of $M_{HI}$ Predictions from Single and Multiple Linear Regression Models ................................. 130

4.4 Coefficients of Orthogonal Fits between pairs of Fundamental Properties ............................................. 134
4.5  Comparison of the Central Slopes of Fundamental Scaling Laws Reported by Different Authors . . . . . . . . . . . . . . . 135
Overview

This dissertation deals with extragalactic detections of the 21-cm line of the neutral hydrogen atom (HI) obtained from both array and single dish radio telescopes. Among the wide range of topics that can be explored from extragalactic HI observations, which are one of the mainstays of our understanding of galaxy evolution and large-scale structure of the Universe, the work presented here is mostly framed within the scope of environmental studies. In particular, we present the results of two projects focused on galaxies lying in opposite environments. The first project consists on the study of the HI distribution and distance estimates of a handful of gas-poor spirals in the Virgo Cluster region from observations carried out with the Very Large Array (VLA). On the other hand, the second part of this thesis is devoted to analyze the optical properties and HI content of a sample of field galaxies detected by the Arecibo Legacy Fast ALFA (ALFALFA) blind 21-cm line survey. The aims of our research regarding these two projects are complementary: while the first part is focused on gaining insight on the effects that high density environments have on galaxies, the second part seeks to set up new standards of normalcy for the HI and optical properties of galaxies for future environmental studies. In what follows, we outline the way in which the presentation of this work has been organized, and a summary of the results obtained in each part.

In Chapter 1, a general introduction to the topics treated in this thesis is presented, giving special emphasis to the main precursor studies of our research. The aperture synthesis observations in the 21 cm line of pointings
centered on the Virgo Cluster region spirals NGC 4307, NGC 4356, NGC 4411B, and NGC 4492 using the VLA in its CS configuration are presented in Chapter 2. These galaxies were identified in a previous study of the three-dimensional distribution of HI emission in the Virgo region as objects with a substantial dearth of atomic gas and Tully–Fisher (TF) distance estimates that located them well outside the main body of the cluster. For these four objects, and two other galaxies (NGC 4411A and VCC 740) detected within the fields and observed bands of the main targets, we provide detailed information of the gas morphology and kinematics. Our new data confirm the strong H I deficiency of all the main targets but NGC 4411B, which is found to have a fairly normal neutral gas content. Besides, the VLA observations have also been used to discuss the applicability of TF techniques to the five largest spirals we have observed. This analysis leads us to conclude that none of them is actually suitable for a TF distance evaluation, whether due to the radical trimming of their neutral hydrogen disks, which is the case of the targets NGC 4307, NGC 4356, and NGC 4492, or to their nearly face-on orientation (NGC 4411A and B).

The statistical analysis of H I and optical properties of field galaxies (Part II) has been divided in two chapters. In Chapter 3, we report the results from the study of the H I content and stellar properties of nearby galaxies based on the ALFALFA survey observations on two declination strips covering a total area of 9 hr × 16 deg in the north Galactic hemisphere. From these data, we have assembled a large homogeneous sample of 5496 galaxies with healthy amounts of neutral gas located in environments of low local galactic density. These are all galaxies with a counterpart in the Sloan Digital Sky Survey (SDSS), which has also granted us access to a vast number of their optical attributes. In this chapter, we focus exclusively on the distributions of morphology indicators such as structural parameters and colors, as well as their relationship with H I emission in an effort to shed more light on the nature of extragalactic H I sources in the local universe. Our analysis confirms the extended view that red, passive objects of early type in the field tend to have less neutral hydrogen than bluer, star-forming late-type galaxies, a trend that gets emphasized with increasing distance in flux-limited datasets
due to the positive correlation between global color and intrinsic luminosity. We find, however, that this latter relationship involves exclusively galaxies undetected in H I, whereas the color distribution for gas-healthy galaxies is essentially luminosity-independent. Moreover, our study demonstrates that the best way to discriminate between early and late-type galaxies is through the presence of significant H I emission, since sources with red colors, a true early-type morphology and substantial reserves of cold gas are rare exceptions to the norm.

Two smaller homogeneous samples containing several hundreds of isolated ALFALFA galaxies each have also been extracted by applying different systematic identification criteria. These two subsets, which show neutral gas content distributions indistinguishable from those found in our larger dataset of galaxies in low density environments, are nonetheless characterized by a higher relative abundance of objects with optical early-type attributes. This last trait, also observed in other catalogs of isolated galaxies existing in the literature, highlights the selection bias arising from the aforementioned connection between color and luminosity and the application of methods for identifying isolated galaxy candidates partial to the brightest and, hence, the reddest members of a dataset.

In Chapter 4, we deal with a homogeneous subset of about three thousand objects extracted from the sample of H I-healthy galaxies in environments of low local galactic density assembled in Chapter 3 obeying the condition $F_{\text{H}I} \geq 1$ Jy km s$^{-1}$, for which we know that ALFALFA is complete. We apply strategies of multivariate data analysis to these gas-rich galaxies in order to i) investigate the correlation structure of their main extensive 21-cm and optical properties and reveal the true dimension of the space generated by them, ii) identify the intrinsic parameters that best define their H I content, and iii) explore the constraints that the joint distributions of our observables put on the scaling relations among fundamental variables. Our analysis confirms previous claims that gaseous galaxies form essentially a one-parameter set, as well as that the most precise predictors of their H I mass are the stellar diameter, followed by the total luminosity (both in the $r$-band), and the maximum rotation speed. The most interesting findings regarding the asso-
ciated scaling laws are that we get 90% confidence intervals of $-7.54^{+1.1}_{-0.86}$ for the slope of the relation between magnitude and log rotation speed when the analysis is restricted to the galaxies with the highest inclinations ($i \geq 60^\circ$), in good agreement with Tully-Fisher studies, and of $1.68^{+0.09}_{-0.08}$ for the log slope of the H I mass-galaxy size relation. Given the completeness, homogeneity, and broad morphological coverage of our galaxy database, the latter result suggests that the constancy of the global (hybrid) H I surface density that some authors advocate for the entire spiral population is just a crude approximation of reality.

Some ancillary results are relegated to Appendix A, where we display examples of rare, non-conventional, gas-rich early-type galaxies lying in low density regimes as well as to Appendix B, that is devoted to the Hubble test for the ALFALFA catalog. The main conclusions of our research are summarized Chapter 5, where we also put forward some aspects that should be addressed in the future on the basis of the work already done.

Finally, the rest of this preface corresponds to a summary of this dissertation in Catalan, official language of the Universitat de Barcelona.
La tesi que aquí es presenta tracta amb observacions extragalàctiques de la línia de 21 cm de l’hidrogen neutre (HI) obtingudes mitjançant interferòmetres i radiotelescopis d’una única antena. Entre les qüestions que es poden explorar a partir d’observacions extragalàctiques d’HI, les quals conformen un dels puntals del nostre coneixement pel que fa a l’evolució de les galàxies i l’estructura a gran escala de l’univers, el treball que presentem s’emmarca en l’àmbit dels estudis de l’entorn galàctic en les propietats fundamentals d’aquests objectes. En concret, presentem els resultats de dos projectes enfocats a l’estudi de galàxies en regions de densitats extremes. El primer d’aquests consisteix en l’estudi, a partir d’observacions de síntesi d’apertura obtingudes amb el Very Large Array (VLA), de la distribució de l’HI i les estimacions de distància d’un conjunt de galàxies pobres en gas en la regió del cúmul de Virgo. D’altra banda, en la segona part de la tesi, s’analitzen les propietats òptiques i del gas neutre d’una mostra de galàxies espirals en el camp detectades pel sondeig cec de fonts d’hidrogen neutre Arecibo Legacy Fast ALFA (ALFALFA). Els objectius de la recerca desenvolupada en aquests dos projectes són complementaris: mentre que en la primera part es persegueix ampliar els nostres coneixements en l’efecte que els entorns d’elevada densitat de l’univers tenen en les galàxies que hi habiten, la segona part cerca establir nous estàndards de normalitat en el contingut de gas neutre de les galàxies que es puguin utilitzar com a punt de partida en futurs estudis que explorin els efectes evolutius que pateixen les galàxies motivats per l’entorn en què resideixen.
A continuació adjuntem un breu resum dels dos projectes que conformen aquesta tesi.

I Observacions amb el VLA de galàxies pobres en gas a la regió del cúmul de Virgo

En aquesta primera part de la tesi (Capítol 2), presentem les observacions de síntesi d’apertura de la línia de 21 cm, dutes a terme amb el VLA en la seva configuració CS, de quatre apuntats a les galàxies espirals NGC 4307, NGC 4356, NGC 4411B i NGC 4492. Aquestes galàxies, que es troben a la regió del cúmul de Virgo, pertanyen a la mostra de tretze objectes seleccionats per Sanchis et al. (2004) amb deficiències gasoses importants, comparables a les mesurades en les espirals del centre de Virgo, però que aparentment es troben localitzades en la perifèria de l’estructura, d’acord amb les estimacions de les seves distàncies a partir del mètode de Tully & Fisher (1977). La selecció d’aquesta mostra es va fer en funció d’estudis previs de la distribució tridimensional del contingut en hidrogen neutre de les galàxies d’aquesta regió (Solanes et al. 2002), així com dels resultats d’estudis teòrics que exploraven la distància màxima que pot assolir una galàxia respecte el centre d’un cúmul un cop ha passat per aquest centre per primera vegada (Sanchis et al. 2002; Mamon et al. 2004). D’acord amb aquests estudis teòrics, els membres de la mostra seleccionada per Sanchis et al. (2004) tenien una baixa probabilitat d’haver patit interaccions amb el medi intracumular, llevat que les estimacions de les seves distàncies radials tinguessin associats errors molt per sobre dels estimats en un inici.

Per aquests quatre objectes i dues galàxies més (NGC 4411A and VCC 740) detectades en els cubs de dades associats a aquests, incloem informació detallada de la morfologia i cinemàtica del gas. Les nostres deteccions de síntesi d’apertura confirmen la forta deficiència en gas neutre tots els objectes principals, llevat per NGC 4411B, per la qual estimem un contingut en gas neutre relativament normal. A més, les dades obtingudes amb el VLA per avaluar l’aplicabilitat del mètode Tully & Fisher (1977) per obtenir es-
timacions de distàncies a les espirals detectades (llevat VCC 740). Aquest anàlisi ens porta a concloure que cap dels objectes estudiats es adient per aquest tipus d’estimacions, ja sigui per la presència de discs d’hidrogen fortement retallats, com és el cas de NGC 4307, NGC 4356 i NGC 4492, ja sigui per la seva baixa inclinació respecte la línia de visió (NGC 4411A i B).

II Estàndards de contingut en gas neutre a partir d’observacions del sondeig ALFALFA

La segona part de la tesi (Capítols 3-4) correspon a l’anàlisi anàlisi estadístic de les propietats de l’H I i propietats òptiques de galàxies en el camp a partir de les deteccions en la línia de 21 cm del sondeig ALFALFA (Giovanelli et al. 2005b) extretes d’una regió del cel de 9 hr × 16° en l’hemisferi nord Galàctic. Partint de les dades inicials en aquesta zona del cel, hem obtingut una mostra homogènia de 5496 galàxies relativament riques en hidrogen neutre en entorns de baixa densitat galàctica. Tots aquests objectes tenen una contrapartida òptica en el sondeig òptic Sloan Digital Sky Survey Data Release 7 (SDSS DR7; Abazajian et al. 2009), la qual cosa ens permet l’accés a un gran nombre de paràmetres òptics. En primera instància, i amb l’objectiu de caracteritzar en detall la naturalesa de les deteccions d’hidrogen neutre, hem inspeccionat les distribucions de diversos indicadors de morfologia, com ara colors i paràmetres estructurals. El nostre anàlisi confirma la visió general que les galàxies vermellloses i passives de tipus primerenc en el camp tendeixen a tenir menys gas neutre que les galàxies blavoses amb brots de formació estel·lar. Aquesta tendència s’emfatitza amb la distància en mostres limitades en flux, a causa de la correlació positiva entre lluminositat intrínseca i color. De tota manera, aquesta relació sembla afectar només a les galàxies no detectades en H I, mentre que la distribució de colors dels objectes rics en gas resulta essencialment independent de la lluminositat. D’altra banda, donat que objectes de tipus morfològic primerenc amb reserves substancials de gas són rares excepcions a la norma, el nostre estudi demostra que la detecció d’una emissió significativa en la línia de 21 cm és el millor discriminador
entre objectes de morfologia primerenca i objectes de tipus tardà.

Hem explorat també caracteritzacions alternatives del que hauria de ser una mostra de control formada per galàxies que presentin una mínima o nul·la evidència d’haver patit efectes ambientals en la seva evolució mitjançant l’aplicació de dos criteris diferents d’aïllament. El primer criteri, similar a l’aplicat per Prada et al. (2003), combina la informació espetroscòpica i fotomètrica per a seleccionar aquells objectes que, en un determinat volum en l’espai de velocitats, no tenen galàxies veïnes emphi dinàmicament rellevants. Una segona mostra de galàxies aïllades és obtinguda a partir d’un criteri estrictament fotomètric. En aquest cas, adoptem el criteri definit per Allam et al. (2005) – una adaptació del criteri de Karachentseva (1973) per a SDSS – basat en la comparació de diàmetres angulars aparents.

Les dues mostres de galàxies aïllades obtingudes, que presenten una distribució de continguts en gas neutre compatible amb la de la mostra de galàxies en entorns de baixa densitat, es caracteritzen per contenir una fracció relativament més alta d’objectes amb característiques primerenques. Aquest últim fet exemplifica els biaixos de selecció resultants d’aplicar criteris d’aïllament que trien candidats entre els objectes més brillants, i per tant, més vermells, d’una mostra. La constatació d’aquest biaix, juntament amb la mida relativament modesta de les mostres de galàxies aïllades i el reduït rang dinàmic que subtendeixen les propietats observades, ens indiquen que la utilizació de la densitat local és la millor opció a l’hora de seleccionar una mostra de control per al contingut d’H\textsubscript{I}.

Tenint en compte aquest fet, l’última part de la tesi analitza una submostra de les deteccions d’H\textsubscript{I} en entorns de baixa densitat, a les quals hem imposat la condició que el flux integral $F_{\text{H}\textsubscript{I}} \geq 1$ Jy km s\textsuperscript{-1}, valor pel qual sabem que el catàleg d’ALFALFA satisfà el test de Hubble (Apèndix B). Aquesta submostra, formada per unes tres mil galàxies, s’utilitza per dur a terme un anàlisi multivariable amb l’objectiu de i) investigar l’estructura de correlació entre els principals paràmetres obtinguts a partir de la línia de 21 cm i de l’òptic i derivar la dimensionalitat de l’espai que generen, ii) identificar els paràmetres intrínsecs de les galàxies que millor defineixen el seu
contingut d’H\textsc{i}, iii) explorar les restriccions que les distribucions multivariades entre els observables imposen en les relacions d’escala entre variables fonamentals.

El nostre anàlisi ha permès confirmar resultats previs pel que fa a la unidimensionalitat de l’espai dels paràmetres de les galàxies detectades en H\textsc{i}, així com el fet que el diàmetre estel·lar, lalluminositat total (ambdós en la banda $r$ del SDSS), i la velocitat màxima de rotació del gas, són, en aquest ordre, els millors predictors de la massa d’H\textsc{i}. Pel que fa a les relacions d’escala que obtenim del nostre espai de paràmetres, l’ajust del pendent entre la magnitud i la velocitat de rotació, quan restringim les inclinacions dels objectes a $i \geq 60^\circ$, proporcionen un valor de $-7.54^{+1.1}_{-0.86}$ en unitats de magnitud i per un interval de confiança del 90\%, en consonància amb els estudis Tully-Fisher. D’altra banda, el pendent obtingut per la relació d’escala entre la massa d’H\textsc{i} d’una galàxia i el seu diàmetre estel·lar, $1.68^{0.09}_{-0.08}$, pel mateix interval de confiança, suggereix que prèvies afirmacions sobre el valor constant de la densitat superficial (híbrida) de l’hidrogen neutre per la totalitat de la població de galàxies espirals és tan sols una aproximació.

La tesi inclou també dos curts apèndix que presenten alguns resultats complementaris relacionats amb la recerca dels pocs objectes de la mostra de galàxies en entorns de baixa densitat els quals, malgrat presentar característiques òptiques típiques d’objectes primerencs, contenen quantitats considerables d’hidrogen neutre (apèndix A), així com el test de Hubble pel catàleg de deteccions d’ALFALFA.
A general introduction to the topics treated in this thesis for the non-expert reader is presented in this chapter. In order give a general context, the basic physics of the 21-cm line is explained in Section 1.1, followed by a brief introduction to the field of extragalactic HI detections and its state of the art, which includes a selection of references to published reviews and key studies in the field (Section 1.2). The use of the Tully-Fisher relation as a distance indicator is summarized in Section 1.3, while in Section 1.4 we explain how the HI content of a galaxy is estimated through the use of the HI deficiency parameter. The effects of the environment on the HI content of galaxies are treated in Section 1.5, as well as in Section 1.7, where we pay special attention to the precursor studies carried out in the Virgo Cluster Region, whose overall structure is presented first in Section 1.6.

Unless otherwise stated, in this work we will assume a flat standard concordant CDM cosmology with parameters \((h, \Omega_M, \Omega_\Lambda) = (1, 0.3, 0.7)\). All sky coordinates are equatorial coordinates referred to the J2000 equinox.


1.1 The 21-cm transition line

The symbol ‘H1’ denotes the neutral hydrogen atom. Atomic hydrogen has two states in its ground level, corresponding to the electron spinning parallelly or anti-parallelly with respect to the nuclear spin. Parallel spins give a slightly higher energy than antiparallel ones, and decay by emission of a photon with a frequency $\nu_{H1} \sim 1420.406$ MHz (when transformed into a wavelength $\lambda_{H1} = c/\nu_{H1} \sim 21$ cm, where $c \sim 299792.5 \text{ km s}^{-1}$ is the speed of light) in the radio domain of electromagnetic spectrum.

In general, the HI photons emitted by a galaxy will be observed at a reddened wavelength because of the Doppler effect:

$$\lambda_{\text{observed}} = \lambda_{\text{emitted}}(z + 1),$$

where $z = v/c$ is the redshift of the galaxy, and $v$ its observed radial velocity.

The upper level of energy of the HI atom is populated by collisions, and has a decay half-life of about $3.5 \times 10^{14}$ seconds $\simeq 1.1 \times 10^{7}$ years. Given the vast number of HI atoms found in a typical galaxy, together with the transparency of our atmosphere to radio wavelengths from about 10 mm ($\sim 30$ GHz) to 60 m ($\sim 5$ MHz), the signal of this phenomenon can be detected with Earth-based radio telescopes for galaxies in the Local Universe ($z \lesssim 0.06$).\(^1\)

The signal from all the HI atoms traveling at velocities close to the mean velocity of the galaxy will result in a broadening of the 21-cm emission line, and the shape of the observed profile will depend on the distribution and the dominant movement of the atoms (e.g., Roberts 1978). If the motions of the atomic gas are ordered in a rotational disk, the corresponding signal will appear as a double-peaked shaped profile, with two horns formed from almost vertical boundaries, and a central depression in the central region (e.g. Figure 2.9). On the contrary, single-peaked profiles will be observed if random motions of atoms are comparable or exceed the ordered motion, or

\(^1\)HI detections are hazardous below 1300 MHz because of radio frequency interference. In fact, H1 has not been detected in emission at $z > 0.25$. 

1.1. The 21-cm transition line

if the HI distribution is centrally concentrated.

The HI profile emitted by the atomic gas can be used to estimate properties of the host galaxy. For instance, the linewidth of the profile, $W$, can be used to determine with great accuracy the central velocity of the system, $v_{\text{sys}}$. The maximum rotational velocity of the disk, $V_{\text{max}}$, can be estimated as well through

$$W \sim 2V_{\text{max}}$$ (1.1.2)

(see Roberts 1978, and Section 1.3).

As for the strength of the signal, one can integrate the intensity of the profile, $S(v)$, which is measured in units of Jansky (Jy)\(^2\), along the velocity axis $v$ to obtain the global flux

$$F_{\text{HI}} = \int S(v) \, dv,$$ (1.1.3)

usually expressed in units of Jy km s\(^{-1}\). Under some non very restrictive assumptions regarding the spin temperature and optical depth (see Giovanelli & Haynes 1988, for a discussion on these assumptions), this integrated line profile can be used to derive the total HI mass, $M_{\text{HI}}$, of the galaxy:

$$M_{\text{HI}} = 2.356 \times 10^5 \, d^2 \, F_{\text{HI}} \, M_\odot,$$ (1.1.4)

where $d$ is the distance to the galaxy in $h^{-1}$ Megaparsecs (Mpc)\(^3\) and $F_{\text{HI}}$, expressed in Jy km s\(^{-1}\), must be corrected for self-absorption and instrumental biases (details on the derivation of this formula are given by Roberts 1975). Since $M_{\text{HI}}$ is derived from the square of the distance, uncertainties arising from the scale parameter of the Universe $H_0$ can be made explicit if using the quantity $h^2 M_{\text{HI}}$, being $h = H_0/100$ km s\(^{-1}\)Mpc\(^{-1}\).

\(^{2}\)1 Jy = $10^{-26}$ W/(m\(^2\)⋅Hz) = $10^{-23}$ erg/(s⋅cm\(^2\)⋅Hz)

\(^{3}\)1 Mpc = 3.08568025 × $10^{22}$ m
1.2 Extragalactic H\textsc{i} observations

The history of the field of extragalactic neutral gas detections has been reviewed by Giovanelli & Haynes (1988), and more recently by Giovanelli (2005). As stressed in these reviews, the field of observations of atomic gas has impressively evolved since the first detection of Galactic H\textsc{i} (Ewen & Purcell 1951), and the extragalactic detection of the HI signal from the Magellanic Clouds two years later (Kerr & Hindman 1953).

During the late 1970s and 1980s, one decade after the first large single-dish telescopes got on the scene, the field experimented its most rapid growth thanks to the construction of synthesis instruments such as the Very Large Array and the Westerbork Synthesis Radio Telescope, and the upgrade of the Arecibo single-dish antenna, in parallel to significant improvements in the sensitivity and spectral resolution of the detections. Among the science topics that have been investigated from 21-cm line detections since then, I wish to highlight, for being the most related with this work: the redshift-independent distance estimates that can be obtained from the existing relation between the luminosity and the linewidth of the profile of a galaxy (Tully & Fisher 1977, Section 1.3), the study of the rotation curves of disk galaxies (e.g., Sofue & Rubin 2001, for a review), and the use of HI gas as a tracer of the impact of environment on galaxies (Section 1.5). In particular, we will introduce in Section 1.4 the measurement of the neutral gas content of the spiral galaxies in terms of an H\textsc{i}-deficiency parameter.

Although not dealt with in the present work, I also want to mention the fundamental studies regarding the large-scale structure of the Universe that have been carried out from redshift surveys (for reviews on this topic see, e.g, Giovanelli & Haynes 1991; Salzer & Haynes 1996). In particular, HI redshift surveys, that benefit from the great precision of galaxy velocities derived from HI profiles, are complementary to optical surveys in the sense that, while the latter preferentially favor objects of high central optical surface brightness, galaxies of low optical surface brightness tend to have abundant HI and, at equal apparent magnitude, are more easily detected by radio surveys. The
1.3. Tully-Fisher distance estimates

net result is that HI redshift surveys are specially applicable to objects that cluster least, whereas redshifts of galaxies in clusters – where lenticular and elliptical morphologies are most abundant – will be more easily detected in the optical bands.

Another important feature from which our research also benefits is the excellent results that are providing the last generation of multifeed arrays that have been installed in the telescopes. Current blind HI surveys are capable of mapping the extragalactic sky over cosmologically significant volumes. Our work (Chapters 3–4) will exploit the large sample of HI detections of the Arecibo Legacy Fast ALFA survey (ALFALFA Giovanelli et al. 2005a), that is mapping \( \sim 7000 \, \text{deg}^2 \) of the high galactic latitude sky visible from Arecibo, providing a HI census of the local universe covering the redshift range within \( \sim -2000 \) and \( 18000 \, \text{km s}^{-1} \).

1.3 Tully-Fisher distance estimates

We provide in this section a basic review of the Tully-Fisher relation (Tully & Fisher 1977), one of the most popular redshift-independent techniques used to estimate distances to galaxies (e.g., Jacoby et al. 1992, for a review on extragalactic distances). The reason is that one of the topics tackled in this thesis (Chapter 2) are distances derived through this method, which relies on the underlying relation of both the rotation speed and luminosity of a galaxy with its total mass.

Ignoring the bulge and using the definition of the circular velocity of a gravitationally bound system, we might expect that the total mass of a spiral galaxy, \( M_T \), to follow approximately

\[
M_T \propto R \cdot V_{\text{max}}^2 ,
\]

with \( V_{\text{max}} \) the maximum rotational velocity of its disk, and \( R \) the disk scale length. This relation is proved by applying the virial theorem (e.g., Binney & Tremaine 2008, for a derivation of the theorem). As for the input value
of $V_{\text{max}}$, which can be estimated from a global profile linewidth, $W$, or from rotation curve measurements, it is also required that the measured linewidths reflect only the effect of gravitation, and that the lines are not broadened by high optical depth. Therefore, if 21-cm data are used to estimate the total mass, as first was done by Roberts (1962), the measured linewidths must be properly corrected of observational effects: the inclination of the disk to the line of sight, the cosmological redshift, the turbulent motions of the gas, as well as instrumental corrections (e.g., Springob et al. 2005).

Provided that the total mass of a galaxy is directly proportional to its luminosity, $L$, and assuming that spiral galaxies all have the same central surface brightness $\Sigma_0$ so that $L \propto R^2 \cdot \Sigma_0$, then

$$L \propto V_{\text{max}}^n,$$

(1.3.2)

with $n = 4$. Observational fits to the latter correlation give $n$ in the range of 2.5-3 at optical wavelengths. There is a general trend that the slope steepens at longer wavelengths, climbing to $n \sim 4$ at infrared wavelengths.

An equivalent expression to the latter equation in terms of absolute bolometric magnitude, $M^4$, and profile linewidth, $W$, is

$$M = -a \log W + b,$$

(1.3.3)

where $M$ should be corrected for both the extinction by dust in our Galaxy and by dust in the target galaxy itself, while $W$ is corrected for all the observational effects mentioned above.

In 1977, Tully & Fisher were the first to realize that this intrinsic relation between luminosity and rotational velocity could be used to measure the distances to galaxies. The method they proposed consisted on first fitting the coefficients of Equation 1.3.3 for a set of calibrators, that is, galaxies whose distances were derived through other methods (e.g., Cepheids), and

$$L \over L_\odot = 10^{0.4(M-M_\odot)} ; L_\odot = 3.846 \times 10^{26} W$$
then use the fitted relation to estimate the absolute magnitude of any galaxy with an available linewidth measurement. Next, from the definition of the Distance Modulus, $DM$,

$$DM = m - M - (K + e),$$

where $m$ is the apparent magnitude, and $(K+e)$ is the term accounting for the K-correction and the correction for evolution (since the galaxies that will be analyzed in this work are within 18,000 km s$^{-1}$ in the local volume, we can ignore these corrections from now on), it is straightforward to infer the luminosity distance of the galaxies, $D_L^5$, from the relationship


Therefore, following Tully & Fisher (1977), one can estimate the radial distances for a sample of galaxies with the only information of the observables $W$ and $m$, as long as Equation 1.3.3 has been previously calibrated. However, in order to exploit the full potential of this method, the sample of objects must be selected applying certain observational constraints on these quantities. As for the rotational velocity of a disk, the best estimates are obtained for edge-on galaxies. However, because of extinction, the luminosity of edge-on systems could be underestimated. On the contrary, for face-on galaxies, whose luminosity measurements will be unbiased, the measured linewidths can be affected by projection effects. Thus, the selection of the sample galaxies is important, as well as the optical bands used to estimate their luminosity, being the infrared bands those less affected by absorption (e.g., Masters et al. 2003). On the other hand, the maximum rotational velocity of the disk can be estimated from H$\alpha$ or H$\text{I}$ lines (e.g. Vogt et al. 2004). Likewise, we must note that differences can also arise from the different methods to estimate this quantity (e.g., Verheijen & Sancisi 2001).

The literature regarding TF studies is very extensive, with countless studies that have been carried out on this topic along the last 30 years. For

---

5See Hogg’s notes on distance measures in cosmology (Hogg 1999).
instance, whereas the coefficients fitted to Equation 1.3.2 play a significant role on galaxy evolution (e.g., Zwaan et al. 1995), the the ever-increasing quality and size of galaxy samples have allowed to use this technique as a cosmological tool (see Jacoby et al. 1992, and references therein). From the early studies to date, TF distances to galaxies in the Virgo Cluster Region (e.g. Tully & Fisher 1977; KCT88 1988; Fou90 1990; FTS98 1998, see Chapter 2) and other nearby clusters (e.g., Giovanelli et al. 1997a) have been used to constrain the local value of the Hubble parameter, $H_0$.

1.4 The $\text{H} \text{I}$ deficiency parameter

The comparison of HI content among galaxies requires the set up of a suitable reference that can be used to measure deviations from normalcy. It is quite common to use the term HI content as a synonym for the HI mass.

Since $M_{\text{HI}}$ is derived from the square of the distance (Equation 1.1.4), to eliminate the uncertainty caused by distance error, alternative estimates for the HI content have been proposed in terms of distance-independent axis ratios: the hydrogen mass to luminosity ratio, $M_{\text{HI}}/L$, which is equivalent to the ratio of the corresponding fluxes, or the HI surface density, $M_{\text{HI}}/D^2$, where $D$ is the linear diameter (e.g., Davies & Lewis 1973; Chamaraux et al. 1980a, for early comparisons of HI content of Virgo Cluster galaxies with respect to samples of normal galaxies compiled by Roberts (1975) and Balkowski (1973)).

Haynes & Giovanelli (1984) were the first both to carry out an objective evaluation of the performance of different diagnostic tools for the HI content and to provide a rigorous operational definition of this quantity. In order to establish the reference values for the HI content, they selected a control sample of 288 galaxies listed in the Catalogue of Isolated Galaxies (CIG, Karachentseva 1973) and for which they detected HI emission. The sample, rigorously selected to include objects affected as little as possible by their surrounding environment, was used to define the standards of HI normalcy.
By analyzing the HI and optical properties of these galaxies, these authors demonstrated that, for a given Hubble type, the optical linear diameter is the most accurate diagnostic tool for the HI mass of galaxies.

According to that, the optical diameter of a galaxy $D_{\text{opt}}$, complemented with the information of its morphological type $T$, can be used to estimate its expected HI content. The ratio between that expectation and the actual observed value $M_{\text{HI}}^{\text{obs}}$ of a given galaxy’s HI content was labeled the deficiency factor or the HI deficiency parameter $DEF$. It was expressed as a logarithmic quantity, positive for HI-deficient galaxies:

$$DEF = \langle \log M_{\text{HI}}(D_{\text{opt}}, T) \rangle - \log M_{\text{HI}}^{\text{obs}} ,$$

(1.4.1)

A distance-independent approach for the definition of the deficiency parameter was also adopted on the basis of the almost constancy of the hybrid HI surface density, $\Sigma_{\text{HI}} = M_{\text{HI}}/D_{\text{opt}}^2$, they inferred from the control sample. The latter quantity turns out to show also a mild dependency on Hubble type leading, therefore, to the following distance-independent version of the HI-deficiency parameter:

$$DEF = \langle \log \Sigma_{\text{HI}}(T) \rangle - \log \Sigma_{\text{HI}}^{\text{obs}} ,$$

(1.4.2)

where the hybrid surface density $\Sigma_{\text{HI}}^{\text{obs}}$ has the advantage of being easily calculated from the ratio of the observables $F_{\text{HI}}$ [Jy km s$^{-1}$] and $a_{\text{opt}}$ ['], the integral HI flux and the apparent optical diameter, respectively. Both the definition of the $DEF$ parameter and its distance-independent version provided the basis for subsequent quantitative studies of the HI content of galaxies in various intergalactic environments (see Sections 1.5 and 1.7).

New expressions for the standards of HI content were later derived in an unbiased way by Solanes, Giovanelli, & Haynes (1996) using a larger, HI flux-limited sample of 532 galaxies from the Catalog of Galaxies and Clusters of Galaxies (CGCG, Zwicky et al. 1968) located in the lowest density environments of the Pisces-Perseus supercluster region. Taking into account
that the distribution of values of $M_{\text{H}I}$ for field galaxies approximately follow a Gaussian law with a scale of 0.25 independently of morphological type (Solanes et al. 1996), $DEF$ values of $1\sigma$, $2\sigma$ and $3\sigma$ correspond to a reduction of 44, 68 and 82% on the expected H I content, respectively.

1.5 Environmental effects on the H I of galaxies

We can find galaxies inhabiting a diversity of environments: galaxies can form pairs, triplets, compact or loose groups (gravitationally bound systems of tens to hundreds of galaxies), clusters (gravitationally-bound aggregates of up to thousands of galaxies) or, on the contrary, be relatively isolated. At larger scales, the spatial distribution of galaxies traces what is called the large scale structure of the Universe: voids (where the the lowest density of galaxies take place), filaments, walls, and superclusters (large aggregations of galaxy groups and clusters). As has been largely demonstrated, the properties of galaxies are related with their environment. The best example of this environmental dependence is perhaps the Morphology-Density relation: the distribution of morphological types among galaxies differs between the central parts of rich clusters and the field. In dense environments, the fraction of elliptical and lenticular galaxies (S0) is higher than the fraction of spiral and irregular galaxies, whereas the latter types dominate low-density environments. This relation, that was known as early as 1931 (Hubble & Humason 1931), was first quantified by Oemler (1974) and Dressler (1980), and has been successively confirmed by subsequent works (e.g., Whitmore et al. 1993; van der Wel 2008; Bamford et al. 2009).

The fact that Hubble type and local density are closely inter-related rises the question of whether this correlation is strictly a consequence of the circumstances in the era of galaxy formation (nature) or if it reflects processes that occurred late in the lives of these galaxies (nurture). In other words, were the morphological fractions at the epoch of galaxy formation the same as in the present, or, alternatively, galaxies can evolve along the Hubble se-
1.5. Environmental effects on the HI of galaxies

sequence by means of environmental processes? Particularly intriguing is the case of lenticular galaxies, which are abundant in clusters of galaxies but barely found in low density environments.

Apart from the different mix of morphological types, there have also been detected many differences in the properties of the spiral population itself: late-type galaxies inhabiting clusters differ systematically from their field counterparts regarding their neutral gas content, star formation activity, molecular gas content, metallicity, cold dust content, kinematic perturbations, and radio continuum synchrotron emission (e.g., see Boselli & Gavazzi 2006, for a detailed review on the effects of environment on late-type galaxies). In order to disentangle the origin of all these differences, (i.e., whereas it is nature or nurture) we need to understand the effects of environment on galaxy evolution, and HI has been proved to be a key tool in this research. Since, in general, HI extends to larger radial distances from galaxy centers than other baryonic components, it is more vulnerable to external perturbations and can be used as a tracer of relatively recent interactions. However, we must take into account that atomic, ionized and molecular gas components are distributed differently in a galaxy’s potential well, and they are vulnerable in different degrees to the action of external forces. Therefore, the use of HI to trace recent traumas, does not mean that whole interstellar gas component had the same fate.

Since the pioneering observations of Davies & Lewis (1973), all the investigations on the HI content of cluster galaxies generally agree with the general conclusion that, on average, cluster spirals tend to have less neutral gas than their field counterparts, and also in finding evidence for a correlation of the HI deficiency with clustercentric distance, with HI-poor disks typically situated close to the cluster cores and galaxies removed from those regions showing normal gas contents (e.g., Giovanelli & Haynes 1985; Solanes et al. 2001). Plenty of these studies focus their attention on nearby clusters such as Coma (Bravo-Alfaro et al. 2000) and Virgo (e.g., Haynes & Giovanelli 1986; Cayatte et al. 1990, 1994; Solanes et al. 2002; Chung et al. 2008, see

\[^{6}\text{Galaxies with a central bulge and a disk, but the latter apparently lacking spiral arms and substantial amounts of interstellar material.}\]
Section 1.7 for specific studies in the region). Detailed imaging studies carried in these nearby clusters show spirals with truncated gas disks, smaller than their undisturbed stellar disks, as well a trend of the extent of the HI disks with location in the cluster, being the most reduced those located near the core (e.g., Cayatte et al. 1994; Chung et al. 2009).

Figure 1.1 Top: HI-deficient fraction in bins of projected radius from the cluster center for the superposition of all the HI-deficient clusters but Virgo. Vertical error bars correspond to 1σ confidence Poisson intervals. The abscissae show medians and quartile values of the bins in radial distance. Bottom: Same as top panel for the measured HI deficiency. Displayed are the medians and quartiles of the binned number distributions in HI deficiency. Small dots show the radial variation of HI deficiency for individual galaxies, while the arrows identify nondetections plotted at their estimated lower limits. From Solanes et al. (2001).

Among the most recent statistical works, Solanes et al. (2001) investigated the HI content of the spirals in 18 nearby clusters using data extracted from the *Arecibo General Catalog* (AGC), an all-sky database maintained by M. P. Haynes and R. Giovanelli (Cornell University) that contains 21-cm and optical data for a large number of galaxies. They found that two-thirds of the clusters showed a large fraction of gas-poor objects in their inner parts. For these neutral gas deficient clusters, the radial extent of the regions with
1.5. Environmental effects on the H I of galaxies

significant gas ablation could reach up to 2 Abell radii\textsuperscript{7}, thus confirming previous results by Giovanelli & Haynes (1985). Within this zone, the proportion of H I-deficient spirals increases towards the cluster center (see Figure 1.1). They also found evidences that H I-deficient objects have more radial orbits than gas-rich ones, something already suggested by Dressler (1986).

Table 1.1 Some popular gas sweeping mechanisms

<table>
<thead>
<tr>
<th>Mechanisms</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Environment-independent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bregman (1978)</td>
</tr>
<tr>
<td>II. Environment-dependent</td>
<td></td>
</tr>
<tr>
<td>2. Galaxy-galaxy interactions</td>
<td></td>
</tr>
<tr>
<td>2.1 Direct collisions</td>
<td>Spitzer &amp; Baade (1951)</td>
</tr>
<tr>
<td>2.2 Tidal encounters</td>
<td>Toomre &amp; Toomre (1972)</td>
</tr>
<tr>
<td></td>
<td>Richstone (1976)</td>
</tr>
<tr>
<td>2.3 Mergers</td>
<td>Roos &amp; Norman (1979)</td>
</tr>
<tr>
<td>2.4 Harassment</td>
<td>Moore et al. (1996)</td>
</tr>
<tr>
<td>3. Galaxy-cluster medium interactions</td>
<td></td>
</tr>
<tr>
<td>3.5 Ram Pressure Stripping</td>
<td>Gunn &amp; Gott (1972)</td>
</tr>
<tr>
<td>3.6 Thermal evaporation</td>
<td>Cowie &amp; Songaila (1977)</td>
</tr>
<tr>
<td>3.7 Turbulent viscous stripping</td>
<td>Takeda et al. (1984)</td>
</tr>
<tr>
<td>3.8 Starvation or strangulation</td>
<td>Larson et al. (1980)</td>
</tr>
</tbody>
</table>

Note. — Modified version of Table III in Haynes (1988)

All these observational results indicate that, at least in the densest regions of cluster, gas-sweeping occurs. Since the 1970s, when the first models of tidal interactions between galaxies (Toomre & Toomre 1972) and interactions of the inter-stellar medium (ISM) of galaxies with the dense intra-cluster medium (ICM) were proposed (Gunn & Gott 1972), theoretical models and numerical simulations have provided a wealthy variety of possible scenarios. Some of the most popular gas-sweeping processes with accompanying references are listed in Table 1.1. The physics of most of these

\footnote{\textit{7} \( R_A = 1.72/z' \), where \( z \) is the cluster’s redshift, and is equivalent to \( 1.5h^{-1} \) Mpc radius at the distance of the cluster.}
mechanisms is detailed in Boselli & Gavazzi (2006). The reader is also referred to Haynes et al. (1984), who reviewed the effects of environment on HI content of galaxies, as well as to van Gorkom’s review on interactions of galaxies with the ICM (van Gorkom 2004).

In general, the results of most of the aforementioned HI studies, such as the radial pattern of HI deficiency (e.g., Giovanelli & Haynes 1985; Haynes & Giovanelli 1986; Solanes et al. 2001), the radial orbits of HI-deficient spirals (e.g., Dressler 1986; Solanes et al. 2001), synthesis imaging of truncated HI disks, and detected HI tails (e.g., Chung et al. 2007), among others, in combination with results at other wavelengths (e.g., Koopmann & Kenney 2004) point towards ram pressure stripping (RPS) of the interstellar gas of galaxies by the hot ICM as the most plausible mechanism acting on the spiral galaxy population in today’s clusters. However, although simple models of RPS are, in general, consistent with galaxy data for some galaxies, for other galaxies more than one mechanism appears to be necessary to explain the observations (e.g., Vollmer 2003), whereas in other cases the results suggest that the effects of the intracluster medium reach further than expected from simple ICM-ISM models (e.g. Crowl et al. 2006; Crowl & Kenney 2006; Bravo-Alfaro et al. 2009; Levy et al. 2007). Another mechanism that might affect cluster disks is galaxy harassment (Moore et al. 1996), proposed as a mechanisms to transform late-spirals into S0, and recently claimed to be responsible for the morphology of NGC 4254, as unveiled by the ALFALFA detection of this system (Haynes et al. 2007).

The measurements of HI content have been made extensive to other density regimes, showing that gas-depleted disks are not exclusive of cluster spirals: HI-deficiency has also been detected in Hickson Compact Groups (Verdes-Montenegro et al. 2001), in X-ray bright groups (e.g., Sengupta et al.

---

8 As first proposed by Gunn & Gott (1972), ram pressure can effectively remove the ISM of a galaxy if it overcomes the gravitational pressure anchoring the gas to the disk:

$$\rho_{IGM} V_{gal}^2 \geq 2\pi G \Sigma_{star} \Sigma_{gas},$$

where $\rho_{IGM}$ is the density of the IGM, $V_{gal}$ is the galaxy velocity inside the cluster, $\Sigma_{star}$ is the star surface density, and $\Sigma_{gas}$ is the gas surface density. Thus, the ISM could be removed from galaxies moving at 1000 km s$^{-1}$ through a hot ($\sim 10^7$–$10^8$ K) and dense ($\sim 10^{-3}$–$10^{-4}$ atoms cm$^{-3}$) intergalactic medium by means of a ram pressure mechanism.
as well as in loose groups (Sengupta & Balasubramanyam 2006). In these density regimes, in which the relative velocities of the galaxies are low, the tidal forces in close galaxy encounters can result in the removal of significant portion of a galaxy’s HI mass. Precisely, the efficiency of this mechanism in groups of galaxies has been used to favor the idea that galaxies suffer a pre-processing prior to entering the cluster environment (e.g., Fujita 2004). There are observations supporting this scenario, such as the compact group falling towards Abell 1367 (Cortese et al. 2006), or of the Eridanus group (Omar & Dwarakanath 2005).

As has been explained in this section, the stellar and gaseous disks of spirals are sensitive probes of their environment. Apart from observations, an increasing number of models and numerical simulations have has improved our knowledge of the different mechanisms that could affecting spirals in high density environments (e.g. Bekki et al. 2005; Kronberger et al. 2007, 2008; Vollmer 2009) However, it is still not clear to what extent the strong HI deficiency of a spiral is an unequivocal probe of cluster residency. Given the different scenarios in which the disk of a spiral galaxy can be moderately to severely affected, we still do not know precisely how far from the center of the cluster are effective the different galaxy-cluster interactions. For this reason, it is important to study late-type galaxies in intermediate and high density environments, specially in nearby groups and clusters, where good quality detections can be achieved (see Section 1.7).

1.6 The Virgo Cluster Region

The Virgo Cluster, which takes its name from the constellation where it lies in the sky, and with more than one thousand known members, is our nearest rich cluster of galaxies. It constitutes the nucleus of the Local Supercluster in whose outskirts our Milky Way Galaxy is situated.

We describe here the structural and dynamical complexity of the the Virgo Cluster Region, which shows a high degree of irregularity and subclus-
tering, in order to illustrate the difficulties arising when trying to determine the individual location of the galaxies located in this region.

1.6.1 Overall structure

Along the line of sight (LOS), the Virgo Cluster Region is located in the foreground of a small void\(^9\) that extends from \(\sim 3000 \text{ km s}^{-1}\) towards the Coma/A1367 Supercluster (Gregory & Thompson 1978; Oort 1983; Giovannelli & Haynes 1988), at an heliocentric velocity \(v_{\text{hel}} \sim 7000 \text{ km s}^{-1}\). This leaves a well-defined range of radial velocities, within \(-500\) and \(\sim 3000 \text{ km s}^{-1}\), beyond which there are certainly no cluster members (see Figure 1.2). Within this range of velocities, we find several subunits, more or less aligned with the central part of the cluster and with similar radial velocities, which make a non-trivial task to isolate the Virgo Cluster itself from the substructures around it.

Studies of the spatial distribution of galaxies in the Virgo region started with de Vaucouleurs (1961), who was the first to indicate the existence of several substructures, some of them in the near background of the cluster, designated as *clouds*. Since then, a large number of researchers have made progress on disentangling the complexity of this region of the sky (de Vaucouleurs & de Vaucouleurs 1973; Paturel 1979; Tully 1982; Ftaclas et al. 1984; Teerikorpi et al. 1992; Binggeli et al. 1993; Gavazzi et al. 1999, among others). However, there is a lack of agreement in the nomenclature and spatial delimitation of such structures, which makes, for instance, that the same name has been used in the literature to designate different subunits. Therefore, in order to avoid the confusion that could result from a description of all previous works, we will restrict ourselves to the study carried out by Binggeli et al. (1993), which is one of the most complete and comprehensive.

Binggeli et al. proposed a separation of the cluster region into several clouds and subunits (see Figure 1.2). That was done by identifying the possible cluster members from the available information on morphology and

\(^9\) Region where the density of galaxies is less than the mean density of the Universe
1.6. The Virgo Cluster Region

Figure 1.2 Distribution of galaxies listed in the last release of the VCC (Binggeli et al. 1993) extracted from GoldMine (Gavazzi et al. 2003). **Top:** Map of all galaxies listed in the VCC with known heliocentric velocity. The location of W, W' and M clouds is shown, as well as the location of the Southern Extension (SE). Five galaxies are identified with cross signs for reference, from North to South: M100, M86, M87, M60 and M49. **Bottom:** Histogram of heliocentric velocities of all galaxies in the VCC with known velocities. The bin size is 250 km s\(^{-1}\).
kinematics provided by an extension of the well-known Virgo Cluster Catalog (VCC, Binggeli et al. 1985). The Virgo Cluster proper extends approximately within a circumference in the sky of about 6° of radius centered in the equatorial coordinates \((\alpha, \delta) \sim (12^h 30^m, 13^\circ)\). Among the clouds apparently lying in the near background, the most prominent is the W cloud, at \((\alpha, \delta) \sim (12^h 15^m, 6^\circ)\), already identified by de Vaucouleurs (1961). The W cloud would be located at about twice the distance to the Virgo Cluster, and could be connected to the cluster itself through a second smaller cloud known as W' \(((\alpha, \delta) \sim (12^h 25^m, 7^\circ))\). Another small and well-defined cloud, named M cloud by Ftaclas et al. (1984), is located at the North-West of the principal cluster, at \((\alpha, \delta) \sim (12^h 15^m, 13^\circ)\), and would lie at a distance similar to the W cloud. Finally, Binggeli et al. defined the Southern Extension (SE), an enormous structure that stretches down to the Southern Hemisphere \((\delta < 5^\circ)\) and mainly consisting of late-type galaxies and that would be in the process of falling towards the cluster, according to Tully (1982), who was the first in identifying it.

In parallel with the identification of the clouds we have described above, there has also been achieved significant progress in the determination of the structure of the central part of the cluster. The clumpiness of this central part is indicative of the dynamical youth of the system (Binggeli et al. 1993). According to Schindler et al. (1999), the innermost part can be decomposed in two principal subclusters: the largest and most massive centered on the cD galaxy M87, which coincides with the maximum of the X-ray emission, and a second subcluster centered on M86 (see Figure 1.3). There is also a third subcluster towards the south centered in M49. Both M86 and M49 subclusters seem to be in a state of merging with the M87 subcluster. Schindler et al., using X-ray data from ROSAT All-Sky-Survey (Böhringer et al. 1994), which traces the hot intracluster gas, and optical data from the VCC (Binggeli et al. 1993), which traces the galaxies, estimated that the intracumular gas in the M87 subcluster is about three times the mass in the galaxies. In contrast, they estimated that the M49 subcluster has much more mass on galaxies than in gaseous state: its gas-mass fraction is only about 1%, which is an unusually low value for clusters (and even for groups of galaxies).
1.6. The Virgo Cluster Region

Figure 1.3 Comparison between the X-ray image and the optical appearance of the Virgo Cluster. The greyscale corresponds, in logarithmic scale, to the X-rays emission detected by the Rosat All Sky Survey (Böhringer et al 1994) in the 0.5-2.0 keV band, smoothed with a Gaussian filter. Contours show the linear grow of the density of members in the VCC. The lowest contour corresponds to $1.5 \times 10^{-3}$ galaxies per $\square''$. The image is $12.8^\circ \times 12.8^\circ$ of the sky. The darkest region in the X-rays coincides with the location of M87. From Schindler et al. (1999).

Further observations in X-rays have revealed the existence of a unusual hot region in the cluster located between M87 and M49 (Kikuchi et al. 2000). According to the authors, this heating of the gas could have been the result of the collision of the two subclusters. In fact, from the estimate that M49 is moving at a velocity of $\sim 1300$ km s$^{-1}$ towards M87 (Irwin & Sarazin 1996), they argue that the kinetic energy of the M49 subcluster itself would be enough to heat this region.
1.6.2  Distribution of morphological types

According to Binggeli et al. (1987), the distribution of Hubble types in the Virgo Cluster follows the general pattern of the well-known morphology-density relation (Dressler 1980): the fraction of early-type galaxies increases monotonically with local density, reaching the largest values in the cluster center. The differences in the spatial distribution of Hubble types in the Virgo Cluster were already suggested in the first works by de Vaucouleurs (1961), and confirmed later by Schindler et al. (1999), from whom we include Figure 1.4. This Figure shows the projected spatial distribution of the galaxies according to their morphological type. We can observe that the distribution of dwarf ellipticals (dEs) is similar to that of the whole sample of cluster members, although the latter could be explained by the fact that dEs represent the 75% of the total VCC catalog. As for the rest of early-type galaxies, ellipticals (E) and lenticulars (S0), we can see they are more concentrated in the central part of the cluster, following a distribution that resembles the X-ray emission (Figure 1.3). This is in contrast with the distribution of spiral (S) and irregular (Irr) galaxies: the distribution of late-types is more extended and does not correlate with the X-rays. Differences in the fractions of morphological types among the subclusters are also detected, being evident the richness in late-type galaxies of the M49 subcluster.

Apart from the significant differences in the spatial distributions of the morphological types we have just mentioned above, there are also significant kinematic differences (de Vaucouleurs 1961; Tammann 1972): gas-rich objects show a larger dispersion in velocity, $\sigma_v = 890$ km s$^{-1}$, than early types, $\sigma_v = 570$ km s$^{-1}$, according to the values reported by Binggeli et al. (1987). These kinematic and spatial differences among populations, together with the predictions from models of the falling of the SE towards the cluster, drove Tully & Shaya (1984) to conjecture that, essentially, all the spirals and irregulars in Virgo are in a process of incorporation into the cluster and, therefore, would have arrived more recently than the early-type objects.
1.6. The Virgo Cluster Region

Figure 1.4 Distribution of morphological types in the Virgo Cluster. The Figure shows the number density of VCC members smoothed with a Gaussian kernel of $\sigma \sim 24'$. (a) All members. (b) dE and dS0 galaxies. (c) E and S0 galaxies. (d) S and Irr galaxies. The contours correspond to a linear grow in the density of members of VCC, with increments of (a) $1.5 \times 10^{-3}$ galaxies/$\square'$, (b) $1.3 \times 10^{-3}$ galaxies/$\square'$, (c) $2.2 \times 10^{-4}$ galaxies/$\square'$, (d) $3.6 \times 10^{-4}$ galaxies/$\square'$, respectively. The first contour level of each map is equal to the corresponding increment. Five galaxies are identified with filled symbols for reference, from North to South: M100, M86, M87, M60 and M49. From Schindler et al. (1999).

Regarding the issue of the morphological classification, we must also highlight the possible errors in the Hubble type assignments of the spirals in the Virgo Cluster reported by Koopmann & Kenney (1998): in general, Virgo spirals systematically show a reduced global star formation rate (SFR) when compared to isolated galaxies\(^\text{10}\). That could pose a trouble to the determin-

\(^{10}\)In fact, the reduction in the SFR is severe in the outer parts of their disks, whereas
nation of their Hubble type, since spirals with reduced global SFR are often assigned early-type spiral classifications, irrespective of their central light concentrations. For instance, about half of the spirals in Virgo classified as of Sa type according to their low SFR, have light concentration indexes characteristic of isolated Sb-Sc galaxies. This suggests that the traditional observational parameters used in the Hubble classification might not be the most adequate to evaluate the morphology of galaxies inhabiting Virgo and, for extension, other clusters. In fact, it is possible that these misleading classifications could be contributing to the morphology-density relation (Dressler 1980, and Section 1.5 in this thesis).

1.7 H I studies of the Virgo Cluster

The proximity of the Virgo Cluster makes it our best laboratory to study ICM-ISM interactions. In the 1970’s decade, the first studies on the H I content of spiral galaxies detected that the disks in the Virgo Cluster show a lack of H I when compared to galaxies in the field of the same morphological type (e.g. Davies & Lewis 1973; Huchtmeier et al. 1976; Krumm & Salpeter 1979; Chamaraux et al. 1980b). This result has been confirmed, by studies based on an increasing number of detections of 21-cm line emitters in Virgo (e.g., Giovanardi et al. 1983; Haynes & Giovanelli 1986; Hoffman et al. 1998; Dressler 1986; Warmels 1986; Magri et al. 1988; Valluri & Jog 1991), and other clusters, as already pointed out in Section 1.5. The proximity of this cluster has also allowed detailed imaging of the galaxies in the region, revealing the existence of undisturbed stellar disks with truncated gaseous distributions (e.g. Cayatte et al. 1990), which is usually accompanied by truncated star formation in the outer stellar disks (although with normal or even enhanced star formation rates in the inner parts, Koopmann & Kenney 2004).

Among the very large amount of studies devoted to the study of the H I properties of galaxies in the Virgo region, we now restrict ourselves to those
works that have motivated the research carried out in the first part of this thesis, and that are summarized in the following paragraphs.

Figure 1.5 2D distribution of HI deficiency in the Virgo Cluster region. The map shows contours for the DEF parameter (Equation 1.4.2) from a sample of 287 spiral galaxies listed in the AGC (solid dots). The greyscale corresponds to the X-rays (0.4-2.4 keV) Rosat All Sky Survey (Böhringer et al 1994). Dotted lines enclose substructures in the region: W, W’ and the M cloud. Five galaxies are identified with cross signs for reference, from North to South: M100, M86, M87, M60 and M49. From Solanes et al. (2002).

As already mentioned in Section 1.5, Solanes et al. (2001) investigated the HI content of the spirals in 18 nearby clusters using data extracted from
the *Arecibo General Catalog* (AGC). In the case of Virgo, since a larger sample was available, they carried out a study of the bi-dimensional pattern of the H\textsc{i}-deficiency parameter (Section 1.4). The resulting map supported the hypothesis that H\textsc{i} deficiency of spirals is caused by the interaction of their disks with the intracumular hot gas around M87.

![Figure 1.6](image)

Figure 1.6 3D distribution of H\textsc{i} deficiency in the Virgo Cluster region. Fraction of spiral galaxies with \( DEF > 2\sigma \), \( F_{DEF} \), with respect to the 3D radial distance \( r \) to the center of the cluster. Vertical error bars correspond to 1\( \sigma \) confidence Poisson intervals. The abscissas show medians and interquartile ranges of the bins in distance determined from 16 galaxies, with the remaining one added to the last bin. The horizontal dotted line represents the expected value of \( F_{DEF} \) for field spiral galaxies if \( DEF \) follows a Gaussian distribution. The long-dashed curve illustrates the radial run of the medians of the binned number distributions in the measured \( DEF \). From Solanes et al. (2002).

The investigation initiated by Solanes et al., was expanded by subsequent works on the spatial distribution of the H\textsc{i} in the Virgo Cluster region. Thus, in 2002, Solanes et al. (2002) used the H\textsc{i} contents and Tully-Fisher radial distances of a sample of spirals in the Virgo region to study the two- and three-dimensional map of the H\textsc{i} distribution in the cluster and its outskirts. As a measure for the H\textsc{i} content, they used the distance-independent H\textsc{i}-deficiency parameter \( DEF \) defined in Equation 1.4.2. The bi-dimensional map of \( DEF \) they derived from 287 spirals listed in the AGC is shown in (Figure 1.5). This figure illustrates that the maximum \( DEF \) is located at
the center of Virgo, but also shows secondary HI deficiency peaks in other regions dominated by substructures apparently lying in the background of the cluster, such as the W, W' and M clouds (e.g., Binggeli et al. 1993). In order to disentangle the origin of such peaks, which could simply be the result of projection, Solanes et al. proceeded to determine the distribution of the HI deficiency along the LOS by using radial distances to the galaxies. Since Cepheid distances were only available for a short list of objects in this region (e.g., Freedman et al. 2001), they chose to deal with the catalogs of radial Tully-Fisher distances that had been published to date. After a process of homogenization of the TF data (details are given in Solanes et al. 2002), they ended up with a sample of 161 spirals with both HI content and radial distance information from which they derived the 3D map of the DEF parameter in rectangular coordinates (Figures 1.6-1.7). In agreement with the bi-dimensional map, the cube shows that the highest peaks of HI deficiency are consistent with the location of M87, and a global pattern of HI deficiency decreasing with increasing radial distance to the center, as found in other clusters (Solanes et al. 2001). There are, however, two secondary peaks of deficiency in the foreground and background, at \( \sim 5 \) and \( 10 \) Mpc from the center, respectively. The peak of HI deficiency in the foreground of M87 corresponds to a set of galaxies with radial distances within \( 5 - 10 \) Mpc from the Local Group, traveling at high speed with respect to the cluster, and with no apparent relation among them. On the other hand, most of the galaxies responsible for the peak of DEF in the background are strongly HI-deficient objects with homogenized radial distances between \( \sim 25-30 \) Mpc from the Local Group, most of them with systemic heliocentric velocities similar to the cluster mean (\( \sim 1100 \) km s\(^{-1}\)) and a projected location in the plane of the sky coincident with the area dominated by the M49 subcluster, and the W and W' clouds.

Ensuing efforts carried out by Sanchis et al. (2002), were aimed at finding an explanation for the presence of these strongly HI-deficient spirals apparently located outside the cluster body \(^{11}\). Using Tolman-Bondi models

\(^{11}\)Given that Virgo is the most explored cluster, an observational bias could explain that secondary peaks of HI deficiency in the outskirts had not been detected in more distant
Figure 1.7 Voxel projection of the three-dimensional distribution of H I deficiency in the Virgo Cluster region. The plot is in rectangular equatorial coordinates with distances given in Mpc. The x-y plane corresponds to Decl=0°, the x- and y-axes point to R.A. = 12h and 18h, respectively, and the z-axis points to the north. The central bright spot is associated with the cluster, with M87 being right at its center. The other two enhancements are peripheral regions of HI deficiency in the frontside and backside of the cluster. Our position is at the origin of the coordinate system. From Sanchis et al. (2002).

For spherical collapse, Sanchis et al. explored the possibility that the HI-deficient objects in the outskirts had passed through the cluster core and lost their gas by means of ICM-ISM interactions. They concluded that this scenario was not statistically unfeasible, although it would imply that after the interaction, these objects would have maintained unaltered both their HI deficiency and morphology for several Gyr.

To complement the latter work, Mamon et al. (2004) used analytical calculations and N-body simulations to estimate the maximum rebound radius clusters, where the HI observations would have been carried out closer to their centers than for Virgo.
1.7. HI studies of the Virgo Cluster

Figure 1.8 Velocity-distance diagrams of simulations and observations in the Virgo Cluster region. The observer is at (0,0) and sees a cone of angular radius 9°. Dots represent the velocity field traced by the particles in the cosmological N-body simulation by Mamon et al. (2004), incorporating Gaussian relative distance errors of $\sigma(\ln D) = 0.2$ for the dark matter particles. Superposed are galaxies of the Virgo cluster. Circles and triangles represent early-type and late-type galaxies, respectively. The size of the triangles informs about the HI deficiency of the spiral galaxies measured in units of the mean standard deviation for field objects (=0.24). HI-deficient galaxies deviating more than 3$\sigma$ from normalcy apparently outside the body of the cluster are shown as filled red triangles. Squares represent additional spiral galaxies with no HI-deficiency data. Open symbols indicate galaxies with uncertain distances (with distance errors greater than 5 Mpc or spirals with only one distance estimate). Dashed lines show the unperturbed Hubble flow with $H_0 = 66.7$ and 70 km s$^{-1}$ Mpc$^{-1}$, respectively (going upwards). From Sanchis et al. (2004).
of galaxies after they cross for the first time the core of a virialized system. They concluded that this rebound radius could be as much as 2.5 times the virial radius of the system, \( r_{\text{vir}} \), in agreement with Balogh et al. (2000) (see also the more recent work by Gill et al. 2005). For the Virgo cluster, with \( r_{\text{vir}} \sim 1.65 \) Mpc (according to X-ray observations), the maximum rebound radius would be \( \sim 4 \) Mpc, implying that, unless the radial distances of the HI-deficient outliers were grossly incorrect, they were galaxies that could not have lost their gas in a previous ICM-ISM interaction.

The role of distance errors was investigated in an accompanying paper by Sanchis et al. (2004). On it, the predicted velocity-distance diagrams of the N-body simulations of Mamon et al. (2004) were compared with that derived observationally from the sample of spirals gathered by Solanes et al. (2002). Since ellipticals trace better the distribution of the ICM hot gas (Schindler et al. 1999), kinematic information for 61 elliptical galaxies in the region was also included in the observational diagrams. While Sanchis et al. (Figure 1.8) concluded that a 20% relative error in radial distances would explain most of the elongation of the Virgo region detected along the LOS, they identified 13 strongly HI-deficient galaxies \( (\text{DEF} > 3\sigma) \) with a low probability of being within the maximum rebound radius. These authors considered different explanations for the gas deficiency and location of these HI-outliers, such as errors in the DEF parameter, the presence of massive nearby companions that could have caused the dearth of gas by means of tidal interactions, or grossly incorrect distances with errors larger than 30%. Among the options investigated, the latter one was pointed as the most feasible, even though HI-deficient objects did not show systematic deviations in their homogenized Tully-Fisher distances with respect to the whole sample.

The last step done so far in this direction is actually the starting point of this thesis project. It consists on carrying out dedicated VLA observations of the HI distribution of some of these 13 objects, seeking for clues that can help to elucidate the origin of their HI deficiency (Chapter 2). We note that 7 of our 13 HI outliers have already been observed by the VIVA Survey Chung et al. (2009), which has confirmed their strong HI deficiency and proposed
possible mechanisms that could explain it. In particular, they attribute the observed deficiency to a global cluster mechanism, such as RPS or turbulent viscous stripping, maybe in combination with tidal interactions (Chung et al. 2007).
Part I

VLA observations of gas-poor spirals in the Virgo Cluster region
We present here the VLA observations centered on the Virgo Cluster region spirals NGC 4307, NGC 4356, NGC 4411B, and NGC 4492, as well as of two other galaxies located in two of our fields and observed bands, the spiral NGC 4411A and the dwarf spiral VCC 740. This chapter, which presents the investigation already published in Toribio & Solanes (2009), begins with an introduction and the description of the targets selection (Section 2.2). The acquisition and reduction of the 21-cm line data, as well as the steps followed in the derivation of the HI synthesis results are discussed in Section 2.3. In Section 2.4, we analyze, case by case, the HI properties of all the galaxies showing 21 cm emission in the selected fields of view, while in Section 2.5 we discuss the applicability of the TF technique to the five large spirals that have been observed. Finally, the results and conclusions of this work are given in Section 2.6.
2.1 Introduction

The regions around rich clusters are the most obvious sites to evidence the transformation of galaxy disks driven by the surrounding intracluster medium (ICM). The increased density of both hot gas and galaxies, as well as the high relative velocities of the latter, set the scene for dramatic effects on their fragile interstellar medium (ISM).

In the local universe, the nearby Virgo Cluster region is an ideal place to quantify these nurturing effects, because its proximity makes it possible to probe the gaseous disks with higher sensitivity and resolution than in any other cluster. Another characteristic that makes this galaxy system very appealing for studies of galaxy evolution is its relative dynamical youth: Virgo has a central region with several substructures in the process of merging, surrounded by suburbia dominated by late-type galaxies that might fall into the cluster during the next Hubble time. As noted by Vollmer et al. (2001) and Solanes et al. (2001), environmental mechanisms such as ram-pressure stripping may see their effectiveness increased during the built-up of clusters.

There is a long list of studies covering a broad stretch of the electromagnetic spectrum that have investigated the impact of cluster residency on the late-type galaxy population. Plenty of them use data from the 21-cm emission line of the abundant, and easy to strip off, neutral hydrogen (HI) of the disks, as the most direct approach to measure the affectation of the ISM. These investigations generally agree in indicating that Virgo spirals, like those inhabiting other rich clusters, tend to have less neutral gas than their field counterparts, and also in finding evidence for a correlation of the HI deficiency with clustercentric distance, with HI-poor disks typically situated close to the cluster cores and galaxies removed from those regions showing normal gas contents (e.g. Haynes & Giovanelli 1986; Cayatte et al. 1994; Solanes et al. 2001; Chung et al. 2008). The lack of atomic gas, which usually affects the outer disks, is frequently accompanied by an even more severe truncation of the Hα emission and the corresponding quenching of star formation beyond that truncation radius (e.g. Koopmann & Kenney 2004;
Crowl & Kenney 2008). Recently, there are also evidences than it could be associated with H$_2$ reduction too (Fumagalli et al. 2009).

Virgo is also the first, and so far the only, cluster region for which the spatial distribution of the HI deficiency has been mapped (Solanes, Sanchis, Salvador-Solé, Giovanelli, & Haynes 2002, hereafter Sol02). By using homogenized Tully-Fisher (TF) distance moduli and 21 cm data from single-dish observations for 161 galaxies, these authors confirmed that the neutral gas deficiency in the Virgo Cluster decreases with increasing 3D barycentric distance. This study, however, also revealed the presence of an unexpectedly large fraction of strongly HI-deficient spirals with TF radial distances pointing to a location well outside the cluster body.

Ensuing studies based on both analytic infall models (Sanchis, Solanes, Salvador-Solé, Fouqué, & Manrique 2002) and N-body simulations (Mamon, Sanchis, Salvador-Solé, & Solanes 2004) investigated the possibility that some of the gas-poor spirals in Virgo’s suburbia had lost their gas content in a previous passage through the cluster core and were now lying near the apocenter of their orbits. Both works lead to the identification by Sanchis et al. (2004) of 13 extremely HI-poor spiral galaxies for which the lack of cold neutral gas could hardly be attributed to ISM-ICM stripping, unless their radial distances were affected by relative errors much larger than the typical uncertainty attributed to TF measurements. Other possibilities for the origin of these HI outliers, such as gas deficiency caused by gravitational interactions (tides or mergers) with companion galaxies, or errors in the HI deficiency estimates arising from morphological misclassifications, were also investigated and considered less probable.

With the aim of shedding more light on this matter, we initiated some time ago a program of dedicated observations of some of the outlying HI-deficient Virgo Cluster spirals found in Sanchis et al. (2004). In this chapter, we attempt to improve the results of the aforementioned study, which were based on the analysis of integral galaxy properties retrieved from public databases, by investigating the neutral gas distribution and kinematics, as well as the TF distances, for 4 of these objects by means of deep 21 cm
synthesis observations carried out with the Very Large Array (VLA) in its CS configuration\textsuperscript{1}.

### 2.2 Galaxy selection

![Sky distribution of our 4 main targets](image)

Figure 2.1 Sky distribution of our 4 main targets. The positions of five dominant Virgo cluster galaxies are marked by open diamonds (\textit{top to bottom}: M100, M86, M87, M60, and M49). Overlaid on the figure it is the greyscale map of the X-ray emission in the cluster from the \textit{ROSAT All-Sky Survey} in the 0.4–2.4 keV band (Böhringer et al. 1994). The location of W and W\textquotesingle clouds is also shown.

We have used the VLA to observe Virgo Cluster galaxies that are faint in the 21 cm line. Three of the targets, NGC 4307, NGC 4356, and NGC 4492,

\textsuperscript{1}The VLA is a facility of the National Radio Astronomy Observatory.
Table 2.1. Properties of the target galaxies

<table>
<thead>
<tr>
<th>Galaxy</th>
<th>R.A.</th>
<th>Dec.</th>
<th>Type</th>
<th>$v_{hel}$ (km s$^{-1}$)</th>
<th>$a \times b$ (')</th>
<th>$i$ (°)</th>
<th>$m_{B}^{tot,c}$ (mag)</th>
<th>DEF</th>
<th>$d$ (Mpc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main targets:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGC 4307</td>
<td>12$^{h}$22$^{m}$05$^{s}$.6</td>
<td>+09$^{\circ}$02$^{\prime}$37$^{\prime\prime}$</td>
<td>Sbc</td>
<td>1050</td>
<td>3.95 × 0.72</td>
<td>90.0</td>
<td>11.83</td>
<td>1.24</td>
<td>27.4$^{+3.5}_{-3.1}$</td>
</tr>
<tr>
<td>NGC 4356</td>
<td>12 24 14.4</td>
<td>+08 32 09</td>
<td>Sc</td>
<td>1137</td>
<td>3.2 × 0.5</td>
<td>71.4</td>
<td>13.24</td>
<td>1.35</td>
<td>29.4$^{+3.7}_{-3.3}$</td>
</tr>
<tr>
<td>NGC 4411B</td>
<td>12 26 47.2</td>
<td>+08 53 05</td>
<td>Sc</td>
<td>1271</td>
<td>3.45 × 3.45</td>
<td>26.7</td>
<td>12.81</td>
<td>0.63</td>
<td>27.9$^{+0.6}_{-0.7}$</td>
</tr>
<tr>
<td>NGC 4492</td>
<td>12 30 59.8</td>
<td>+08 04 41</td>
<td>Sa</td>
<td>1777</td>
<td>1.96 × 1.96</td>
<td>30.0</td>
<td>13.02</td>
<td>0.78</td>
<td>27.9$^{+8.7}_{-6.7}$</td>
</tr>
<tr>
<td>Secondary targets:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VCC 0740$^{a}$</td>
<td>12 24 39.9</td>
<td>+08 30 05</td>
<td>Sm</td>
<td>875</td>
<td>0.71 × 0.32</td>
<td>69.1</td>
<td>15.27</td>
<td>· · ·</td>
<td>· · ·</td>
</tr>
<tr>
<td>NGC 4411A$^{b}$</td>
<td>12 26 30.1</td>
<td>+08 52 18</td>
<td>Sc</td>
<td>1282</td>
<td>2.79 × 2.79</td>
<td>54.4</td>
<td>13.04</td>
<td>0.41</td>
<td>15.1</td>
</tr>
<tr>
<td>VCC 0933$^{b}$</td>
<td>12 26 44.3</td>
<td>+08 49 25</td>
<td>dE</td>
<td>1381$^{c}$</td>
<td>0.5 × 0.39</td>
<td>· · ·</td>
<td>17.00$^{d}$</td>
<td>· · ·</td>
<td>· · ·</td>
</tr>
<tr>
<td>VCC 0976$^{b}$</td>
<td>12 27 11.9</td>
<td>+08 50 24</td>
<td>dE</td>
<td>1259$^{c}$</td>
<td>0.28 × 0.18</td>
<td>· · ·</td>
<td>18.26$^{d}$</td>
<td>· · ·</td>
<td>· · ·</td>
</tr>
</tbody>
</table>

$^{a}$: in the same field of NGC 4356.

$^{b}$: in the same field of NGC 4411B.

$^{c}$: derived from optical measurements as listed in the Galaxy On-Line Database Milano Network (GOLDMine).

$^{d}$: photographic magnitude.
belong to the subset of 13 HI-outliers identified by Sanchis et al. (2004)\(^2\). These are galaxies with neutral gas deficiencies deviating by more than \(3\sigma_{\text{DEF}}\) from normalcy. A fourth pointing has been centered on NGC 4411B, another spiral with a less extreme HI deficiency. All these objects are among the most gas-poor spiral galaxies lying on the sky between the M49 subcluster and the W’/W cloud region (see Fig. 2.1) and have TF estimates of their radial distances suggestive of a possible background location far from the cluster core, provided one adopts the currently preferred Virgo mean distance of \(d_{\text{Virgo}} \sim 16 - 17\) Mpc, as suggested by both the measurements of \(H_0\) from HST observations of Cepheids and the spatial distribution of the early type galaxy population and X-ray gas (e.g. Gavazzi, Boselli, Scodeggio, Pierini, & Belsole 1999; Freedman, Madore, Gibson, Ferrarese, Kelson, Sakai, Mould, Kennicutt, Ford, Graham, Huchra, Hughes, Illingworth, Macri, & Stetson 2001; Sanchis, Mamon, Salvador-Solé, & Solanes 2004; Mei, Blakeslee, Côté, Tonry, West, Ferrarese, Jordán, Peng, Anthony, & Merritt 2007). Another characteristic these galaxies have in common is that their systemic velocities do not differ much from the mean cluster velocity. Galaxy properties are compiled in Table 2.1, including a preliminary estimate of their neutral gas deficiency using the following distance-independent calibrator

\[
\text{DEF}_{\text{HI}} = \langle \log \Sigma_{\text{HI}}(T) \rangle - \log \Sigma_{\text{HI}},
\]

which compares the logarithms of the expected and observed values of the hybrid HI surface density calculated from the ratio between the intrinsic integrated HI flux and the square of the apparent major optical diameter of a galaxy of morphological type \(T\). We have followed Solanes et al. and adopted for \(\langle \log \Sigma_{\text{HI}}(T) \rangle\) the values: 0.24 for Sa, Sab types; 0.38 for Sb; 0.40 for Sbc; 0.34 for Sc; and 0.42 for later types in units of Jy km s\(^{-1}\) per arcmin square. Thus, taking into account that the rms scatter in DEF\(_{\text{HI}}\) for field galaxies is \(\sigma_{\text{DEF}} = 0.24\), an object with \(\text{DEF}_{\text{HI}} > 3\sigma_{\text{DEF}}\) has less than 20% of the expected HI mass for a galaxy of its morphology. Description of columns:

\(^2\)Seven other members of this subset are among the targets of the VIVA (VLA Imaging of Virgo in Atomic gas) survey by J. Kenney, J. van Gorkom, and cols., in which the gas is imaged down to a column-density sensitivity of a few times \(10^{19}\) cm\(^{-2}\), similar to that of the present observations.
2.3. Data acquisition and processing

(1) NGC or VCC designations.
(2)–(3) J2000 position listed in the Lyon-Meudon Extragalactic Database (LEDA).
(4) Morphological type from GOLDMine.
(5) Heliocentric radial velocity obtained from H\textsc{i} profile measurements.
(6) Angular size of the semi-major and semi-minor optical axes determined at the surface brightness of 25 mag arcsec$^{-2}$ from GOLDMine.
(7) Inclination between the line-of-sight (LOS) and polar axis of the optical image listed in the LEDA.
(8) Total apparent $B$ magnitude corrected from inclination and extinction effects as listed in the LEDA.
(9) H\textsc{i} deficiency parameter calculated using the definition in equation (2.2.1) and the optical sizes and H\textsc{i} fluxes from conventional single-dish measurements attributed in the Arecibo General Catalog, a private compilation of 21 cm line measurements maintained by M. P. Haynes and R. Giovanelli at Cornell University.
(10) TF radial distance estimate from Solanes et al..

We have also included in Table 2.1 four more galaxies located in some of the fields of view of our main targets. These are: VCC 740, a small spiral in the vicinity of NGC 4356; the first component of the pair NGC 4411A/B, very close to its companion galaxy in both projected position and radial velocity (their centers are separated only by 4 ′and 11 km s$^{-1}$, respectively), but whose estimated TF radial distance of $\sim$ 16 Mpc (Solanes et al.) indicates that it is likely a Virgo Cluster member; and two dwarf ellipticals, VCC numbers 933 and 976, lying close to this pair on the sky.

2.3 Data acquisition and processing

2.3.1 Observations

The observations published here consist of data obtained at the VLA in its C configuration between July and October 2005. The H\textsc{i} spectral line was
observed with the correlator in 4IF mode using on-line Hanning smoothing.

The observational strategy was designed to achieve the best velocity resolution given the bandwidth needed for each galaxy. For observations where the primary target was an edge-on galaxy (NGC 4307 and NGC 4356), we chose to overlap partially the two IFs, each one with a bandwidth of 1.526 MHz and a spectral resolution of 24.4 kHz ($\sim 5.2$ km s$^{-1}$), whereas for fields with main targets oriented face-on (NGC 4411B and NGC 4492), the two IFs were centered on the heliocentric velocity of the target, the first with a bandwidth of 1.562 MHz and a spectral resolution of 24.4 kHz, and the second using a wider bandwidth of 3.125 MHz, but a lower frequency resolution of 97 kHz ($\sim 20.8$ km s$^{-1}$). The goal was to search for 21 cm line emission also from possible gaseous tidal tails, extraplanar gas or dwarf companions in the neighborhood of the target objects. The pointing of the field containing NGC 4492 was offset by 3′ toward M49, due to its strong, extended radio continuum emission, in order to avoid systematic effects due to the VLA beam squint as well as to pointing uncertainties in individual VLA antennas (Bhatnagar et al. 2008; Uson & Cotton 2008).

The July and August observations started in the afternoon. An incidence in the electric system of a substation on August 12th led to the partial loss of the observing time initially allocated for NGC 4492, which was compensated by 3 hours of diurnal observation on September 18th. The NGC 4411B field was also observed during the daytime on the 1st and 2nd of October (5 and 4 hrs, respectively). Solar interference was therefore significant only for the September and October observations.

Each galaxy was observed for about 6 hours with an overhead of $\sim 2$ hours for calibration which included two 10-minute scans on each day on the primary calibrator 3C286 = 1331+305 (J2000). The rest of the observing sequences consisted of 30-minute scans of the target fields interspersed with 10-minute observations of the corresponding secondary calibrator. For the data acquired between July and September, 3C273 = 1229+020 (J2000), with a flux of $\sim 32$ Jy, was used as a secondary phase and bandpass calibrator. This calibrator was too close to the Sun during our October observations.
of the NGC 4411B field. We therefore modified our strategy and observed J1254+116 as a secondary calibrator. A summary of the observing parameters is provided in Table 2.2.

2.3.2 Data calibration

The raw UV data were reduced using the Astronomical Image Processing System (AIPS) software package distributed by the National Radio Astronomy Observatory.\(^3\)

The observations of NGC 4307, NGC 4356, and NGC 4492 were calibrated in a similar way. First, we discarded corrupted data by inspecting a pseudo-continuum database obtained from the vector average of visibilities for channels 4 to 60 at each time stamp—the remaining channels in the low and high velocity ends of the bandpass were discarded from the beginning. The primary flux calibrator 3C286 was used then to determine an initial bandpass as well as zeroth-order amplitude and phase calibration. Next, pseudo-continuum images of 3C273 were obtained and subsequently used to calculate a secondary phase calibration. The same source was the basis to determine a secondary bandpass calibration. This calibration was applied to the data on the observing fields by linear interpolation of bandpass and global phases. Finally, phase self-calibration of the data on the galaxies was applied, and for NGC 4307 and NGC 4356, the two IFs were joined by means of the task UJOIN (channel-by-channel visibility averages for the overlapping channels; see Matthews & Uson 2008). Regarding the NGC 4492 field, we note that during the make-up observation scheduled on September 18th we acquired 3 hours of data that suffered from solar interference. This forced us to reject baselines shorter than 1 k\(\lambda\) for that day.

The reduction process just described was unsuitable to the observations of the NGC 4411B field. After discarding corrupted data, we calculated amplitude and phase calibrations from the observations of 3C286 and 1254+116. A solution for the shape of the bandpass obtained using 3C286 was applied

\(^3\)Unless otherwise stated, all the quoted routines in capital letters belong to this package.
Table 2.2. Summary of the observations

<table>
<thead>
<tr>
<th>Field</th>
<th>NGC 4307</th>
<th>NGC 4356</th>
<th>NGC 4492</th>
<th>NGC 4411B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Array configuration</td>
<td>CS</td>
<td>CS</td>
<td>CS</td>
<td>CS</td>
</tr>
<tr>
<td>Observing dates</td>
<td>09-07-2005</td>
<td>16-07-2005</td>
<td>12-08-2005 and 18-09-2005</td>
<td>02-10-2005</td>
</tr>
<tr>
<td>Total time on source (hr)</td>
<td>6.23</td>
<td>6.12</td>
<td>6.25</td>
<td>6.3</td>
</tr>
<tr>
<td>Phase center, $\alpha$ (J2000)</td>
<td>$12^h\ 22^m\ 05^s$</td>
<td>$12^h\ 24^m\ 14^s$</td>
<td>$12^h\ 30^m\ 47^s$</td>
<td>$12^h\ 26^m\ 47^s$</td>
</tr>
<tr>
<td>Flux calibrator (J2000)</td>
<td>1331+305</td>
<td>1331+305</td>
<td>1331+305</td>
<td>1331+305</td>
</tr>
<tr>
<td>Phase calibrator (J2000)</td>
<td>1229+020</td>
<td>1229+020</td>
<td>1229+020</td>
<td>1254+116</td>
</tr>
<tr>
<td>Bandwidth (1st/2nd IF pair, MHz)</td>
<td>1.562/1.562</td>
<td>1.562/1.562</td>
<td>1.562/3.125</td>
<td>1.562/3.125</td>
</tr>
<tr>
<td>Number of channels (1st/2nd IF pair)</td>
<td>63/63</td>
<td>63/63</td>
<td>63/31</td>
<td>63/31</td>
</tr>
<tr>
<td>Channel width$^a$(1st/2nd IF pair, kHz)</td>
<td>24.4/24.4</td>
<td>24.4/24.4</td>
<td>24.4/97.6</td>
<td>24.4/97.6</td>
</tr>
<tr>
<td>Central heliocentric velocity$^b$ (1st/2nd IF pair, km s$^{-1}$)</td>
<td>920/1180</td>
<td>940/1180</td>
<td>1777/1777</td>
<td>1309/1309</td>
</tr>
</tbody>
</table>

$^a$: after Hanning smoothing.

$^b$: optical definition.
to the whole data at the same time that the amplitude and phase solution, which was interpolated for NGC 4411B dataset by means of simple linear connection between phases. As for the NGC 4492 September data, we discarded spacings less than 1 kÅ to avoid the strong contamination by solar Radio Frequency Interference (RFI).

2.3.3 Continuum emission subtraction

Continuum emission in our fields was estimated by imaging the vector average of visibilities for line-free channels and subsequently subtracted from the datacube with UVSUB.

We note that comparison of our flux measurements of the brightest sources in each field with the corresponding flux values listed in the NRAO VLA Sky Survey (NVSS, Condon et al. 1998) shows good agreement, although, on average, our flux measurements are $\sim 5 \pm 2\%$ larger than the NVSS values. Given the low signal-to-noise ratio (S/N) of the spectral signal of our targets, this does not affect our total H I flux estimates. These discrepancies, however, were accounted for in the calculation of the corresponding uncertainties.

As mentioned on § 2.3.2, for targets close to the Sun, the images were obtained after excluding the baselines most affected by solar contamination. However, we could not eliminate solar RFI completely, which prompted us to make a second continuum subtraction. The routine UVLIN was used to fit and subtract a first order polynomial to the real and imaginary components of each visibility through the line-free channels. Following the recommendation by Cornwell et al. (1992), we applied UVLIN also to the data not affected by solar contamination in order to subtract any residual continuum emission. Subsequently, we obtained the channel images and examined the statistics of the image cubes to check for artifacts and, in particular, to verify that the distribution of noise in our data cube was Gaussian-like. Finally, we proceeded to concatenate the datasets for those objects with observations split in two different dates (NGC 4492 and NGC 4411B).
2.3.4 HI synthesis results

2.3.4.1 Channel maps

Image cubes were constructed for the NGC 4307, NGC 4356, and NGC 4411B fields using robustness parameters $\mathcal{R} = -1$ (closer to uniform weighting) and $\mathcal{R} = 0.7$ (closer to natural weighting). For all galaxies but NGC 4492, we present the results for $\mathcal{R} = 0.7$ as it provides the best compromise between the S/N and resolution when data have full UV-coverage (Briggs 1995).

The low S/N of the data on NGC 4492 forced us to smooth and taper the observations to improve our sensitivity. Channel maps were obtained from IF 2 (spectral resolution $\sim 20.8$ km s$^{-1}$) by using a robustness parameter $\mathcal{R} = 1$, which results in lower noise levels and a wider, but still acceptable, synthesized beam size. A Gaussian taper (with a 9 k$\lambda$-width at the 30\% level both in the U and V directions) additional weighting was applied on the visibilities to lower the contribution of the long-baseline datapoints.

In all cases, the channel images were CLEANed and the clean components restored with a Gaussian beam similar to the synthesized beam. The characteristics of the deconvolved images are summarized in Table 2.3.

2.3.4.2 Moments

The cleaned image cubes were used to calculate total HI flux images, as well as first and second velocity moments of the 21 cm emission.

We decided whether to keep or not a pixel in the integration by examining a spatially and frequency smoothed version of the datacubes. The spatial smoothing was done by convolving with a Gaussian kernel, whereas a Hanning smoothing was applied in velocity. We selected those pixels from the original datacubes that were above the $3\sigma$ level in the smoothed counterpart, except for NGC 4492, for which the $2\sigma$ level was used (otherwise, almost no signal was left from the galaxy). All flux with absolute value above $0.5\sigma$ was
Table 2.3. Characteristics of the deconvolved images

<table>
<thead>
<tr>
<th>Field</th>
<th>NGC 4307</th>
<th>NGC 4356</th>
<th>NGC 4492&lt;sup&gt;b&lt;/sup&gt;</th>
<th>NGC 4411B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robustness parameter</td>
<td>0.7</td>
<td>0.7</td>
<td>1</td>
<td>0.7</td>
</tr>
<tr>
<td>Gaussian taper width at 30% level (kλ)</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Synthesized beam position angle (°)</td>
<td>−51.1</td>
<td>−55.7</td>
<td>−55.5</td>
<td>−38.8</td>
</tr>
<tr>
<td>rms noise per channel in line-free channels (mJy beam&lt;sup&gt;−1&lt;/sup&gt;)</td>
<td>∼ 0.57&lt;sup&gt;a&lt;/sup&gt;</td>
<td>∼ 0.60&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.35</td>
<td>∼ 0.60</td>
</tr>
<tr>
<td>1σ limiting column density per channel (10&lt;sup&gt;19&lt;/sup&gt; cm&lt;sup&gt;−2&lt;/sup&gt;)</td>
<td>1.32</td>
<td>1.36</td>
<td>1.65</td>
<td>1.46</td>
</tr>
<tr>
<td>rms noise in total intensity map (mJy beam&lt;sup&gt;−1&lt;/sup&gt;km s&lt;sup&gt;−1&lt;/sup&gt;)</td>
<td>10.3</td>
<td>11.1</td>
<td>16.9</td>
<td>9.0</td>
</tr>
<tr>
<td>1σ limiting column density in total intensity map (10&lt;sup&gt;19&lt;/sup&gt; cm&lt;sup&gt;−2&lt;/sup&gt;km s&lt;sup&gt;−1&lt;/sup&gt;)</td>
<td>4.6</td>
<td>4.8</td>
<td>3.7</td>
<td>4.2</td>
</tr>
</tbody>
</table>

<sup>a</sup>: ∼ 0.42 for the cube obtained by averaging each two channels.

<sup>b</sup>: corresponding to images obtained from IF2 (see Table 2.2).
The heliocentric velocity of each panel is given in the upper right corner. The contours are drawn at \(-1.5 \text{ km s}^{-1}\) (dashed), \(-1.0 \text{ km s}^{-1}\) (solid), \(-0.5 \text{ km s}^{-1}\) (dashed), \(-0.0 \text{ km s}^{-1}\) (solid), and \(0.5 \text{ km s}^{-1}\) (dashed), respectively. Only channels containing line emission are plotted, bracketed by two noise channels. The maps were obtained by averaging the channels in pairs. Only channels containing line emission are plotted, bracketed by two noise channels.
Figure 2.2 Continued.
Figure 2.3 Same as in Figure 2.2 but for NGC 4356.
Figure 2.4 Same as in Figure 2.2 but for VCC 740. The contours are drawn at $(-2 \, \text{[absent]}, -1 \, \text{dashed}, 1, 2, 4) \times 3\sigma$, where $\sigma \sim 0.42 \, \text{mJy beam}^{-1}$. 

2. Gas-Poor Spirals in the Virgo Cluster Region

Figure 2.5 Same as in Figure 2.2 but for NGC 4411 B (left) and NGC 4411A (right). We show the individual channel maps from IF1 (resolution $\sim 5.2$ km s$^{-1}$). The contours are drawn at ($-2$ [absent], $-1$ (dashed), 1, 2, 4) $\times 3\sigma$, where $\sigma \sim 0.60$ mJy beam$^{-1}$ is the rms noise level.
2.3. Data acquisition and processing

Figure 2.5 Continued.
Figure 2.6 Same as in Figure 2.2 but for NGC 4492. We show the individual channel maps from IF2 (resolution \( \sim 20.8 \text{ km s}^{-1} \)). The contours are drawn at \((-\sqrt{2} \text{ [absent]}, -1 \text{ (dashed), } 1, \sqrt{2}, 2) \times 3\sigma\), where \(\sigma \sim 0.35 \text{ mJy beam}^{-1}\) is the rms noise level.
2.3. Data acquisition and processing

integrated along the velocity axis to obtain the total HI map, and intensity-weighted first and second order moments were computed. Finally, the HI intensity map was corrected for primary beam attenuation and scaled to the column density $N_{\text{HI}}$ of the gas, assuming optically thin HI. When estimating the errors in the latter map, we scaled the rms noise in the channel maps to the square-root of the mean number of adjacent channels that contributed to the total intensity image, and then added a 5% independent uncertainty arising from calibration and correction for primary beam attenuation.

The diameter of the HI disks was measured at a column density of $10^{20}$ atoms cm$^{-2}$ (no correction for beam smearing was applied). The reported uncertainties take into account variations in position angle of major axis as well as the correlation introduced by the synthesized beam (Table 2.4).

2.3.4.3 Global HI profiles

In order to determine if we have recovered all of the flux from single-dish measurements, simulated line profiles were derived. The latter were calculated by integrating over the spatial axes for each channel the primary-beam attenuation-corrected emission used in the moment calculation (§ 2.3.4.2).

We have measured $W_{20}$ and $W_{50}$, the profile width at the 20% and 50% of the peak intensity, respectively. For those cases in which a clear double-peaked profile is found (VCC 740, NGC 4411A and NGC 4411B), the peak fluxes on both sides were considered separately when calculating the linewidths. In the remaining cases we used the overall peak flux. The HI linewidths were subsequently corrected for broadening effects due to the finite spectral resolution of the instrument by following the considerations of Verheijen & Sancisi (2001). The adopted broadening corrections for the cubes with a spectral resolution of 5.2 km s$^{-1}$ were $\delta W_{20} = 0.86$ km s$^{-1}$ and $\delta W_{50} = 0.56$ km s$^{-1}$, whereas when the resolution was 20.8 km s$^{-1}$, we adopted $\delta W_{20} = 12.0$ km s$^{-1}$ and $\delta W_{50} = 7.88$ km s$^{-1}$.
### Table 2.4. Results from our VLA data and recent Arecibo measurements

<table>
<thead>
<tr>
<th>Galaxy</th>
<th>$W_{20}$ (km s$^{-1}$)</th>
<th>$W_{50}$ (km s$^{-1}$)</th>
<th>$f$ $S$ $dv$ (Jy km s$^{-1}$)</th>
<th>DEF$_{HI}$</th>
<th>$v_{sys}$ (km s$^{-1}$)</th>
<th>$D_{HI}$ (km)</th>
<th>$W_{50}$ (Jy km s$^{-1}$)</th>
<th>$f$ $S$ $dv$ (Jy km s$^{-1}$)</th>
<th>DEF$_{HI}$</th>
<th>$v_{sys}$ (km s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGC 4307</td>
<td>327±8</td>
<td>286±15</td>
<td>1.48±0.09</td>
<td>1.32</td>
<td>1048±4</td>
<td>1.83±0.16</td>
<td>314±41</td>
<td>1.21±0.10</td>
<td>1.41</td>
<td>1065±21</td>
</tr>
<tr>
<td>NGC 4356</td>
<td>204±7</td>
<td>154±27</td>
<td>0.44±0.04</td>
<td>1.71</td>
<td>1120±6</td>
<td>1.41±0.14</td>
<td>237±25</td>
<td>0.87±0.08</td>
<td>1.41</td>
<td>1092±12</td>
</tr>
<tr>
<td>NGC 4411B</td>
<td>93±1</td>
<td>77±1</td>
<td>18.91±0.95</td>
<td>0.13</td>
<td>1270±0</td>
<td>4.03±0.29</td>
<td>82±2</td>
<td>17.57±0.07</td>
<td>0.16</td>
<td>1272±1</td>
</tr>
<tr>
<td>NGC 4402$^a$</td>
<td>141±12</td>
<td>79±25</td>
<td>0.43±0.04</td>
<td>1.21</td>
<td>1768±7</td>
<td>1.38±0.17</td>
<td>182±9</td>
<td>0.59±0.07</td>
<td>1.07</td>
<td>1740±5</td>
</tr>
<tr>
<td>VCC 740$^b$</td>
<td>132±3</td>
<td>109±5</td>
<td>1.54±0.09</td>
<td>-0.21</td>
<td>875±1</td>
<td>1.47±0.14</td>
<td>97±5</td>
<td>1.39±0.06</td>
<td>-0.17</td>
<td>877±2</td>
</tr>
<tr>
<td>NGC 4411A$^c$</td>
<td>107±1</td>
<td>90±1</td>
<td>12.47±0.63</td>
<td>-0.03</td>
<td>1278±0</td>
<td>3.84±0.27</td>
<td>89±0</td>
<td>13.6±0.07</td>
<td>-0.07</td>
<td>1278±0</td>
</tr>
</tbody>
</table>

$^a$: Columns (2)–(6) show measurements from IF1 integrated profile (resolution $\sim$ 5.2 km s$^{-1}$), whereas (7) was measured on IF2 images.

$^b$: in the same field of NGC 4356.

$^c$: in the same field of NGC 4411B.

---

**Note.** — Column description:

1. Target name.
2–3. Width of the H$\text{I}$ line profile measured at 20\% and 50\% of the peak flux, respectively. Adopted broadening corrections correspond to a spectral resolution of 5.2 km s$^{-1}$ (see text).
4. Integrated total flux observed with the VLA.
5. Corresponding value of the H$\text{I}$ deficiency parameter recalculated from our VLA measurements.
6. Heliocentric systemic velocity.
7. H$\text{I}$ diameter measured at a column density of $10^{20}$ atoms cm$^{-2}$ (no correction for beam smearing is applied).
8. Width of the H$\text{I}$ line profile at 50\% of the peak flux measured by ALFALFA.
9. Integrated total flux measured by ALFALFA.
10. Corresponding value of the H$\text{I}$ deficiency parameter recalculated from ALFALFA measurements.
11. Heliocentric systemic velocity measured by ALFALFA.
The heliocentric systemic velocity, \( v_{\text{sys}} \), was derived as the average of velocities of channels at 20 and 50% of the peak flux of the profile, and the total HI flux was obtained by integrating the 21 cm line profiles along the velocity axis.

Since some of the profiles have low S/N, we decided to measure the kinematic parameters defined above by running Monte Carlo simulations for each one of the profiles. The values quoted in Table 2.4 correspond to the mean and 1-\( \sigma \) deviation of the distribution of measurements from one thousand random realizations of the profiles by taking into account the flux in each channel and its rms error. The latter was estimated from the rms noise in the channel maps and the correlation introduced by the synthesized beam. The quoted error in the total HI flux has been estimated by adding in quadrature the error estimates from the Monte Carlo technique and a 5\% uncertainty arising from the calibration and the correction for primary beam attenuation.

We find, in general, good agreement with single-dish observations (Table 2.4). Especially remarkable is the close match between the shapes, linewidths, and total flux densities of our VLA HI line profiles and their counterparts in the ongoing Arecibo Legacy Fast ALFA (ALFALFA) extragalactic HI survey (Giovanelli et al. 2005b), which recently released the results for the strip of the Virgo Cluster region where our targets are located (Kent et al. 2008).

2.3.4.4 Position-velocity diagrams and rotation curves

Position-velocity (PV) diagrams along the major axis of each target were also obtained by taking slices of the data cubes through the optical centers and estimating their position angle from the moment maps or from the rotation curves in those cases where this was possible.

Major-axis velocities as a function of angular radius were exclusively derived for the two components of the pair NGC 4411A/B and the dwarf spiral VCC 740, since these were the only targets detected with a high S/N (see
next section). To infer the rotation curves we have fitted both a Brandt model and the standard iterative tilted-ring algorithm (Begeman 1989).

2.4 Case by case analysis of the H I synthesis data

Table 2.4 summarizes the results inferred from our analysis of the VLA observations for all the galaxies detected. The corresponding images are shown in Figures 2.2–2.10.

We now comment on the properties of the target galaxies from our VLA observations.
2.4. Case by case analysis of the HI synthesis data

Figure 2.7 HI synthesis results for NGC 4307. (a) HI column density contours overlaid on the optical image of the galaxy obtained from the DSS. The contours levels are $(-2 \text{ absent}, -1 \text{ (dashed)}, 1, 2, 4, 8) \times 31 \text{ mJy beam}^{-1} \text{ km s}^{-1}$. The size of the synthesized beam is plotted in the lower left corner. (b) Linear gray-scale map of the HI column density. The grayscale range is 31–276 mJy beam$^{-1}$ km s$^{-1}$. (c) Velocity field of the HI with contours and grayscale ranging from 899 km s$^{-1}$ (light gray) to 1261 km s$^{-1}$ (dark gray) in increments of 20.8 km s$^{-1}$. The adopted heliocentric systemic velocity is indicated by a thick black contour. (d) HI position-velocity plot along the major axis. The data have been averaged in velocity (joining the channels in pairs) and spatially along a 8$''$-wide strip through the optical center of the galaxy with a position angle $\Gamma = 25^\circ$ (east of north). Contour levels are at $(-\sqrt{2} \text{ absent}, -1 \text{ (dashed)}, 1, \sqrt{2}, 2) \times 1.44 \text{ mJy beam}^{-1}$. The gray scale is linear in the range 0.96–3.11 mJy beam$^{-1}$. A pair of vertical arrows are drawn to indicate the extent of the optical disk. (e) Integrated HI line profile from our VLA observations (solid line) and ALFALFA (dotted line). The vertical arrow indicates the systemic velocity from the VLA profile. (An extended figure set for all fields of view is available in the online journal).
Figure 2.7 Continued.
2.4. Case by case analysis of the HI synthesis data

Figure 2.8 Same as in Figure 2.7 but for NGC 4356 (left) and VCC 740 (right). (a) The contours levels are \((-\sqrt{2} \text{ [absent]}, -1 \text{ (dashed)}, 1, \sqrt{2}, 2, 2\sqrt{2}, 4) \times 33 \text{ mJy beam}^{-1}\text{km s}^{-1}\) for NGC 4356 and \((-\sqrt{2} \text{ [absent]}, -1 \text{ (dashed)}, 1, \sqrt{2}, 2, 2\sqrt{2}, 4, 4\sqrt{2}) \times 33 \text{ mJy beam}^{-1}\text{km s}^{-1}\) for VCC 740. (b) The grayscale range is 33–242 mJy beam\(^{-1}\)km s\(^{-1}\). (c) Contours and grayscale range from 1039 km s\(^{-1}\) (light gray) to 1227 km s\(^{-1}\) (dark gray) for NGC 4356 and from 817 km s\(^{-1}\) (light gray) to 929 km s\(^{-1}\) (dark gray) for VCC 740 in increments of 20.8 km s\(^{-1}\) and 10.4 km s\(^{-1}\), respectively. (d) Spatial averages of data are along a 8\(''\)-wide strips through the optical centers of the galaxies with position angles of, respectively, \(\Gamma = 50^\circ\) and \(\Gamma = 115^\circ\) (east of north). Contour levels are at \((-\sqrt{2} \text{ [absent]}, -1 \text{ (dashed)}, 1, \sqrt{2}, 2) \times 1.38 \text{ mJy beam}^{-1}\) for NGC 4356 and at \((-\sqrt{2} \text{ [absent]}, -1 \text{ (dashed)}, 1, \sqrt{2}, 2, 2\sqrt{2}) \times 1.38 \text{ mJy beam}^{-1}\) for VCC 740. The gray scales range is 0.92–5.18 mJy beam\(^{-1}\). (f) Radial run of the rotation curve for VCC 740 in the form \(V_{\text{rot}}(\theta)\sin i\). The data for the approaching (open circles) and receding (filled circles) sides are presented superposed, together with a thick solid curve following the average value. No correction for beam smearing is applied.
Figure 2.8 Continued.
2.4. Case by case analysis of the H I synthesis data

Figure 2.8 Continued.
Figure 2.9 Same as in Figure 2.7 but for NGC 4411B (left) and NGC 4411A (right). (a) The contours levels are \(-\sqrt{2}\) [absent], \(-1\) (dashed), \(1, \sqrt{2}, 2, 2\sqrt{2}, 4, 4\sqrt{2}, 8\) × 27 mJy beam\(^{-1}\) km s\(^{-1}\). (b) The grayscale range is 27–279 mJy beam\(^{-1}\) km s\(^{-1}\). (c) Contours and grayscales range from 1226 km s\(^{-1}\) (light gray) to 1332 km s\(^{-1}\) (dark gray) in increments of 10.4 km s\(^{-1}\). (e) The data is averaged spatially along a 4′′-wide strips through the optical centers of the galaxies with position angles of \(\Gamma = 233^\circ\) and \(\Gamma = 217^\circ\) (east of north) for NGC 4411B and NGC 4411A, respectively. Contour levels are at \(-\sqrt{2}\) [absent], \(-1\) (dashed), \(1, \sqrt{2}, 2, 2\sqrt{2}, 4\) × 1.80 mJy beam\(^{-1}\). The grayscale range is 1.20–9.62 mJy beam\(^{-1}\). (f) Radial run of the rotation curves in the form \(V_\text{rot}(\theta)\sin i\). The data for the approaching (open circles) and receding (filled circles) sides are presented superposed, together with a thick solid curve following the average value. No correction for beam smearing is applied.
Figure 2.9 Continued.
Figure 2.9 Continued.
2.4. Case by case analysis of the H I synthesis data

Figure 2.10 Same as in Figure 2.7 but for NGC 4492. (a) The contours levels are \((-\sqrt{2} \text{ [absent]}, -1 \text{ (dashed)}, 1, \sqrt{2}, 2, 2\sqrt{2}) \times 49 \text{ mJy beam}^{-1} \text{ km s}^{-1}\). (b) The grayscale range is 49–149 mJy beam\(^{-1}\) km s\(^{-1}\). (c) Contours and grayscales range from 1726 km s\(^{-1}\) (light gray) to 1809 km s\(^{-1}\) (dark gray) in increments of 10.4 km s\(^{-1}\). (d) The spatial average of data is along a 11"-wide strip through the optical center of the galaxy with a position angle of \(\Gamma = 205^\circ\) (east of north). Contour levels are at \((-\sqrt{2} \text{ [absent]}, -1 \text{ (dashed)}, 1, \sqrt{2}) \times 1.05 \text{ mJy beam}^{-1}\). The gray scale range is 0.7–1.91 mJy beam\(^{-1}\). Panels (a)-(d) were obtained from IF2 (resolution \(\sim 20.8 \text{ km s}^{-1}\)), whereas (e) shows the integrated H I line profile from IF1 (resolution \(\sim 5.2 \text{ km s}^{-1}\)).
2.4.1 NGC 4307

The H I gas disk of NGC 4307 has very small dimensions and is found only deep within the optical disk, in agreement with the strong gas-deficiency estimated for this galaxy. The deficiency parameter $\text{DEF}_{\text{HI}}$ (eq. [2.2.1]) measured from our VLA flux (see Table 2.4) shows a pretty good consistency with the value of 1.41 one can infer from the observed H I flux provided by (Kent et al. 2008). We remark that this new value is also in agreement with the extremely H I-deficient status reported in previous works (e.g. Solanes et al.; Gavazzi et al. 2005).

The 0th-order moment map shows that the maximum of the 21 cm emission is displaced SE with respect to the optical center, in a direction nearly perpendicular to the major axis, with an excess of emission at the approaching side. This asymmetry shows up also in the global H I profile, which has a peaked appearance indicative of centrally concentrated gas. The estimated offset of $\sim 7 \arcsec$ is less than half of both the synthesized beam size and the mismatch expected to arise from edge-on ram-pressure stripping in hydrodynamic simulations (Kronberger et al. 2008).

There are no other pieces of evidence susceptible to being interpreted as environmental effects down to the sensitivity limit of the measurements.
Moreover, the galaxy is too inclined to allow one to disentangle whether the observed asymmetries can be ascribed to noncircular motions associated with a spiral arm.

The PV diagram along the major axis is still rising at the edge of the measured HI distribution, as if the gas was rotating almost as a pure solid body. Our VLA data is compatible to a large extent with the Hα rotation curve derived by Gavazzi et al. (1999), which shows a hint of a turn over on both sides of the galaxy. We discuss in Section 2.5 the risks of using extremely HI-deficient galaxies like this one in a TF analysis.

2.4.2 NGC 4356 and VCC 740

The properties of the neutral hydrogen in NGC 4356 resemble those of NGC 4307. This is a nearly edge-on galaxy that like the former one, and even more strongly so, has a very small gas distribution compared with its optical dimensions. The PV diagram along the major axis also rises steeply and shows no signs of turning over, as in the Hα rotation curve measured by Gavazzi et al. (1999). There is similarly a misalignment between the distribution and motion of the atomic gas and the optical disk. As in the former case, this offset is less than half the synthesized beam size and, hence, relatively small and not necessarily indicative of ram-pressure pushing.

For this galaxy, our VLA linewidth and flux values are somewhat smaller than the most recent estimates by Kent et al. (2008). Given the low S/N of this galaxy, the observed difference in total fluxes, which is only of tenths of a mJy km s$^{-1}$, does not necessarily imply that some flux has been lost due to the missing short baselines in our synthesis aperture images. Instead, the discrepancies can be the result of low-baseline ripples that seem to affect the single-dish profile.

VCC 740, another highly inclined galaxy detected in this same pointing, has, in contrast, quite a strong S/N and a total HI flux that is about the same that in NGC 4307. Its H1 map suggests that the neutral gas has a
sharp cutoff at the optical radius of the disk on the approaching side, while
the other side has a more gradual falloff and is somewhat more extended. In
spite of giving the impression that part of the gas might be missing, both ours
and the ALFALFA flux measurements actually result in a negative deficiency
parameter $\text{DEF}_{\text{HI}}$ of $\sim -0.2$, indicative of a perfectly normal HI content.

The gas velocity field of VCC 740, on the other hand, exhibits some
warping, as well as inner contours parallel to the minor axis, indicating that
$V_{\text{rot}}(R) \propto R$, as the PV diagram shows. Further out on the SE side, the con-
tours show the classical $V$-shape, suggesting that at the sensitivity limit of
the measurements the rotation speed is just about to become nearly constant.
In the rotation curve model fits, the warp of the velocity field increases the
position angle on the external part of the disk. Both Brandt and tilted-ring
models yield an estimated inclination of $\sim 70^\circ$ for the internal disk region
that drops to $\sim 45^\circ$ on the outside. The fact that the velocity field of this
galaxy shows a clear rotation pattern reinforces its morphological classifica-
tion as a dwarf barred disk instead of the IB type assigned in the LEDA.
The angular resolution of our VLA measurements, however, is insufficient to
observe the effects of the bar on the gas velocities. No signs of interaction
are found between VCC 740 and NGC 4356.

2.4.3 NGC 4411A/B, VCC 933 and VCC 976

Contrarily to what has happened with NGC 4307 and NGC 4356, the two
low surface brightness spirals NGC 4411A and B have produced integrated
HI fluxes that are inconsistent with the values quoted in Solanes et al.,
which were estimated in earliest mapping attempts done with the Arecibo
antenna (Haynes 1981). Discrepancies with this and other old single-dish
measurements (e.g. Hoffman et al. 1989) are ascribed, however, solely to the
amplitudes, as the shapes of the global HI profiles look very similar.

Our new data, reinforced by the newest single-dish measurements done at
Arecibo by the ALFALFA team (Kent et al. 2008), imply that NGC 4411A
and NGC 4411B loose their initial status of objects with moderate and strong
H I deficiency, respectively (see Table 2.1), to be both reclassified as galaxies with quite a normal gas abundance, in accordance with the visual impression obtained from the H I contours overlaid on the Digitized Sky Survey (DSS) image.

These maps show that these two galaxies have H I distributions extending beyond their optical disks, which for NGC 4411A reach up to nearly twice the optical radius, except on the NE side where the H I contours appear compressed. In this latter galaxy the H I is concentrated in a ring with two important regions of emission that emanate perpendicularly from the ends of the bar. The 0th-order map of NGC 4411B shows an even wider major ring-like structure with a noticeable excess of emission at its northern half near the outer edge. We do not find important displacements from the optical disks, the most remarkable feature being the presence of a depression in the center of both galaxies, as found in other LSB galaxies (de Blok, McGaugh, & van der Hulst 1996).

We fail in detecting neither gas bridges nor significant intergalactic H I signals between these two targets. Given that both galaxies show normal disk emission with only mild alterations of the symmetry, comparable to those seen in more isolated objects (Kornreich et al. 2000), we feel compelled to classify the system NGC 4411A/B as a visual pair (see also § 2.5).

Regarding the kinematics of the neutral gas, we note that the PV diagram of NGC 4411B shows a steep rise well within the stellar disk (of small amplitude given its near face-on orientation) followed by a sharp bend towards a flat part, —with indications of a modest decline on the receding side at the largest radii where the H I is detected—, consistent with the shape of the radial velocity contours and with the well behaved double-horned global profile. In NGC 4411A, the turnover in the rotation velocity is not complete and the 21 cm line profile is affected by a larger asymmetry. The observed behavior of the PV diagrams is typical of objects with a compact distribution of their luminous matter.

For these two galaxies we have gone a step further and inferred also
major-axis rotation velocities with the aim of estimating the orientations of the gaseous disks. The apparent inclination inferred from the tilted-ring technique for NGC 4411A is $29^{+5.2}_{-3.7}$ degrees, while for NGC 4411B we get $i = 26^{+4.4}_{-4.7}$ degrees. Brandt curve fits to the whole velocity fields result in similar inclinations of $\sim 27^\circ$ and $\sim 28^\circ$, respectively. Yet it shouldn’t be forgotten that the bumps in the HI line data associated with spiral arms and, especially, the near face-on orientation of these disks, make rotation curve model fits to the velocity fields uncertain and, caveat lector, unable of giving a precise inclination angle. In the next section, we discuss the problems arising from the derivation of this parameter in low-inclination galaxies and provide alternative estimates based on optical images.

Our VLA observations have not detected 21 cm line emission coming from any of the two dwarf elliptical galaxies VCC 933 and VCC 976 also present in this field.

### 2.4.4 NGC 4492

In spite of the strong image-degrading effects of the Sun for the second half of the observations of NGC 4492, we succeed in detecting its HI signal at a quality level comparable with the previous single-dish observations carried out by Haynes & Giovanelli (1986) and Hoffman et al. (1989)—and not too different from the one achieved by the ALFALFA measurements, which have a S/N of 4.6.

In this galaxy, the detected neutral hydrogen is located within the optical radius and shows an important elongation in the SE-NW direction, almost perpendicularly to the galaxy major axis. The HI distribution is strongly asymmetric, with the peak of the HI emission shifted some 30" to the East from the optical center of the galaxy. This gives rise to a synthesized line profile with decreasing flux toward the approaching side. By integrating the latter, we find a total HI flux about a 30% smaller than the ALFALFA’s value (see Table 2.4). Again, one may wonder whether possible flux losses in our estimation arising from too a strict rejection of the short baselines could
explain this difference. In this respect, we note that during the application of
the Gaussian UV-taper weighting to obtain the final datacubes some testing
done varying the width of the tapering function showed that the resulting
total HI flux was not significantly affected. Therefore, given the low S/R
of the detections, the discrepancies in total flux and linewidths with respect
to singledish measurements could have been originated by low-level baseline
ripples that seem to affect the ALFALFA profile. With this caveat in mind,
our VLA measurements assign a new HI deficiency parameter of 1.21 to
NGC 4492, which indicates that the gas content for this galaxy could be less
than 6% of the expectation value.

The asymmetries on the spatial distribution of the gas for this galaxy
are reproduced in the HI dynamics. Thus, the velocity map of NGC 4492
reveals a possible displacement of the dynamical center from the light cen-
ter (consistent with that observed in the 0th-order map), while on the PV
diagram along the major axis most of the emission is found coming from the
receding side. The low S/N of the datacubes, the small size of the HI disk
with respect to the beam size, as well as a not too favorable orientation of
this galaxy, prevent any attempt to fit a rotation curve model to the velocity
data.

2.5 TF-based distance estimates to our galaxies

The most striking aspect of the Virgo’s galaxies selected in the present study
is not their high deficiency of neutral gas, but the possibility that this has
been attained outside the cluster environment. While in the outskirts of
clusters gas stripping can happen by galaxy-galaxy interactions in infalling
groups or through collisions with lumps of intergalactic gas (see, for instance,
Crowl & Kenney 2006, 2008), strong HI deficiencies in the periphery of
clusters are expected to be an exception to the rule.

The classification by Sanchis et al. (2004) of a spiral galaxy as an HI
outlier relied on its TF distance. The latter was inferred from the disk’s
maximum rotation speed $V_{\text{rot}}$, —which is expected to be a proxy to the total mass of the galaxy and, therefore, also to its intrinsic luminosity—, measured via the width of the HI spectral line. It is then of fundamental importance to assess the feasibility of this technique in galaxies like ours which have gaseous disks deeply altered.

Figure 2.12 TF distances to the disk galaxies NGC 4411A (a) and NGC 4411B (b). The open diamonds indicate the radial distance to the galaxies published in several TF studies used in Solanes et al., quoted in the horizontal axis (see caption of Figure 2.13). Vertical bars show the ranges of distances allowed by these studies when our VLA HI linewidth estimates and the most extreme values of the inclination ever inferred for each object are used. Double pointed arrows encompass the ranges of distances corresponding to the values of inclinations and errors derived from our HI data. Distance estimates resulting from the structural decomposition of the SDSS r-band images are indicated by asterisks.

From Guhathakurta et al. (1988) to Cortés et al. (2008) the literature is full of TF studies highlighting the risks of using the gas kinematics in the determination of distances to HI-deficient galaxies. The most straight reasoning being that truncated gas disk measurements could underestimate the rotation velocity of a galaxy and, therefore, its mass, biasing low the derived radial distance. However, depending on the interaction mechanism and its geometry, the HI that does not get dislocated may also lost temporarily its equilibrium within the global galactic potential —externally induced disturbances on the kinematics of disks are erased in about 1 Gyr (e.g. Dale et al.
2.5. TF-based distance estimates to our galaxies

2001), a time during which two interacting galaxies can move hundreds of kiloparsecs apart. These effects, as well as induced noncircular or nonplanar gas motions, which may even lead to an overestimation of the true $V_{\text{rot}}$ (Kronberger et al. 2007), possible changes in the observed luminosity resulting from alterations in the star formation rate or, simply, the fact that even for undisturbed galaxies in many cases there is evidence of noncircular motions in their central regions (for instance, due to bars; Valenzuela, Rhee, Klypin, Governato, Stinson, Quinn, & Wadsley 2007), could make the application of the TF technique in strong H I-deficient galaxies totally inefficient.

All this drives us to regard critically the distance estimates of the three galaxies in our sample that exhibit severely truncated gas disks: NGC 4307, NGC 4356, and NGC 4492. Certainly, we have not found evidence of a recent gravitational encounter in our 21 cm line imaging data in the form of gaseous tails and bridges for any of them. Nor the closeness of the systemic velocity of the first two galaxies to the mean cluster velocity supports a recent ram-pressure event inside the core. Yet the fact we do not detect a flat part in their rotation curves, as well as the irregularities in the spatial distribution and kinematics of the HI evidenced in the VLA maps of these objects do not allow us to state confidently that the dynamical equilibrium of the neutral gas has been fully restored after the removal event. Neither can we assert that their luminosities have not been affected. Therefore, we believe that it is not legitimate to use the TF technique to derive the radial distance to any of these three galaxies and that, consequently, their published TF distance measurements might well be largely in error.

Compared to the previous objects, NGC 4411B and its close neighbor NGC 4411A exhibit regular and symmetric gas velocity fields, with flat extended outer parts that appear to satisfy the basic tenet that underlies a TF study. In this case, however, the attempts of estimating the radial distance to these two galaxies are thwarted for an unfavorable viewing angle. At low apparent inclinations ($i \lesssim 40-45^\circ$), estimates of the orientation of disks become more uncertain, led to deprojected quantities with divergent errors as a face-on orientation is approached, and are skewed towards larger values by nonaxysymmetric features in the images (see Andersen & Bershady
As a result, the total fractional error in the radial distance for low-inclination galaxies largely exceeds 15%, a value usually adopted as representative of the typical uncertainty in the distance arising from good-quality TF data.

Figure 2.13 B-band TF templates considered in Sol02. The acronyms correspond to Yasuda, Fukugita, & Okamura (1997) (YFO97), Kraan-Korteweg, Cameron & Tammann (1988) (KCT88), Fouqué et al. (1990) (Fou90), Federspiel, Tammann, & Sandage (1998) (FTS98), and Ekholm et al. (2000) (Ekh00). The relation resulting from the homogenization of the eight TF catalogs used by Sol02 is also included.

This situation is illustrated in Figure 2.12, where we show the radial distances to NGC 4411A and B reported in some of the TF catalogs used by Sol02. Our synthesized HI line profiles have been used to estimate the intrinsic rotational velocities, which we have calculated by exactly following the same prescriptions adopted in the referenced studies that in some cases apply non-null turbulent motion corrections. The uncertainty resulting from inclination measures is shown by a vertical bar whose extent is set by the most extreme values of this angle ever assigned to our galaxies in the literature (which range from \(\sim 20^\circ\) to \(\sim 55^\circ\)). Open diamonds in the plots
indicate radial distances published in the cited references. It is obvious from this figure that the inability to infer accurate inclinations prevents us from establishing the location of both galaxies along the LOS within the entire Virgo Cluster region.

We have also depicted in Figure 2.12 the HI-based radial distances inferred from our own measurements of the inclination of these galaxies (Section 2.4.3). Double pointed arrows encompass the ranges of distances corresponding to the values of inclinations and errors derived from inspection of the residuals of our tilted-ring model fit to the HI velocity fields. The well-known limitations of the weighting scheme included in the modeling of the HI rotation curves when it comes to measuring inclinations below 40° (Andersen & Bershady 2003), have led us to make as well independent fits to the orientation parameters of these disks from Sloan Digital Sky Survey (SDSS) images. We have used the package GALFIT (Peng et al. 2002) to decompose the r-band images of both galaxies into several components. The sky contribution apart, we have fitted a bulge plus an exponential disk to the image of NGC 4411B, while for NGC 4411A a bar component has been added too. We have followed Binney & de Vaucouleurs (1981) and adopted intrinsic axis ratios of 0.18 for NGC 4411A (SBc) and of 0.13 for NGC 4411B (Scd). All this gives estimated inclinations of about 24° for the stellar disk of NGC 4411A and of ~ 16° for the one of NGC 4411B. The uncertainties in these values are difficult to determine, as the analysis of the optical images involves a large number of parameters. Resulting distances are indicated in the plots by asterisks. Comparison with the results from the HI shows that for nearly face-on disks in the Virgo Cluster region inclinations differing by ~ 10° can lead, depending on the TF relationship adopted, to differences in distance exceeding 10 Mpc. The very uncertain radial distances and the lack of unmistakable signs of an ongoing interaction do not make it possible to assert that these galaxies are physically connected in spite of their proximity in z-space.

Another feature of Figure 2.12 that draws one’s attention is the existence of considerable author to author fluctuations in the estimated TF radial distances that cannot be just ascribed to the uncertainties in the observational
parameters entering this relationship. For highly inclined galaxies, such discrepancies can be traced back mostly to the adopted TF template, which for a given passband can show systematic variations among different authors, even when similar samples of calibrators are used. We note, for instance, that small datasets are rather sensitive to the trimming of the data and, hence, to the, somewhat subjective, identification of the most deviant measurements. Besides, differences on the fitting methods, or on the adopted data weighting, as well as morphological and incompleteness biases (see, for instance, Giovanelli et al. 1997b), may give raise to significant variations in the slope and zero-point coefficients of the TF relation. As shown in Figure 2.13, where we compare the $B$-band TF template relations defined in the studies of the Virgo Cluster considered by Solanes et al., the absolute magnitude assigned to a given galaxy, regardless of its inclination, can vary up to $\sim 1.5$ mag depending on the calibration adopted. This exceeds by far the typical error of $\sim 0.1$–$0.2$ mag usually assigned to the zero-point calibration of individual TF templates.

2.6 Summary and conclusions

This is the first 21 cm synthesis survey of spiral galaxies ever made in which the targets have been specifically chosen on the basis of their expected dearth of cold gas. It has been motivated by the detection, in previous investigations of the neutral gas content on spirals in the Virgo Cluster region, of a significant number of severely HI-deficient disks supposedly located, according to their TF distance estimates, beyond the maximum rebound radius galaxies can bounce after infall. According to this location, these galaxies could not owe their HI deficiency to interactions within the cluster environment.

Our high-sensitivity VLA observations have been aimed at characterizing in detail the spatial distribution and kinematics of the neutral gas in four galaxies suspected of being HI-outliers, in a first effort to gain a better understanding of the origin of this class of objects. At the same time, these synthesis observations in the 21 cm line have provided direct evidence of the
risks involved in the application of the TF relationship to disturbed or nearly face-on disks, which can render the derived distances unreliable.

We have detected a total of six galaxies within the four fields initially selected. The main conclusions of the analysis of our VLA data are as follows:

1) We confirm the strong HI deficiency of three of our four main targets, NGC 4307, NGC 4356, and NGC 4492 (inferred HI contents are a factor $\gtrsim 20$ lower than their corresponding standard values), which is pronounced through the reduced extension of the gaseous disks, a characteristic typical of ram-pressure-stripped galaxies. In contrast, we find that the integrated HI fluxes of our fourth target, NGC 4411B, and its companion, NGC 4411A, are $\sim 2$–3 times larger that the old single-dish values used to estimate their HI deficiency in Solanes et al.. Our measured VLA fluxes, —which for all our targets are compatible with those inferred from the new, sensitive ALFALFA extragalactic HI survey—, indicate that the HI contents of these last two galaxies deviate less than $1\sigma$ from normalcy. This is consistent with our observation that their HI disks extend beyond their optical counterparts, so that only the outermost portions of the cold gas distributions have been affected, if at all. A sixth galaxy with a healthy amount of cold gas, the dwarf spiral VCC 740, has been detected in the field of NGC 4356.

2) Visual inspection of the images of the most gas-deficient galaxies has revealed signs of asymmetries and lopsidedness, as well as small offsets of the dynamical centers with respect to the optical ones. These are all suggestive, albeit not conclusive, indications of possible gravitational interactions and/or ram-pressure effects, as deviations from flat, axisymmetric disks are also known to prosper in isolated galaxies. This, and the fact that we have not found evidence of gaseous tails or bridges within the limit we have been able to trace the HI ($\sim 3$–5 $\times 10^{19}$ cm$^{-2}$ channel$^{-1}$ at the 3$\sigma$ level) appear to indicate that none of the galaxies investigated has undergone recent gravitational interactions. This means, in particular, that our VLA observations reinforce the classification of NGC 4411A/B as a virtual pair in spite of their closeness on the observational space phase.
3) Our three targets with highly truncated gas disks exhibit rotation velocities that are still rising at the last measured points. Moreover, the observational evidence gathered do not allow us to assert with complete confidence that the gas remaining tied to the disks has regain dynamical equilibrium, nor the extent to which the luminosity of these H1-deficient galaxies could have been affected. The classification of these objects as H1-outliers could therefore simply obey to the inefficient estimate of their radial distances by means of the TF relationship. The fourth target, NGC 4411B, as well as it space phase neighbor NGC 4411A, show, in contrast, extended gas disks with regular and symmetric velocity fields. In spite of being galaxies presumably in virial equilibrium, their TF-based distances are also problematic because of their nearly face-on apparent orientation, which results in the inability to determine accurate inclinations. This translates into a considerable uncertainty —larger than the resolution necessary to determine unambiguously the region (infall or cluster) where a galaxy belongs— when it comes to placing these two galaxies along the LOS within the Virgo Cluster region.

Aperture synthesis observations in the 21 cm line like the ones presented here are fundamental for probing the impact of cluster residency on the spiral population. Further insight into the identification of the physical processes disturbing the disks can be gained by supplementing this type of data with multifrequency observations.
Part II

Standards of HI content
from the ALFALFA Survey
This chapter is devoted to the definition of a suitable control sample to derive new standards of HI normalcy from the ALFALFA survey. After presenting the two main catalogs used in our investigation, we describe the steps followed to cross-correlate the radio and optical sources and to calculate their radial velocities (Section 3.2), and explore different characterizations of the environment of ALFALFA detections that allow us to define suitable comparison samples for the HI content of galaxies (Section 3.3). Next, we carry out a comparative study of the distribution of the optical structure, colors, and morphologies of the HI sources included in the resulting control samples (Section 3.4). Some ancillary results are relegated to Appendix A, where we display examples of rare, non-conventional, gas-rich early-type galaxies lying in low density regimes. The analysis will be continued in Chapter 4 by applying strategies of multivariate data analysis to the flux-limited subset of HI emitters in low density environments defined here.
3.1 Introduction

One of the defining characteristics of extragalactic research in the last few decades is the execution of wide-area galaxy surveys, which have in many ways revolutionized our understanding of both large-scale structure and galaxy evolution. Most of those surveys have been conducted in the optical/near infrared regime, where stellar light dominates. Multi-color imaging surveys of virtually the entire sky have been complemented by spectroscopic surveys with more than one million objects, notably the Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006), or others, such as the fourth edition of the Sloan Digital Sky Survey (SDSS) Quasar Catalog (Schneider et al. 2007), spanning in redshift from the local universe to near the reionization epoch ($z \lesssim 6$) when the first galaxies may have appeared.

The improvement of technological capabilities in the past decade has enabled the astronomical community to build large statistical samples of extragalactic objects in other regions of the electromagnetic spectrum as well, the radio wavelength regime being one of the most prolific in this respect, thus allowing a more thorough account of the extragalactic realm. While early blind neutral hydrogen (H I) surveys, like the Arecibo H I Strip Survey (Zwaan et al. 1997) or the Arecibo Dual-Beam Survey (Rosenberg & Schneider 2002), yielded useful estimates of the H I mass function from shallow samples containing a relatively modest numbers of detections, the advent of focal plane arrays at 21-cm allowed wide solid angle surveys that led to a significant increase in the depth and, especially, size of the databases. Between 1997 and 2001, the H I Parkes All-Sky Survey (HIPASS; Meyer et al. 2004) used a thirteen-beam array to sample the entire southern sky and the north up to Decl. $< +25^\circ$, with a beam of 15.5 arcmin and a sensitivity of 17 mJy per beam at a spectral resolution of 18 km s$^{-1}$. With a velocity range between $-1280$ and 12,700 km s$^{-1}$ and more than 4000 H I sources identified in the southern sky alone, HIPASS was the first large untargeted H I survey approaching the dimensions necessary to sample large-scale structure on a cosmologically fair volume.
3.1. Introduction

Nowadays, a seven-beam array named ALFA (Arecibo L-band Feed Array), installed at the Arecibo telescope in 2004, is being used to carry out the Arecibo Legacy Fast ALFA Survey (ALFALFA; Giovanelli et al. 2005b), in an effort to map in the HI line 7000 square degrees of the high galactic latitude sky between 0° and +36° in declination. ALFALFA offers significantly higher angular (FWHM ∼ 3′.5) and spectral (∼ 5.5 km s⁻¹) resolution, as well as higher sensitivity (σrms = 2.2 mJy/beam at 10 km s⁻¹ resolution) than any previous surveys of its kind. It is also deeper, with a radial extent (−2000 ≲ czhel ≲ 18,000 km s⁻¹) and a median depth (czhel ≃ 9000 km s⁻¹) enough to sample satisfactorily the largest structures seen in numerical simulations in a ΛCDM world model (∼ 100 h⁻¹ Mpc). While HIPASS detected 1 galaxy every ∼ 5 deg², ALFALFA is detecting ∼ 5 galaxies deg⁻². The current detection rate suggests that ALFALFA will be able to detect some in excess of 30,000 extragalactic HI line sources when completed. An account of the ALFALFA survey goals and technical characteristics, as well as its main science contributions, can be found in Giovanelli et al. (2005b) and in the public website http://egg.astro.cornell.edu/alfalfa/ and references therein.

In this work, we deal with a compilation of the new ALFALFA data, partially released already in several installments (Giovanelli, Haynes, Kent, Saintonge, Stierwalt, Altaf, Balonek, Brosch, Brown, Catinella, Furniss, Goldstein, Hoffman, Koopmann, Kornreich, Mahmood, Martin, Masters, Mitschang, Momjian, Nair, Rosenberg, & Walsh 2007; Saintonge, Giovanelli, Haynes, Hoffman, Kent, Martin, Stierwalt, & Brosch 2008; Kent, Giovannelli, Haynes, Martin, Saintonge, Stierwalt, Balonek, Brosch, & Koopmann 2008; Stierwalt, Haynes, Giovannelli, Kent, Martin, Saintonge, Karachentsev, & Karachentseva 2009; Martin, Giovannelli, Haynes, Saintonge, Hoffman, Kent, & Stierwalt 2009), in two separate sky strips of the northern hemisphere covering a total area of 9 hr in right ascension and 16 deg in declination. This portion of the survey is especially interesting because it overlaps almost entirely with the SDSS footprint, thus offering the possibility of cross-correlating observations of the cold-gas component with multi-wavelength optical measurements. This synergy between both surveys will be exploited
in this chapter and in Chapter 4 to carry out a systematic analysis of the fundamental properties of nearby galaxies and their inter-relations with the hope of gaining a more complete understanding of the structure of the baryonic component in gaseous galaxies and setting up unbiased benchmarks for the normalcy of their \( \text{H} \text{I} \) content that rely on integral attributes easily accessible to observation.

3.2 Main 21-cm and optical catalogs

We describe here the primary datasets from which we have assembled the galaxy samples used in the present study and examine their statistical properties. This section also outlines the steps followed to add to the available measurements information on the cosmic radial distances of the selected galaxies.

3.2.1 Statistical Properties of Catalogs and Cross-identification of Sources

The main catalog used in this thesis contains ALFALFA H\( \text{I} \) measurements distributed in two regions of the high Galactic latitude sky: \( 07^\text{h}30^\text{m} < \text{R.A.} < 16^\text{h}30^\text{m} \), \( +4^\circ < \text{Decl.} +16^\circ \) and \( +24^\circ < \text{Decl.} < +28^\circ \) (J2000.0 coordinates).

The ALFALFA H\( \text{I} \) detection threshold and completeness limit depend on the source signal-to-noise ratio, \( S/N \), given by the expression (cf. Martin, Giovanelli, Haynes, Saintonge, Hoffman, Kent, & Stierwalt 2009):

\[
S/N = \frac{1000 F_{\text{H} \text{I}}}{W_{50}} \frac{u_{\text{limo}}^{1/2}}{\sigma_{\text{rms}}},
\]

(3.2.1)

where \( F_{\text{H} \text{I}} \) is the 21-cm line flux integral in Jy km s\(^{-1}\), which in combination with the cosmological distance \( d \) to the source in \( h^{-1} \) Mpc (see Section 3.2.2) provides its total neutral gas mass in \( h^{-2}M_{\odot} \) units

\[
M_{\text{H} \text{I}} = 2.356 \times 10^5 d^2 F_{\text{H} \text{I}},
\]

(3.2.2)
with $W_{50}$ the observed velocity width of the source line profile at the 50% level of the two peaks corrected for instrumental broadening, $w_{\text{smo}}$ a smoothing width (either $W_{50}/20$ for $W_{50} < 400 \text{ km s}^{-1}$ or 20 for $W_{50} \geq 400 \text{ km s}^{-1}$), and $\sigma_{\text{rms}}$ the rms noise figure measured in mJy at 10 km s$^{-1}$ resolution. ALFALFA detections are coded according to a scheme dependent on $S/N$ and availability of a prior, e.g., an optical source matching in angular coordinates or redshift. Code 1 sources are detections of $S/N > 6.5$ and thought to be reliable with high confidence level; code 2 sources have $4 < S/N < 6.5$ typically, but the marginal HI detection is corroborated by a match in position and redshift of an optical source. By searching through the ALFALFA data for sources coded 1 and 2, we have come up with 15,043 extragalactic objects selected purely on their HI content in the sky region defined above and over the full survey bandwidth.

The present portion of the ALFALFA survey overlaps with other major galaxy searches like 2MASS and SDSS. The latter—which has scanned about one quarter of the northern Galactic cap sky in five bands $u, g, r, i,$ and $z$, reaching objects up to a total magnitude 22.2 in $r$ band (Stoughton et al. 2002)—is especially interesting for our purposes, as it provides a uniform, very extensive list of precise photometric and spectroscopic measurements suitable to carry out an exhaustive investigation of the correlations between the stellar and gaseous properties of galaxies. Here we make use of SDSS data coming from the full survey, Data Release Seventh (SDSS DR7; Abazajian et al. 2009), which also provides radial velocity measurements for a considerable number of its main targets. The spectroscopic observations, which are complete to a Petrosian $r$-band magnitude of 17.77, after correcting for Galactic extinction (but not for internal reddening), offer a redshift coverage deep enough to make it possible the identification of the most probable optical counterpart to all ALFALFA detections found out to 250 Mpc ($h = 0.72$) within the sky region where the two surveys overlap\(^1\), whose geometry we have defined using mangle polygons (Hamilton & Tegmark 2004). A graphical representation of the positions of the ALFALFA and SDSS DR7 spectroscopic sources in the overlapping region is provided in Figure 3.1 by

\(^1\)See \url{http://egg.astro.cornell.edu/index.php/} for an image of the footprint of ALFALFA on the SDSS.
Figure 3.1 Sky distribution of ALFALFA (top) and SDSS DR7 spectroscopic sources (bottom) in the sky region where both surveys overlap.
3.2. Main 21-cm and optical catalogs

means of wedge diagrams in right ascension. The total number of HI detections in this portion of the sky amounts to 11,239 once those objects lying closer than 1.5 deg to M87, where HI sources cannot be reliably detected because of the strong radio continuum emission, are discarded.

Some of the main statistical properties of this data set are displayed in Figure 3.2. In the top-left panel, we show the distance distribution of the ALFALFA sources. The most important contributions to the large-scale structure are the various peaks between 60 and $110\,h^{-1}\text{Mpc}$ produced by several clusters and groups of galaxies embedded in the Great Wall of the CfA Survey (Ramella et al. 1992; see also Figure 3.1). The shape of this histogram follows closely that produced by the distance distribution of SDSS galaxies, except as regards the narrow artificial dip seen near $\sim 88\,h^{-1}\text{Mpc}$ and the broader one between $\sim 150$ and $160\,h^{-1}\text{Mpc}$, both related to the important loss in sensitivity of the ALFALFA survey due to the RFI coming from the San Juan airport’s radar (Giovanelli et al. 2007). The HI mass distribution in $h^{-2}\text{M}_\odot$ units is shown in the top-right panel of Figure 3.2. The vertical dotted line has been drawn to illustrate the minimum HI mass arising from the lower heliocentric velocity cutoff of 2000 km s$^{-1}$ adopted to avoid the most uncertain values of this quantity (note that this also leaves out of the present study most of the nearby dwarfs rich in neutral gas). Below, the bottom-left panel features the Spähnauer diagram, detailing the HI mass vs. CMB distance for the objects in the ALFALFA catalog. The broad gap in detections associated with the region most affected by the intrusion of RFI is clearly visible near the right end. Finally, the bottom-right panel of Figure 3.2 shows the (logarithm of the) intrinsic velocity width distribution in km s$^{-1}$. As in the top panels, the hatched histogram indicates the distribution arising from code 1 objects only.

Contrary to what happens with other HI surveys, thanks to the improved angular resolution of ALFALFA —which leads to positional accuracies typically better than 20 arcsec (Giovanelli et al. 2007)—, the identification of the optical counterpart of a given HI source is usually an unambiguous task. Besides, we have taken advantage of the fact that, for each HI detection, the ALFALFA catalog provides the centroid of the optical galaxy that appears to
be the most plausible counterpart according to proximity in $z$-space, color, and morphology (Giovanelli et al. 2007), to perform a semi-automated catalog cross-correlation process. Specifically, we have assumed that a HI and a spectroscopic SDSS detection are the same object if their projected separation on the plane of the sky is not larger than the size of the optical source and the difference in the reported velocities is less than 300 km s$^{-1}$. We allow for such a relatively large discrepancy between the observed heliocentric velocities in order to account for the fact that SDSS spectroscopic measure-
ments in an extended object do not always reflect the radial velocity of the object’s centroid, so a positive match between two galaxies relies essentially in the similarity of their sky coordinates. In practice, however, the dispersion in the difference between the radial velocities of matched pairs is only of $\sim 35$ km s$^{-1}$ (see Figure 3.3). For the cases in which ALFALFA sources were lacking a spectroscopic counterpart, the photometric criterion was applied. Human intervention was reduced just to doubtful identifications.

Figure 3.3 Left: Distribution of differences between HI and SDSS sky positions for ALFALFA detections with an SDSS photometric counterpart. Right: Distribution of differences between HI and SDSS radial heliocentric velocities for ALFALFA detections with an SDSS spectroscopic counterpart.

Table 3.1 lists the numbers (and fractions) of ALFALFA objects with a photometric and spectroscopic counterpart in the SDSS in different bins of radial heliocentric velocity. Only $\sim 2\%$ of the HI sources are not associated with galaxies listed in the SDSS photometric catalog, while there is a somewhat larger percentage, $\sim 15\%$, of HI detections without a spectroscopic counterpart. Note also how the mild incompleteness in the SDSS DR7 data set affecting mainly the galaxies with the largest apparent brightness (e.g., Strauss et al. 2002), leaves its imprint in the form of reduced fractions for the first two velocity bins of the third and fourth columns of the Table. Also included in Table 3.1 are the numbers (and fractions) of ALFALFA objects detected within the ‘SDSS spectroscopic galaxy sample’, SDSS-spec sample for short. From now on, we will use this latter designation to refer to the subset of SDSS sources with both photometric and spectral morphology classified as of galaxy type, and having $r < 17.77$ mag. About one in every five
Table 3.1. Distributions of ALFALFA and SDSS Galaxies in Radial Velocity Bins

<table>
<thead>
<tr>
<th>$\Delta v_{\text{hel}}$ (km s(^{-1}))</th>
<th>ALFA $N_{\text{gal}}$</th>
<th>ALFA/SDSS-phot $N_{\text{gal}}$ (%)</th>
<th>ALFA/SDSS-spec* $N_{\text{gal}}$ (%)</th>
<th>SDSS-spec $N_{\text{gal}}$</th>
<th>SDSS-spec/ALFA $N_{\text{gal}}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 500</td>
<td>71</td>
<td>65 (91%)</td>
<td>50 (77%)</td>
<td>37</td>
<td>18 (49%)</td>
</tr>
<tr>
<td>500–4000</td>
<td>1523</td>
<td>1417 (93%)</td>
<td>1024 (72%)</td>
<td>1702</td>
<td>910 (53%)</td>
</tr>
<tr>
<td>4000–7500</td>
<td>3104</td>
<td>3066 (99%)</td>
<td>2529 (82%)</td>
<td>5102</td>
<td>2420 (47%)</td>
</tr>
<tr>
<td>7500–11,000</td>
<td>3385</td>
<td>3353 (99%)</td>
<td>2947 (88%)</td>
<td>9196</td>
<td>2853 (31%)</td>
</tr>
<tr>
<td>11,000–14,500</td>
<td>2280</td>
<td>2261 (99%)</td>
<td>2025 (89%)</td>
<td>9341</td>
<td>1977 (21%)</td>
</tr>
<tr>
<td>14,500–18,000</td>
<td>876</td>
<td>871 (99%)</td>
<td>807 (93%)</td>
<td>10187</td>
<td>787 (8%)</td>
</tr>
<tr>
<td>Totals</td>
<td>11,239</td>
<td>11,033 (98%)</td>
<td>9382 (85%)</td>
<td>47,026</td>
<td>8965 (19%)</td>
</tr>
</tbody>
</table>

Note. — Numbers in different velocity bins of ALFALFA galaxies (ALFA), H\(_I\) detections with a photometric counterpart in SDSS (ALFA/SDSS-phot), and H\(_I\) detections with a spectroscopic counterpart in SDSS (ALFA/SDSS-spec*), regardless of the morphological classification of the SDSS spectrum. SDSS-spec designates the sample of SDSS sources with both photometric and spectral morphology classified as of galaxy type and having $r < 17.77$ mag, while SDSS-spec/ALFA identifies the subset of SDSS-spec galaxies with an ALFALFA counterpart. Fractions are given in parentheses.
galaxies in this subset is also an ALFALFA detection, although the fraction rises to one in two for $c_{\text{hel}} \leq 7500$ km s$^{-1}$. The small value quoted in the higher velocity entry of the last column of the Table reflects also the impact of the substantial incompleteness of the ALFALFA data due to RFI above $\sim 15,000$ km s$^{-1}$. The effects on the completeness of a few other narrower RFI features present in the spectral bandpass are comparatively modest.

In Figure 3.4 we investigate how the galaxies in both catalogs distribute in a global color-light concentration parameter space. In this plot the density contours delineate the behavior of the galaxies within the SDSS-spec sample, while the data points correspond to those objects that are also ALFALFA’s members. It is obvious from this figure that SDSS galaxies show a clear bimodality separating into two main classes with comparable numbers of objects that can be addressed as the blue and red-sequence galaxies (e.g. Blanton et al. 2003). In contrast, ALFALFA detections populate mostly the locus corresponding to late-type, star forming systems, in very good agreement with the similar offset found by West et al. (2009) using 195 HI-selected galaxies from the Parkes Equatorial Survey (Garcia-Appadoo et al. 2009). We will show in Section 3.4.1 that, while the bimodality of the optically selected sources gets reinforced when looking at the global trend with increasing radial velocity, and hence with increasing typical luminosity of galaxies, HI detections distribute in a similar way independently of their distance to the observer.

3.2.2 Cosmic velocity and radial distance of sources

To calculate radial velocities and distances referred to the Cosmic Microwave Background (CMB) standard of rest for the galaxies in both the ALFALFA and SDSS spectroscopic datasets, we correct first the heliocentric velocities listed in these two catalogs for peculiar motions. This is done by following the prescriptions of the NYU Value-Added Galaxy Catalog$^2$ (NYU-VAGC; Blanton et al. 2005), which uses a model of the local velocity field deter-

$^2$Currently, distances corrected for peculiar velocities are only available for the Sixth Data Release of the SDSS.
mined by Willick et al. (1997) from the 1.2 Jy redshift survey of the Infrared Astronomical Satellite (IRAS; Fisher et al. 1995). This model predicts non-negligible peculiar motion corrections for objects within 6400 km s$^{-1}$ in the Local Group (LG) frame of reference. The transformation to the LG frame is performed by applying a correction of 308 km s$^{-1}$ in the direction ($l^{II}, b^{II}$) = (105°, −7°) (Yahil et al. 1977). For transformations between heliocentric or LG and CMB references we apply the corrections quoted in Kogut et al. (1993). For the conversion to a CMB luminosity distance we assume Euclidean properties of space, an approximation that is maintained for the calculation of any galactic property involved in the present study.

To minimize the impact of redshift-space distortions in the assignment of distances (and, consequently, in local density estimates; see Section 3.3.1) to galaxies lying in the vicinity of the clusters and groups present in the cosmic volume targeted by ALFALFA, we adopt an strategy that can be

![Figure 3.4 Inverse light concentration index in the r band, $C_{59,r}$, versus $(u - r)$ color for members of the SDSS spectroscopic sample (black contours), and for those that are also ALFALFA detections (blue dots). Contours increase by a factor of 40 in density. The histograms on the right and top sides of the main panel show the distributions of the corresponding variables for these two datasets normalized to peak values.](image-url)
3.3. HI control samples

easily applied in a systematic way to any virialized aggregation of galaxies independently of its scale. It consists of estimating the cluster/group virial radius and its velocity dispersion along the line-of-sight, and then using this information to assign to all galaxies with projected distances to the center of one of these overdense environments less than 1.5 $r_{\text{vir}}$ and systemic velocities with respect to the system mean velocity less than 3$\sigma$ the corresponding system distance. For the former quantity, we use an approximation to the halo density-radius relation based on a singular isothermal sphere with mean interior density 200 times the current value of the critical density of the universe (cf. Carlberg, Yee, & Ellingson 1997). In units of physical Mpc:

$$r_{\text{vir}} \simeq \frac{\sqrt{3} \sigma}{1000h}.$$  \hfill (3.2.3)

Values for the characteristic velocity dispersion, $\sigma$ [km s$^{-1}$], of clusters mostly come from the compilation of Abell cluster properties by Struble & Rood (1991). For those galaxy aggregations that have not quoted $\sigma$ in the latter catalog, we employ the membership assignments carried out by Springob et al. (2007) in the estimator (e.g., Tago et al. 2008)

$$\sigma^2 = \frac{1}{(1 + \bar{z})^2} \sum_{i=1}^{N_v} \frac{(cz_i - c\bar{z})^2}{N_v - 1}$$ \hfill (3.2.4)

where the sum is intended over the $N_v$ members of the aggregation with measured velocity $cz_i$, and $\bar{z}$ is the mean system redshift.

### 3.3 HI control samples

The definition of a benchmark suitable for HI content measurements (see Chapter 4) necessitates a control sample as homogeneous as possible formed by galaxies expected to show little or no evidence of interaction with their surroundings. In this section, we investigate different characterizations of the local environment of ALFALFA detections, define three different, possible control samples, and explore their relative degree of adequacy.
3.3.1 Local number density measurements for ALFALFA and SDSS galaxies

The environment of a galaxy can be defined in terms of the density of galaxies located in its immediate vicinity. To calculate the underlying three-dimensional galaxy number density at a given position, we use the galaxies in the SDSS spectroscopic catalog as the nodes of a $n$th-nearest neighbor network. This calculation is carried out for both ALFALFA and SDSS sources, excluding galaxies lying in those regions in which it is affected by larger errors or that are useless for the goals of the present study. In particular, we impose a lower heliocentric velocity limit of 2000 km s$^{-1}$ to avoid those galaxies likely having the largest relative uncertainty in distance arising from peculiar motions not duly accounted for—a constraint that also removes the closest cosmic volume where the SDSS pipeline photometry of large, extended galaxies is most inaccurate—, and a velocity cutoff ranging from 15,000 km s$^{-1}$ up to 16,000 km s$^{-1}$ that embraces the deep trough in sensitivity of the ALFALFA’s spectral window due to radio frequency interference (RFI) (Giovanelli et al. 2007; see also Figures 3.1 and 3.2).

We follow Casertano & Hut (1985) and compute a local, continuous measure of environment from the unbiased estimator

$$\rho_6 = \frac{15}{4\pi d_6^3},$$

(3.3.1)

where $d_6$ is the three-dimensional (comoving) distance to the 6th-nearest neighbor in the SDSS-spec sample. In an effort to correct for redshift-space distortions and improve the accuracy of our density measures in overdense environments $\rho_6$ is calculated, for those galaxies assigned to a group or cluster (Section 3.2.2), using instead projected separations scaled by a factor of $(3/2)^{1/2}$ to convert them into a three-dimensional measure of distance (Cooper et al. 2005).

Local densities calculated this way are then compensated for the magni-
3.3. HI control samples

tude cutoff of the SDSS catalog by multiplying them by the inverse of the selection function

\[ \psi(\alpha, \delta, z) = \frac{\int_{-\infty}^{M_{\text{lim}}(z) + \Delta M(\alpha, \delta)} \phi(M) dM}{\int_{-\infty}^{M_{\text{fid}}} \phi(M) dM}, \] 

(3.3.2)

which we calculate from the differential SDSS luminosity function provided by Blanton et al. (2001). In equation (3.3.2), \( M_{\text{lim}}(z) \) and \( M_{\text{fid}} \equiv M_{\text{lim}}(z_{\text{fid}}) \) are, respectively, the faintest magnitudes detectable because of the optical flux limitation at the galaxy redshift and at a fiducial redshift, \( z_{\text{fid}} \), which we match up with that corresponding to the low-velocity cutoff of 2000 km s\(^{-1}\) imposed to the samples, while \( \Delta M(\alpha, \delta) \) is the extinction correction as a function of right ascension (\( \alpha \)) and declination (\( \delta \)), for which we use the Schlegel, Finkbeiner, & Davis (1998) values loaded in the SDSS database.

Figure 3.5 Three-dimensional local density of galaxies derived using the distance to the 6th-nearest neighbor in the SDSS spectroscopic dataset (\( \rho_6 \)) vs. the corresponding projected density estimator calculated from the 2MASS galaxy distribution in the \( K_s \) band down to 13.0 mag (\( \mu_6 \)). We show results for SDSS galaxies (top) and for those that are also ALFALFA detections (bottom). Galaxies classified as cluster members are drawn using big red dots, whereas the black color is used for the clusters’ outskirts population. In these plots, the values of \( \rho_6 \) for cluster members have not been corrected for redshift-space distortions.

It is worth noting, however, that this measure of the local density cannot be expanded to the full ALFALFA survey, which is also carrying observations within the solid angle \( 22^h \leq \text{R.A.}(\text{J2000.0}) \leq 03^h \) and \( 0^\circ \leq \text{Decl.}(\text{J2000.0}) \leq +36^\circ \), not scanned by the SDSS. In this case, the 2MASS (Skrutskie et al. 2006), with its full sky coverage (98% for the extended source catalog), offers
the most sensible alternative to characterizing the environment in which ALFALFA detections reside. Thus, in anticipation of future research, we have also assigned a surface density parameter $\mu_6$ to each optical and gaseous spectroscopic source from the formula

$$\mu_6 = \frac{5}{\pi s_6^2},$$

(3.3.3)

where $s_6$ is the projected distance to the 6th-nearest neighbor derived taking into account the underlying 2MASS galaxy distribution on the sky in the $K_s$ band down to 13.0 mag. We do not correct this surface density estimator from Galactic absorption, because at the adopted flux limit —the completeness in this extreme near-infrared wavelength extends down actually to 13.5 mag for extended sources— the opacity is insignificant for $|b^{II}| \gtrsim 10^\circ$ (Jarrett et al. 2000). It should not be forgotten, however, that this immunity to the effects of the Milky Way’s dust arising from 2MASS being a survey in the near-infrared, biases its sensitivity regarding the morphology of the detected galaxies toward early-type spirals and ellipticals, whose light is dominated by old stars, in front of late-type spirals, low surface brightness, and compact objects (the latter due to resolution limitations).

In Figure 3.5, we compare the values of $\rho_6$ (uncorrected for redshift-space distortions) and $\mu_6$ for galaxies in the SDSS spectroscopic (top) and ALFALFA (bottom) datasets. The plots show that these two local density tracers are just weakly correlated, with the three-dimensional measure of environment spanning a somewhat large dynamic range. In addition, one can see that the correlation strength for SDSS galaxies is not uniform, growing weaker and less linear as the local density and, hence, the redshift-space distortions on the three-dimensional measure of density increase. In contrast, because of the undersampling of the highest density environments —note the relatively few observations associated with clusters in the bottom plot (see also next section)—, ALFALFA galaxies show essentially a similar correlation strength on all scales.
3.3. HI control samples

3.3.2 Low Density Environment HI galaxy sample

One possibility to define a sample of galaxies whose evolution is hypothetically less influenced by external processes is to select those objects least affected by the clustering phenomenon. According to observational findings the influence of the cluster environment on the properties of galaxies can extend up to about twice the cluster virial radius (e.g., Balogh et al. 1998; Solanes et al. 2001). Therefore, any density threshold adopted to extract a control sample of unperturbed galaxies from a given catalog should be low enough to leave out of it, not only bona-fide cluster and group members, but also those objects located in adjacent regions.

We have studied the relationship between the neutral gas content of ALFALFA galaxies and their local three-dimensional number density measured in the previous section to see if we can identify the density threshold above which environmental interactions start to become important. Following the works by Haynes & Giovanelli (1984) and Solanes et al. (1996), the former property has been quantified by means of a HI-deficiency parameter, DEF$_{HI}$, that compares (the decimal logarithm of) the observed HI mass, $M_{HI}^{obs}$, in solar units, with the value expected from a galaxy free of external influences of the same observed morphological type, $T_{obs}$, and optical linear diameter, $D_{obs}$, expressed in kpc. Specifically:

$$\text{DEF}_{HI} = \langle \log M_{HI}(T_{obs}, D_{obs}) \rangle - \log M_{HI}^{obs},$$

so positive values of DEF$_{HI}$ indicate HI deficiency. We have inferred expectation values for the HI mass for disk galaxies with Sa–Sd morphologies using the maximum likelihood linear regressions of log $M_{HI}$ on log $D$ given in Solanes et al. (1996). Since in that work the standards of normalcy for the HI content were defined specifically for galaxies of type Sa–Sc, we thus have assigned Scd–Sd types to be the same as Sc. This test has been restricted to those ALFALFA detections with Hubble types and (apparent) blue visual sizes available on the Arecibo General Catalog (AGC), a private galaxy database maintained by M. P. Haynes and R. Giovanelli (Cornell University).
3. Selection of a control sample

Figure 3.6 Medians and inter-quartile ranges of the HI-deficiency parameter, DEF$_{HI}$, calculated using the standards of HI content given by Solanes et al. (1996), in equal bins of $\log \rho_6$ up to $\log \rho_6=1$. Above the latter value, only two equal bins are considered. Results for bins containing less than 25 objects are represented by thinner lines. The different panels are for ALFALFA detections of Sa–Sd type with radial heliocentric velocities between 2000 and 18,000 km s$^{-1}$ (top-left), as well as for the three subintervals: 2000–7000 km s$^{-1}$ (top-right), 7000–11,000 km s$^{-1}$ (bottom-left), and 11,000–18,000 km s$^{-1}$ (bottom-right).

We present in Figure 3.6 a plot of the median DEF$_{HI}$ in bins of local number density. We show results inferred using sources with radial velocities between 2000 and 18,000 km s$^{-1}$, as well as separately for three subintervals of this velocity range. Looking at the different panels one can see moderate oscillations in the central values of DEF$_{HI}$, which nonetheless are always well within the quoted uncertainties and, therefore, not statistically significant and that, at best, hint at a mild increase of gas deficiency with density. The observed nearly independence of the gas content of the ALFALFA sources with environment is likely to be the consequence of limiting the effective integration time of the survey, which has not among its goals the detection of the gas-poor galaxies that are expected to reside in the cores of many rich
3.3. HI control samples

clusters (Giovanelli & Haynes 1985; Solanes et al. 2001), to about 48 seconds per beam area (dedicated HI observations of cluster galaxies require usually much larger, typically one hour, integration times). This makes ALFALFA partial to galaxies with relatively large amounts of HI, unless they are located nearby, say at radial distances $\lesssim 20$ Mpc (for an illustration of the detection capabilities of the ALFALFA survey regarding local galaxies with low or very low HI mass see, for instance, the study of early-type galaxies (ETG) in the Virgo Cluster region by di Serego Alighieri, Gavazzi, Giovanardi, Giovanelli, Grossi, Haynes, Kent, Koopmann, Pellegrini, Scodeggio, & Trinchieri 2007). Evidences such as the negative median DEF$_{HI}$ found in all the velocity intervals, or the progressively more negative values achieved by this parameter with increasing radial velocity, support this interpretation (see also Figure 3.5).

The lack of a clear HI-deficiency-density correlation for the portion of the ALFALFA spring-sky survey under study notwithstanding, we prefer to be conservative and, as indicated at the beginning of the section, define a control sample that does not include galaxies located in the vicinity of an overdensity. After testing different density thresholds, we find that the condition $\rho_6 \leq 0.5 h^3$ galaxies/Mpc$^3$ discards all galaxies cataloged as cluster/group members in the surveyed region$^4$, as well as more than 91 (96) per cent of the SDSS objects with system-centric distances between $1.5 - 3 r_{\text{vir}}$ ($1.5 - 2 r_{\text{vir}}$) and systemic velocities within $1 \sigma$ of the cluster/group mean, which may be considered representative of the system outskirts population. When referred to HI detections only, these percentages increase up to 93 and 97, respectively. The resulting Low Density Environment (hereafter LDE) HI galaxy sample contains a total of 5496 ALFALFA objects. Wedge diagrams in right ascension showing the spatial distribution of this data set, as well as of the HI detections discarded for lying in higher density environments, are shown in Figure 3.7.

$^4$The corresponding threshold based on the 2MASS galaxy distribution would be $\mu_6 \sim 10$ galaxies/deg$^2$ (see Figure 3.5).
Figure 3.7 Wedge diagrams in right ascension of the sample of ALFALFA galaxies in Low Density Environments (top) and of the HI detections excluded for lying in environments of higher density (bottom).
3.3. HI control samples

3.3.3 Isolated galaxy samples

While by selecting galaxies in low-density regimes one may confidently exclude objects associated with groups and clusters, it is not possible to guarantee that these galaxies have not interacted at some point with nearby companions. For this reason, we have also investigated the possibility of defining control samples from the ALFALFA catalog that maximize the chances of dealing with field galaxies whose evolution has not been influenced by other objects in the recent past. Here we build two samples of isolated galaxies by applying two different kinds of isolation criteria.

As a first approach, we have devised a selection technique, similar to the one used by Prada et al. (2003) on the SDSS spectroscopic dataset, that combines spectroscopic with photometric information to detect galaxies without dynamically relevant companions inside a given volume in z-space. Thus, we classify a galaxy \( i \) with apparent \( r \)-band total magnitude \( r_i \) at coordinates \((\alpha_i, \delta_i, z_i)\) as a good candidate for a sample of isolated galaxies if it has no neighbor \( j \) with \( r_j \leq r_i + 2.5 \) mag within the cylindrical volume, centered on its position, of radius and semi-axis

\[
R = R_0 \psi^{-1/3} \quad \text{and} \quad V = V_0 \psi^{-1/3} + \sigma_{12}, \tag{3.3.5}
\]

respectively, with \( R_0 = 0.5 \, h^{-1} \) Mpc the linking-length at the adopted fiducial redshift of 2000 km s\(^{-1}\), \( V_0 = H_0 R_0 \), and \( \psi(\alpha_i, \delta_i, z_i) \) the selection function defined in equation (3.3.2) that keeps constant the probability of finding a galaxy in a given volume regardless of distance. In equation (3.3.5), the term \( \sigma_{12} = 350 \) km s\(^{-1}\) is added to account for the distortion on the intergalactic distances along the line-of-sight induced by the one-dimensional pairwise velocity dispersion of field galaxies (Guzzo et al. 1997; Landy 2002). The preference of a \( 0.5 \, h^{-1} \) Mpc (~0.7 Mpc for \( h = 0.7 \)) linking-length has been dictated from the fact that it would take almost 3 Gyr for a galaxy with a peculiar velocity of ~175 km s\(^{-1}\) to cover such a distance, while the magnitude constraint is set to take into account neighbors that have at least one-tenth of the target mass (assuming light is a proxy for mass). This technique identifies 443 HI detections as members of this first isolated galaxy
3. Selection of a control sample

(hereafter IG1) sample. A second isolated galaxy (IG2) sample has been defined by applying one of the isolation criteria typically used for photometric samples—usually much deeper than the spectroscopic ones—, which take into account in a semi-empirical way the distance to companions. Perhaps the best known example is the nearest neighbor algorithm by Karachentseva (1973) based on the comparison of apparent angular diameters. More recently, Allam et al. (2005) have implemented a variation of Karachentseva technique to detect isolated galaxies in the SDSS DR1 that we adopt for the present investigation. According to this criterion, a galaxy $i$ with a $g$-band total magnitude $g_i$ and $g$-band Petrosian radius $R_i$ is considered to be isolated if it satisfies the conditions: $\theta_{i,j} \geq 40R_j$ and $|g_i - g_j| > 3.0$ mag, with $\theta_{i,j}$ the projected sky separation between the target galaxy and any neighboring galaxy $j$. A total of 792 isolated H I emitters are identified by this approach.

It is important to note, however, that because of the apparent magnitude limits of the SDSS, all the galaxies that belong to the IG2 subset have $g$-band magnitudes obeying the constraint $16 \leq g_i \leq 19$ mag. Since the spectro-photometric algorithm imposes the restriction $r_i \leq 15.27$ on the magnitude of the IG1 objects, we find that there is a minimal overlap in the range of apparent brightnesses of the galaxies that can be part of both samples. In practice, this means that not a single ALFALFA detection ends up obeying simultaneously the two isolation criteria. Had we eliminated from the photometric criterion the constraint $g \geq 16$ mag with the intention of facilitating the overlap of both subsets of isolated galaxies, then the number of IG2 members would have risen to 2302. Even so, this enlarged sample would include only 288 of the 443 isolated galaxies selected from spectro-photometric algorithm. It is then obvious that isolated galaxies in $z$-space are not necessarily considered so when their environment is examined in projection and vice versa.

As done in the previous section for the LDE H I galaxy sample, we also investigate the environment of the ALFALFA galaxies in the IG1 and IG2 subsets. Figure 3.8 shows a graphical representation of the H I-deficiency versus local density for the galaxies in the IG1 subset. As before, we focus
3.3. HI control samples

Figure 3.8 Top: DEF$_{\text{HI}}$, calculated using the standards of HI content given by Solanes et al. (1996), vs. log $\rho_6$. ALFALFA detections of Sa–Sd type selected by an spectro-photometric isolation criterion (see text) are shown as big orange diamonds. Dots represent all Sa-Sd ALFALFA detections spanning the same range of $g$-band magnitudes than the isolated objects ($16 \leq g \leq 19$). The subset of galaxies left to the vertical dashed line (galaxies in LDEs) is highlighted in blue color. Bottom: Corresponding medians and inter-quartile ranges of DEF$_{\text{HI}}$. Results for all the ALFALFA galaxies depicted in the main panel are represented by error bars in equal bins of log $\rho_6$ up to log $\rho_6 = 1$, while above the latter value only two equal intervals are considered. Results for bins containing less than 25 objects are drawn using thinner lines. The orange horizontal lines show the median (solid) and upper and lower quartiles (dotted) of DEF$_{\text{HI}}$ inferred from all the galaxies in the isolated subset. Right: Histograms normalized to peak values showing the distribution of DEF$_{\text{HI}}$ for the subsets of LDE galaxies (blue) and isolated sources (orange).

only on sources having optical diameters and spiral types from Sa to Sd listed on the AGC. It is clear from this figure that while isolated galaxies obtained through automated systematic searches avoid the densest regions of the universe, a non-negligible fraction of them inhabits regions of moderate galaxy density having values of $\rho_6$ larger than the upper density threshold adopted to define our LDE HI galaxy sample. This highlights the fact that algorithms defining isolation and a low density environment are not fully-equivalent: lo-
selection of a control sample

This page of the document discusses the selection of a control sample for isolated galaxies. It mentions that number density measurements associated with isolated galaxies can span quite a broad range of values, as noted in Verley et al. (2007). At the same time, Figure 3.8 illustrates that while all these algorithms select substantially different galaxy samples, the resultant H\textsubscript{I} content distributions turn out to be quite similar. There is therefore no evidence that our systematic searches of isolated galaxies have selected objects that are particularly gas-rich.

3.4 Underlying properties of galaxies in the H\textsubscript{I} control samples

In this section, we use several independent classification schemes based on the photometric data of the galaxies included in our three control samples in order to fill in more details that increase our understanding of the objects that belong to these subsets. These are measures sensitive to the morphology that are objective, easily reproducible, and applicable to data sets in which the morphological classification by direct inspection of the galaxy images becomes impractical, given the large number of objects involved.

3.4.1 Structure, colors, and morphology of the galaxies in LDEs

Since we are focusing our attention on subsets of nonclustered, gas-rich galaxies, predictably of late-type, we first apply the criterion used by Maller et al. (2009) to select disk galaxies from the SDSS. As stated by these authors, almost pure disk samples can be gotten by selecting galaxies that obey at least one of these requirements: Sérsic index \( n \leq 3 \) and observed axis ratio \( b/a \leq 0.55 \), where the parameter \( n \), —which in our case has been extracted from the NYU-VAGC—, measures the shape of the observed \( r \)-band luminosity profile of a galaxy fitted using the Sérsic \( R^{1/n} \) formula with elliptical isophotes. We note that these constraints include S0 galaxies into the ‘disk’ class, because the presence of a disk makes their measured axis ratio more im-
3.4. Properties of the HI control samples

Table 3.2. ALFALFA Detections and SDSS Galaxies in LDEs Obeying Maller et al.’s Criterion

<table>
<thead>
<tr>
<th>$\Delta v_{\text{hel}}$ (km s$^{-1}$)</th>
<th>SDSS-spec</th>
<th>SDSS-spec/ALFA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N_{\text{gal}}$</td>
<td>$N_{\text{disk}}$ (%)</td>
</tr>
<tr>
<td>2000–7000</td>
<td>1957</td>
<td>1801 (92.0%)</td>
</tr>
<tr>
<td>7000–11,000</td>
<td>3815</td>
<td>3320 (87.0%)</td>
</tr>
<tr>
<td>11,000–18,000</td>
<td>8215</td>
<td>6737 (82.0%)</td>
</tr>
<tr>
<td>Totals</td>
<td>14145</td>
<td>11991 (84.8%)</td>
</tr>
</tbody>
</table>

Note. — Percentage of SDSS-spec and SDSS-spec/ALFA galaxies (the definition of these subsets can be found in the caption of Table 3.1) in low density environments satisfying Maller et al.’s (2009) disk-identification criterion ($N_{\text{disk}}$). We only consider galaxies with Sérsic index $0.5 < n < 5.9$ and axis ratio $0.15 < b/a < 1$. Fractions are given in parentheses.

important in determining their apparent inclination than the intrinsic ellipticity of the spheroidal component. According to Maller et al., the completeness of their classification scheme reaches 70% for disk galaxies, with a reliability of $\sim 95\%$ when applied to SDSS data. The completeness is the fraction of galaxies of a given type that are successfully selected from the original sample by the classification scheme. The reliability is the fraction of galaxies of the desired type from the selected subset.

In good agreement with expectations, Table 3.2 shows that the fraction of members of the SDSS-spec sample in low density environments ($\rho_6 \leq 0.5 \, h^3$ galaxies/Mpc$^3$) obeying Maller et al. criterion is very high ($\sim 85\%$ on average), and even higher for the subset of ALFALFA detections ($\sim 92\%$). Our global percentages of late-type objects are, however, slightly lower than those obtained by Maller et al., a result that can be attributed to the fact that they were inferred for a very shallow dataset with $M_r < -20.6$ mag (note that the agreement between our results and Maller et al.’s becomes almost perfect if we restrict the comparison to our lowest velocity bin). In
principle, the decrement in the fraction of galaxies assigned to the ‘disk’ class with increasing radial velocity illustrated in Table 3.2 can be interpreted as an indication of the fact that most luminous galaxies tend to be of early type, in qualitative accordance with previous works (e.g., Strateva et al. 2001; Blanton et al. 2003; Baldry et al. 2004). Yet, we find that in the highest velocity bin \( \sim 10\% \) of the ALFALFA galaxies are not classified as disks, while at these relatively large distances HI detections should arise almost exclusively from late-type objects: at 11,000 km s\(^{-1}\) the minimum HI mass detectable in the survey is \( \log(h^2 M_{\text{HI,min}}/M_\odot) \sim 9.0 \). Accordingly, our results seemingly highlight the difficulties in morphology assignment that selection rules like Maller et al.’s, based on measures of galaxy structure, have for faint objects due to the inherent noisiness of images.

The inverse light concentration index, defined as the ratio of the radii containing 50% and 90% of the Petrosian galaxy light, \( C_{59} = R_{50}/R_{90} \), is another measure of galaxy structure that can be used as a morphological divider (Shimasaku et al. 2001; Strateva et al. 2001). As well as the Sérsic index, this parameter is expected to be tightly correlated with galactic morphology. Indeed, both indexes can be interchanged for an ideal galaxy with a Sérsic profile in the absence of seeing. Early-type galaxies are formed by a dominant bulge, so their light is concentrated, while late-type galaxies have their light more dispersed and, hence, larger \( C_{59} \) values, due to the presence of a disk. In practice, however, the concentration index is a morphological separator useful only to split the galactic population into two big morphological classes: early- (E, S0, Sa) and late-type (Sb, Sc, Irr) objects. We follow Strateva et al. (2001) and adopt an \( C_{59,r} = 0.38 \) separator in the \( r \)-band, which provides a cut giving equally complete subsamples of both subpopulations, while at the same time ensures a highly reliable classification that, in the case of the late-type galaxy class, reduces the contamination by early-type objects to \( \sim 12\% \). When applied to galaxies in LDEs, we find that most galaxies have \( C_{59,r} > 0.38 \), indicating a population dominated by disks. As was to be expected, this tendency is again more marked for ALFALFA detections, about 82% of them fall into the light-concentration-index late-type class, than for the optical subset, which has a late-type galaxy fraction of
3.4. Properties of the H I control samples

Figure 3.9 $(u - r)$ color versus absolute magnitude in the $r$-band for ALFALFA galaxies in LDEs. The dashed curve represents the separator adopted by Baldry et al. (2004).

$\sim 70\%$.

Given the correlation between color and morphologies, early-type galaxies being generally redder than late-types, galaxy colors can also be used as a morphological classification tool, as well as to provide a coarse separation of galaxies in terms of their star formation activity. We adopt the classification scheme by Baldry et al. (2004), which utilizes the SDSS model magnitudes (Stoughton et al. 2002) to measure color from the difference between the fluxes in the $u$ and $r$ filters and to calculate the absolute magnitude in the $r$ band. As well as being an optimal color for separating galaxies into red and blue classes (Strateva et al. 2001), $(u - r)$ is a good indicator of star formation activity (Deng & Zou 2009), while the $r$ filter contains the peak of the light curve for most SDSS galaxies. The absolute magnitude $M_r$ is calculated by comparing the observed apparent magnitude with the CMB distance inferred in Section 3.2.2. In this analysis we do not apply $K$-corrections since the samples investigated are shallow.

Following Baldry et al., who split a galaxy sample into a blue and a red
we show in Figure 3.9 the color-magnitude diagram for the ALFALFA detections in the LDE HI galaxy sample with this delimiter superposed. Not surprisingly, we see quite a few more sources on the blue, active population side of the divider (below the curve) than on the red, passive one: 83% and 17% of the total, respectively (92% and 8% if we correct color for internal extinction and keep the same separator). Moreover, we note that most of the HI galaxies above the separator reside in a locus of the color-magnitude diagram around $M_r - 5 \log h \lesssim -20.5$ mag and $(u - r) \gtrsim 2.3$ where the blue and red galaxy distributions identified by Baldry et al. (2004) overlap in this space (see their Figure 9), so it is not unlikely that many of them can be actually blue distribution interlopers.
A slightly coarser color classification scheme, but one that facilitates the analysis of the data, can be obtained by adopting a luminosity-independent $(u - r)$ color cut of 2.22 mag, as suggested by Strateva et al. (2001). In the four panels of Figure 3.10, we show the color distributions of the SDSS-spec galaxies in LDEs separated into HI-detections and non-detections for different bins of radial heliocentric velocity. As illustrated by the histograms, the global fraction of SDSS galaxies in LDEs that are of red-/early-type becomes accentuated notably with increasing distance, in agreement with the well-known color-luminosity-morphology(-star formation activity) correlation (see also Figure 3.9), which appears to hold regardless of local galaxy density (e.g., Ball et al. 2008). Yet, it is obvious from these plots that the observed behavior arises exclusively from the contribution of galaxies undetected in HI, whose fraction of red objects, according to the Strateva et al. (2001) separator, grows steadily from less than 15% in the lowest velocity bin, to about 39% in the highest one, where $\sim 96\%$ of the SDSS galaxies redder than the divider are HI non-detections. In sharp contrast, the histograms of ALFALFA sources are always clearly dominated by blue-/late-type galaxies and do not appear to participate of the color-luminosity relation, to the point that the ratio between the blue and red fractions for this subset is kept roughly constant to $\sim 1/9$ as a function of distance ($\sim 1/19$ if one uses colors corrected for internal extinction) and, hence, of the optical luminosity of the HI detections. A Kolmogorov-Smirnov test provides objective confirmation that the color distributions of HI detections and non-detections are completely different by yielding an essentially null probability that these two subsamples come from the same parent distribution. Fully consistent results are obtained regardless of whether we restrict this comparison to the subset of SDSS galaxies classified as Maller et al. (2009) disks or if we use the Baldry et al. 2004 separator.

The Galaxy Zoo (GZ1) clean catalog (Lintott et al. 2008), which provides visual morphological classifications for nearly one million SDSS galaxies, offers the possibility of inspecting directly the morphology of a fraction of our HI detections. With this purpose, we have focused our attention on the subset of 621 LDE objects located in the lower of the two surveyed strips
(+4° < Decl.(J2000.0) < +16°) that are also members of the sample of visually classified galaxies by Bamford et al. (2009) (selected from the SDSS DR6 spectroscopic survey with 0.03 < z < 0.085). We have been rather strict assigning morphologies to our ALFALFA objects by requiring that they have a debiased spiral or elliptical likelihood, \(p_{sp}\) or \(p_{el}\), respectively, larger than 0.8. Galaxies that do not pass neither of the above thresholds have been classified as of uncertain type. Comparison with the morphologies listed in the AGC indicates that the reliability of this classification scheme for the spiral class is superior to 90%, with a \(\sim 3\%\) contamination from S0/S0a galaxies, in good agreement with estimates by Lintott et al. (2008) and Bamford et al. (2009). In this manner, we infer that about 75\% of our HI detections can be confidently assigned a spiral (disk with spiral arms) morphology, whereas only \(\sim 1\%\) are galaxies with a high probability of belonging to the elliptical class (E and S0 galaxies). The fact that all the objects in the latter class have \(n > 3\) and \(b/a > 0.45\) (\(r\)-band values) implies that criteria like the one used by Maller et al. (2009) are very effective discarding objects of early-type morphology. Indeed, we find that only 2 of the 478 Maller et al. (2009) disks identified in this subset of GZ1 members have \(p_{el} > 0.8\).

We have gone still a step further and examined in some detail the SDSS images of those LDE galaxies showing optical early-type characteristics (see Appendix A). We have concentrated our efforts on objects with extreme inclinations that meet the following criteria: \((u−r) > 2.22\) mag, \(n > 4.0\), and \(b/a < 0.5\) or \(b/a > 0.85\) (the latter two parameters measured in the \(r\)-band). In agreement with Maller et al. (2009), we find that most of our edge-on red concentrated HI sources have a clear 'disky' morphology suggesting that they owe their early-type classification to large amounts of dust reddening. This morphology is also the dominant one among the twenty some face-on gas-rich galaxies identified with large Sérsic indexes and red colors, in which there is plentiful supply of luminous red central regions and large bars. It seems therefore that the best way to determine the basic morphology of a galaxy is through the presence of significant HI emission.
3.4. Properties of the HI control samples

Figure 3.11 \((u - r)\) color distribution for SDSS galaxies obeying the spectro-photometric (top) and photometric (bottom) isolation criteria (see text). The global histograms are split into HI detections (dashed blue) and non-detections (red).

3.4.2 Structure, colors, and morphology of the isolated galaxies

The exercises done in the previous section for ALFALFA’s detections and non detections in LDEs are repeated now for the two samples of isolated HI galaxies.

As illustrated by Figure 3.11, we find that, regardless of the isolation criterion adopted, the population of HI detections is notably bluer than that of non-detections, in good agreement with our results for galaxies in LDEs. Nevertheless, it is also noticeable that the two subsets of isolated galaxies have, regardless of their HI content, color distributions somewhat redder than those in the LDEs galaxy sample, a tendency that is especially evident in the nearly unimodal histograms of the spectro-photometric subset. Likewise, we find that IG1 galaxies tend to have both smaller light concentration indexes and a smaller fraction of Maller et al. (2009) disks than galaxies in low density regions, as would be expected from a population more dominated by early-type objects. Indeed, if we restrict our attention to just the ALFALFA galaxies with Hubble types listed in the AGC, we find that the LDE HI galaxy sample is considerably richer in Scd and Sd spiral subtypes than the subsets of isolated objects.

We believe, however, that this weird behavior does not stem from funda-
3. Selection of a control sample

mental differences between the properties of isolated and non-isolated galaxies, but from the adoption of isolation criteria that put limits on the relative brightness that the neighbors of a given galaxy can have. Therefore, we caution the reader that selection techniques relying on magnitude differences will inevitably lead to choose the isolated galaxies among the brightest objects in a given catalog and, hence, also among the reddest/earliest ones; the shallower the dataset the most notorious the bias (remember that the objects in our spectro-photometric isolated galaxy sample have $r \leq 15.27$ mag). This conclusion is consistent, for instance, with the results by Allam et al. (2005) themselves, who obtained an isolated galaxy sample from the SDSS DR1 catalog with a light concentration index distribution that suggested a fifty-fifty morphological composition between early and late galaxy types. Another consequence of constraining the range of apparent magnitudes is that it directly translates into a reduction in the dynamic range subtended by the intrinsic properties of the selected galaxies.

3.5 Summary and conclusions

We have used measurements in the 21-cm emission line from the ALFALFA blind survey in a region of the sky also scanned by the SDSS DR7 to carry out an investigation of the physical properties of gas-rich galaxies. The goal has been twofold: 1) to improve the understanding of the nature of these extragalactic HI sources; and 2) to define a reference sample suitable to set up standards for the neutral gas content of galaxies.

Different parametric estimates of the environment of HI sources have been explored, towards establishing their suitability in providing a subset of gaseous galaxies whose properties are minimally affected by external influences. Among the environmental measures tested, the local number density estimate based on the distance to the 6th-nearest spectroscopic SDSS neighbor of a galaxy has emerged as the most appropriate. By applying this estimator to the 11,239 HI detections identified in the overlapping region of the radio and optical surveys, we have extracted a sample of 5496 HI sources
in low density environments. We have also defined two different samples of isolated galaxies by applying standard isolation criteria based either on the combination of spectroscopic and photometric information or solely on photometric data. In spite of the marked differences in size and membership between the LDE and isolated subsets, their respective H I content distributions have proven to be all pretty similar.

In combination with SDSS data, we have also investigated the distributions of light concentration indexes, colors, and other proxies of morphology for these three subsets. All the evidence analyzed indicate that H I emission is detected essentially in objects that are structurally similar to late-type galaxies. Thus, in LDEs, more than 90% of H I sources obey the disk selection rule by Maller et al. (2009), while some 82 per cent have inverse light concentration indexes representative of this class. These fractions could be considered actually a lower limit, as the determination of the structural parameters of distant objects is quite imprecise. Likewise, from evidence gathered by studying the color distributions, we have found that at least between 85% and 90% of the ALFALFA detections outside overdense regions are classified as blue, star-forming galaxies. These fractions are likely to be even higher when extinction-corrected colors are used. In contrast, the structural properties and colors of H I non-detections belong to a population substantially richer in early-type objects. The assignment of morphologies to a subset of ALFALFA galaxies in LDEs using GZ1 type likelihoods above a certain threshold corroborates these conclusions, as only a very small fraction (∼1 per cent) of the objects investigated have been classified into the elliptical class with high reliability.

The view that red galaxies and, consequently, passive objects of type S0/Sa or earlier, tend to have less neutral hydrogen than bluer, star-forming late-type systems, is confirmed and observed to steepen with increasing distance, whenever a flux limitation is present. This is due to the positive correlation between global color and intrinsic luminosity. Our investigation, however, has revealed that this latter relationship involves exclusively galaxies undetected in H I, whereas the color distribution for the gas-rich galaxies is essentially distance independent, due to the lack of a significant red
3. Selection of a control sample

component. We have also found that such a connection between color and luminosity can have an important impact on isolation criteria, favoring the selection of the redder/earlier members of a catalog. Awareness of this bias, together with the relatively modest sizes and reduced dynamic range subtended by the properties of the galaxies in the isolated subsets, indicates that the LDE approach may yield a more suitable and statistically sounder control sample. Paper II presents a detailed derivation of HI standards by applying strategies of non-parametric multivariate data analysis to the galaxies in the aforementioned dataset.

Finally, we wish to remark that only a small fraction of our blindly selected HI sources in LDEs show optical characteristics typical of early-type objects. Careful inspection of the SDSS images has revealed that they are indeed mostly edge-on disks affected by large amounts of dust reddening. This leads to the natural conclusion that the best way to discriminate between early and late-type objects is through the presence of significant HI emission. Yet, we have also found a bunch of face-on galaxies for which the simultaneous coincidence of red colors, large Sérsic indexes, and high HI contents appears to reflect their true condition. Offhand, and on the basis of the evidences reported in Appendix A, one would be tempted to divide these exceptional objects into two main classes according to their origin: 1) elliptical or early-type disk galaxies that, in spite of being in LDEs, have just experienced a merger with a gas-rich companion, as proposed by di Serego Alighieri et al. (2007) for some of their ETGs; and 2) spiral galaxies, initially very rich in HI, in which the atomic gas has been efficiently removed from the innermost portion of their disks (e.g., driven inwards by a strong bar that bias the Sérsic index fit towards large values), but that still retain a fair amount of cold gas in their outer regions. In the latter case, these galaxies with red central disks would come to represent an stage of spiral evolution not as advanced as that assigned to the population of red spirals recently studied by Masters et al. (2010), since ours are still quite rich in neutral gas (several of them have \( h^2 M_{\text{HI}} > 5 \times 10^9 M_\odot \)) and showing signs of ongoing star formation. In any event, the origin of these non-conventional HI sources is something that deserves further investigation.
In this chapter, we deal with a homogeneous subset of 3029 objects extracted from the sample of HI-healthy galaxies in environments of low local galactic density assembled in Chapter 3 obeying the condition $F_{\text{HI}} \geq 1 \text{ Jy km s}^{-1}$, for which we know that ALFALFA is complete. After a careful selection of those parameters best describing the optical and HI properties of these objects, we show that the galactic stellar luminosity, size, and rotation speed are the intrinsic galactic properties most closely related to the HI mass (Section 4.2). Next, we apply strategies of non-parametric multivariate data analysis to explore the correlation structure of those variables (Section 4.3.1), to establish standards of normalcy for the HI content of galaxies (Section 4.3.2), and to constrain the fundamental scaling laws in which these extensive parameters participate (Section 4.3.3).

The distance-dependent intrinsic quantities used in our investigation have been calculated assuming $q_0 = 0$ in the computation of the luminosity dis-
stances from the radial velocity of the sources in the cosmic microwave background rest frame.

4.1 Introduction

While the literature abounds with attempts of improving our knowledge about galaxy formation and evolution from the combination of gaseous and stellar information (e.g., Gavazzi, Pierini, & Boselli 1996; Rosenberg, Schneider, & Posson-Brown 2005; Garcia-Appadoo, West, Dalcanton, Cortese, & Disney 2009, to name a few representative examples), the definition of standards of reference for the HI gas abundance in galaxies has received relatively less attention. Early comparisons of the neutral hydrogen content of Virgo cluster galaxies with field objects put the emphasis on using distance-independent measures, such as the $M_{\text{HI}}/L$ and $M_{\text{HI}}/D^2$ ratios (e.g., Davies & Lewis 1973; Chamaraux et al. 1980b), $L$ and $D$ being some optical luminosity and intrinsic linear diameter, which were not always free of systematic biases. Haynes & Giovanelli (1984) were the first both to carry out an objective evaluation of the performance of different diagnostic tools for the HI content and to provide a rigorous operational definition of this quantity. By making use of a control sample of 288 galaxies with 21-cm line emission contained in the Catalogue of Isolated Galaxies (CIG, Karachentseva 1973), these authors demonstrated that, for a given Hubble type, the optical linear diameter is the most accurate diagnostic tool for the HI mass of galaxies. New expressions for their standards of HI content were later derived in an unbiased way by Solanes, Giovanelli, & Haynes (1996) using a larger, HI flux-limited sample of 532 galaxies from the Catalog of Galaxies and Clusters of Galaxies (CGCG, Zwicky et al. 1968) located in the lowest density environments of the Pisces-Perseus supercluster region.

One serious limitation of these studies, made in a time where wide-field, systematic redshift surveys were still in their infancy, is that they had to deal with heterogeneous datasets assembled from incomplete catalogs affected by different and complex selection biases that could jeopardize their findings.
4.2. Selection of galaxy properties

In this chapter, we conduct a detailed multi-wavelength investigation of the global properties of nearby galaxies that, for the first time ever, grows out from a combination of two large and complete databases which have homogeneously mapped the distribution of extragalactic sources over a cosmologically significant volume of the local universe. These are the first major public data release of the Arecibo Legacy Fast ALFA Survey (ALFALFA) blind 21-cm line survey (Giovanelli et al. 2005b), providing HI measurements in a z-space volume in the high Galactic latitude sky of about 9 hr × 16 deg × 18,000 km s\(^{-1}\) in the general Virgo cluster direction, and the spectroscopic subset of the Sloan Digital Sky Survey Data Release Seventh (SDSS DR7; Abazajian et al. 2009) that, together with additional data from the NYU Value-Added Galaxy Catalog (NYU-VAGC; Blanton et al. 2005), are used to complement the radio detections with abundant information on the properties of their stellar counterparts.

The cross-correlation of the radio and optical observations is carried out in Chapter 3. In that work, we place special emphasis on investigating the distributions of the structural parameters, colors, and morphologies of the selected galaxies as a function of their HI emission in an effort to shed more light on the nature of the extragalactic gaseous sources of the local universe. Chapter 3 also details the steps followed to define what we call the Low Density Environment (hereafter LDE) HI galaxy sample, a dataset of 5496 HI emitters located in environments of low local galactic density whose properties are presumably unaffected by external influences. A flux-limited subset of this control sample is used in the present chapter with the aim of both exploring inter-variable correlations and determining the linear combinations of intrinsic properties that best define the HI content of galaxies with healthy amounts of neutral gas.

4.2 Selection of galaxy properties

Central to this investigation is the search of possible relationships between the gaseous content and other global properties of HI emitters. We have
chosen an extensive set of physical parameters that, besides being suitable to characterize gas-rich objects, are inferred from good quality measurements and either publicly available or easy to compute.

Before we go ahead with the investigation of linear correlations between galaxy properties, we would like to comment briefly several aspects of the procedure adopted. In the first place, we want to stress the fact that the large effective ALFALFA search volume of $\sim 5,350,000 \text{ Mpc}^3$, with dimensions in two of the three coordinate directions (R.A. and $cz$) that are considerably larger than the largest structures seen in numerical simulations in a $\Lambda$CDM world model ($\sim 100h^{-1} \text{ Mpc}$), but with a depth not so large that evolution effects are important, will be assumed to represent a fair sample of the universe. Second, it must no be forgotten that the ALFALFA survey is essentially noise-limited and, hence, naturally biased against sources with both low fluxes and large velocity widths. To solve this problem, we will carry out our analysis on a subset of 3029 members of the LDE $\text{H}_1$ galaxy sample whose 21-cm integrated flux densities, $F_{\text{H}_1}$, are larger or equal than 1 Jy km s$^{-1}$, for which we know that the ALFALFA survey is complete (Stierwalt et al. 2009; see also Appendix B). This flux-limited LDE $\text{H}_1$ galaxy sample will be then made representative of the true space density of the objects under investigation by weighting each of its members by the inverse of the maximum volume within which it could have been observed. And third, it is important to realize that the large size of our sample results in very low critical values for the Pearson’s correlation coefficients, $r_P$, on a t-test of significance, i.e., for the elements of the various correlation matrices that will be inferred in the next section. For instance, $r_{P, \text{crit}} \sim 0.047$ for a level of significance of 0.01 on a two-tailed test with 3000 degrees-of-freedom. This implies that variables that ideally should be essentially independent of each other may end up exhibiting a statistically significant linear relationship according to this test, even in the presence of attenuation from measurement errors. For this reason, as well as to avoid working with an excessive number of parameters, we have decided to consider that two given properties are truly connected in practice only if they have a Pearson’s $r$ that, aside of being statistically significant, is also minimally strong ($|r_P| > 0.3$).
In our quest for the more relevant attributes defining the H\textsubscript{I} content, we want to investigate the existence of linear correlations among the ALFALFA measurements and the many intrinsic optical parameters provided by the SDSS DR7 and NYU-VAGC catalogs, which in some cases use different estimates of the same galaxy property (e.g., Petrosian and model magnitudes for luminosity, isophotal diameter and Petrosian radius for galaxy size, and so on). After a first screening of all the parameters available, paying attention to factors such as the dynamical range of the variables and the size of their measurement errors, or the suitability and robustness of the measurements for the galaxies under scrutiny, we have selected the following quantities as the most convenient for our study:

- The two main properties derived from ALFALFA observations: the velocity width of the source H\textsubscript{I} line profile at the 50\% level of the two peaks, \(W_{50}\), and the H\textsubscript{I} mass estimated from the equation: \(M_{\text{H}1} = 2.356 \times 10^5 d^2 F_{\text{H}1}\), where \(F_{\text{H}1}\) is the 21-cm line flux integral expressed in Jy km s\(^{-1}\) and \(d\) is the cosmological distance to the source in \(h^{-1}\) Mpc (see also Section 3.2.1 in the previous chapter). Note that \(W_{50}\) represents here the intrinsic width corrected for redshift broadening, turbulence and inclination (see Springob et al. 2005). Examination of the variation of H\textsubscript{I} surface density with axial ratio for ALFALFA galaxies has lead us to neglect the effects of internal H\textsubscript{I} self-absorption on \(F_{\text{H}1}\).

- Luminosities (and their corresponding absolute magnitudes) in the five SDSS bands from Petrosian apparent magnitudes, which are specially suited for bright, extended objects. Petrosian magnitudes lead to recover almost all the light from late-type galaxies and around 80\% for early types (Blanton et al. 2001). They are not corrected for the seeing effect.

- The 25 mag arcsec\(^{-2}\) isophotal major-axis diameter \(D_{25}\) in the five SDSS bands. Isophotal sizes provide a more continuous measure of the scale of galaxies than the Petrosian radii \(R_{50}\) and \(R_{90}\) — strongly correlated with the first ones—, which have some 'attractor' values due
to the procedure followed to determine them (see, e.g., Appendix A in Strauss et al. 2002). In addition, the minimum velocity cutoff of 2000 km s$^{-1}$ adopted for the LDE HI galaxy sample (see Chapter 3 for the justification) leaves out of this analysis more than half of the mostly nearby, blue faint objects having unrealistically small isophotal radii in the SDSS database.

- Colors from model magnitudes. Here, we explore the combinations $(u-g)$, $(g-r)$, $(r-i)$, $(i-z)$, $(u-r)$, $(u-i)$, $(u-z)$, $(g-i)$, $(g-z)$, and $(r-z)$.

- The Sérsic index, $n$, from the NYU-VAGC catalog, which measures the shape of the observed $r$-band luminosity profile of a galaxy fitted using the Sérsic $R^{1/n}$ formula with elliptical isophotes. Available only for galaxies with $r < 18$ mag.

- The (inverse) index of light concentration $C_{59} = R_{50}/R_{90}$ in the five SDSS bands. Not corrected for seeing.

- The observed axis ratio $b/a$, in the five SDSS bands too, which given its apparent nature should be uncorrelated with any of the other properties. Included only as a control variable.

In the present analysis, the color, as well as the Sérsic and light concentration indexes, are used as objective proxies of morphology. Attempts to work with ‘classical’ indicators of morphological type, such as the continuous de Vaucouleurs numerical code listed in the HyperLeda database, have been thwarted by lack of completeness: about half of the SDSS galaxies do not have information on their Hubble types, while the same is true for $\sim 16\%$ of the ALFALFA sources.

Note also that with the exception of the axis ratio, we are dealing with intrinsic properties that must be inferred from the corresponding observed properties listed in the catalogs. In particular, SDSS absolute magnitudes have been corrected to face-on values following Shao et al. (2007). Similarly to Maller et al. (2009), $D_{25}$ has been corrected for inclination by using
transformations of the form

\[ \log D_{25} = \log D_{25}^{\text{obs}} + \beta \log \left( \frac{b}{a} \right), \]

(4.2.1)

where \( D_{25} \) and \( D_{25}^{\text{obs}} \) are the intrinsic and observed values of this variable in a given band, respectively, whereas the coefficient \( \beta \) measures the strength of the corresponding attenuation. For instance, we obtain \( \beta_r = 0.25 \) in the \( r \)-band, a value not far from the inclination correction estimated for \( R_{50} \) (\( \beta_r = 0.20 \)) by Maller et al. (2009).

In order to examine the correlation structure, we feed in the logarithms of all these basic variables but \( C_{59} \) and \( b/a \). In this manner, we should be able to find any scaling law that might exist among them (this is also a must when variables have a lognormal distribution). This means, in particular, that it is not necessary to include explicitly in the analysis interesting composite parameters, such as the stellar mass given by Gavazzi et al. (2008)\(^1\), \( \log \left( \frac{M_*}{M_\odot} \right) = -0.152 + 0.518 (g - i) + \log L_i \), that are linear combinations of (the logarithm of) some of the input variables. This is also the case for the mean surface brightness, isophotal or Petrosian, \( \bar{\mu} \equiv M + 5 \log R \), which can likewise be expressed as a function of two of the measurements listed above.

The correlation matrices inferred from different subsets of the selected measurements demonstrate that colors and luminosities (absolute magnitudes) are strongly correlated among themselves, as was to be expected. This allows us to discard all but one of the colors and all but one of the luminosities. Among these photometric variables, the ones that show the largest correlation coefficients with the HI mass are the \( g \), \( r \), and \( i \) band luminosities, as well as all the optical colors obtained from combinations of them. Given that the photometric errors in these three bands are rather similar (Strateva et al. 2001), we have taken into account both the dynamical ranges of the different colors and the economy in the number of involved bands to select finally the \( r \)-band luminosity and the \((g - r)\) color as the most adequate representatives of these two fundamental stellar properties.

\(^1\)A similar formula based on the \((g - r)\) color is provided by Bell et al. (2003).
The correlation analysis has also evidenced that the different measurements of both the isophotal diameter and the Sérsic index in the five SDSS bands are degenerated. As before, the superior quality of the SDSS photometry in the central $g$, $r$, and $i$ bands results in stronger correlations of these properties with the H1 mass at these wavelengths. Consistency with the adopted bandpass for measuring the amount of light, has lead us to select the $r$-band estimates of $D_{25}$ and $n$ to represent these quantities as well.

Overall, we find that the isophotal $r$-band diameter, the intrinsic $r$-band luminosity, and the $W_{50}$ linewidth are tightly related to the $M_{\text{HI}}$. A second group of attributes, the $(g - r)$ color and the $r$-band Sérsic index, are only marginally aligned with this quantity ($|r_p| \sim 0.3-0.4$; see Disney et al. 2008 for a similar conclusion regarding color). While we have decided to keep provisionally the color within the set of parameters that may be needed to describe the cold gas content of a galaxy, we have chosen to drop from this list the Sérsic index, as it is available only to galaxies in the NYU-VAGC catalog (i.e., with $r \leq 18$ mag). Finally, the third of the morphological separators investigated, the index of light concentration, shows a negligible linear correlation with the H1 mass, of similar strength ($|r_p| \lesssim 0.15$), for instance, to the correlations involving the apparent inclination, so it will not be taken into account either throughout the remainder of this study.

### 4.3 Relationships among H I and optical properties

We now complete the characterization of gas-rich galaxies by circumscribing ourselves to the subset of five properties that according to the results of the previous section best define the H1 content, i.e., that are more strongly correlated with this quantity.

To begin with, we will investigate the dimensionality of this multivariate space and then proceed to derive relations between the neutral hydrogen content and the remaining properties. We will follow two different approaches in order to achieve this latter goal. On the one hand, we will seek for the most
probable value of $M_{\text{HI}}$ assuming the other properties are precisely known. This corresponds to solve a multivariate regression problem that will provide us with equations useful to establish standards of normalcy for the HI content of galaxies—we shall define as 'normal' the set of HI properties characteristic of our LDE galaxy sample—from a set of diagnostic parameters easily accessible to observation. On the other hand, we will also infer the best-fitting axis of all the different pairs of selected parameters that can be built and the constraints they impose on the scaling relations among the global properties of galaxies.

To reduce the impact of possible outliers on the inferred relationships, which could be substantial especially if they are low HI-mass objects, we have restricted the axis ratio of the sources to be within the range $0.15 < (b/a)_r < 0.9$. This discards $\sim 6$ per cent of the ALFALFA detections in LDE with $F_{\text{HI}} \geq 1 \text{ Jy km s}^{-1}$ having an optical counterpart in the SDSS. We have also removed the handful of objects with missing coordinate values, thus eliminating an additional $\sim 2$ per cent of sources. After this trimming, the galaxy sample to be analyzed in this section contains a total of 2796 HI emitters.

### 4.3.1 PCA results

Here we deal with the correlation matrix of the properties $\log M_{\text{HI}}[h^{-2}M_\odot]$, $\log W_{50}[\text{km s}^{-1}]$, $\log D_{25,r}[h^{-1}\text{kpc}]$, $M_r - 5 \log h[\text{mag}]$, and $(g - r)[\text{mag}]$ to perform a principal component analysis (PCA) in this parameter space. Unlike the covariant matrix, the use of the correlation matrix entails the standardization of the original variables putting them on an equal footing: all samples of variables are get to have zero mean and unit standard deviation. With this scaling of the measurements one avoids the creation of spurious interrelations arising from the preponderance (i.e., larger dynamical range) of certain variables. For the PCA calculations, we employ the IDL procedure *pcapro*\(^2\), implemented following the description by Murtagh & Heck (1987),

\(^2\)http://idlastro.gsfc.nasa.gov/ftp/pro/math/pcapro
which has been slightly modified in order to account for the flux limitation of our sample, as well as for the measurement error of the estimates.

The flux cutoff of the data is compensated by assuming that our observations are distributed over a representative volume of the universe and weighting all galaxies by $1/V_{\text{max}}$, the inverse of the maximum volume in which they could have been observed given the adopted lower limit of 1 Jy km s$^{-1}$. This increases the weight of the low HI-mass objects in the PCA algorithm to compensate for the fact that they are not detected at large distances$^3$. In addition, we have also taken into account that the correlations between variables are weakened in the presence of measurement error. Given two sets of estimates $\hat{X}$ and $\hat{Y}$ with independent measurement errors, the disattenuated correlation between the underlying variables $X$ and $Y$ can be obtained from the formula (cf. Spearman 1904)

$$r_P(X, Y) = \frac{r_P(\hat{X}, \hat{Y})}{\sqrt{R_{XX}R_{YY}}} ,$$  

(4.3.1)

where the reliability coefficients $R_{XX}$ and $R_{YY}$ are defined mathematically as one minus the ratio between the variance on the corresponding measurement error and the total observed variance. In practice, this means that variables are standardized using an estimate of their true, intrinsic standard deviation, instead of the observed one. We have followed Fuller & Hidiroglou (1978) (see also Bock & Petersen 1975) and extended this correction to the multivariate case by imposing the constraint of producing a valid correlation matrix for our disattenuated variables, i.e., a matrix that is at least positive-semidefinite, after checking that their associated measured error scores are essentially independent.

The results of our PCA analysis are summarized in Tables 4.1 and 4.2 in the form of the $1/V_{\text{max}}$-weighted correlation matrix, all its eigenvectors and eigenvalues (i.e., the variances of the data in the directions of the principal axes), as well as the rms residuals between the 5-dimensional space of the observations and the different $p$-dimensional subspaces ($p = 1, ..., 5$) that best

$^3$Our flux-limited dataset represents a volume-limited sample of about 25,000 galaxies.
4.3. Relationships among H I and optical properties

Table 4.1. Principal Component Analysis of H I and Optical Properties for Weighted Data and Disattenuated Correlations

<table>
<thead>
<tr>
<th></th>
<th>( \log M_{H I} )</th>
<th>( \log W_{50} )</th>
<th>( \log D_{25,r} )</th>
<th>( M_r )</th>
<th>( (g - r) )</th>
<th>( C_{59,r} )</th>
<th>( (b/a)_r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \log M_{H I} )</td>
<td>1.00</td>
<td>0.66</td>
<td>0.79</td>
<td>-0.72</td>
<td>0.51</td>
<td>-0.18</td>
<td>-0.01</td>
</tr>
<tr>
<td>( \log W_{50} )</td>
<td>1.00</td>
<td>0.66</td>
<td>-0.71</td>
<td>0.77</td>
<td>-0.32</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>( \log D_{25,r} )</td>
<td>1.00</td>
<td>-0.97</td>
<td>0.77</td>
<td>-0.20</td>
<td>-0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( M_r )</td>
<td>1.00</td>
<td>-0.81</td>
<td>0.30</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( (g - r) )</td>
<td>1.00</td>
<td>-0.19</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( C_{59,r} )</td>
<td>1.00</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( (b/a)_r )</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (( N_{\text{gal}} = 2796 ))</td>
<td>8.93</td>
<td>2.23</td>
<td>0.53</td>
<td>-17.83</td>
<td>0.29</td>
<td>0.44</td>
<td>0.53</td>
</tr>
<tr>
<td>Intrinsic std. dev.</td>
<td>0.42</td>
<td>0.24</td>
<td>0.25</td>
<td>1.56</td>
<td>0.15</td>
<td>0.05</td>
<td>0.18</td>
</tr>
<tr>
<td>Observational error</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.11</td>
<td>0.11</td>
<td>0.01</td>
<td>0.09</td>
</tr>
<tr>
<td>( \sqrt{R_{XX}} )</td>
<td>1.00</td>
<td>0.99</td>
<td>0.98</td>
<td>1.00</td>
<td>0.80</td>
<td>0.98</td>
<td>0.91</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvectors</th>
<th>Eigenvalues (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{e}_1 )</td>
<td>0.41</td>
</tr>
<tr>
<td>( \hat{e}_2 )</td>
<td>0.69</td>
</tr>
<tr>
<td>( \hat{e}_3 )</td>
<td>-0.38</td>
</tr>
<tr>
<td>( \hat{e}_4 )</td>
<td>-0.42</td>
</tr>
<tr>
<td>( \hat{e}_5 )</td>
<td>-0.19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>rms residuals:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Axis</td>
</tr>
<tr>
<td>Principal Plane</td>
</tr>
<tr>
<td>Princ. 3-Plane</td>
</tr>
<tr>
<td>Princ. 4-Plane</td>
</tr>
<tr>
<td>Princ. 5-Plane</td>
</tr>
</tbody>
</table>

Note. — PCA is carried on the first 5 properties only: \( \log(M_{H I}[h^{-2}M_{\odot}]) \), \( \log(W_{50}[km \ s^{-1}]) \), \( \log(D_{25,r}[h^{-1}kpc]) \), \( M_r - 5 \log(h[\text{mag}]) \), and \( (g - r)[\text{mag}] \)

describe them. Table 4.1, which presents the results for the disattenuated correlation matrix, also lists the adopted estimates of the typical measurement error and the resulting square root of the reliability coefficient for each one of the five selected observables. In general, the parameters investigated have a near perfect reliability, except the color, for which the attenuation correction is significant.

The correlation matrices indicate the existence of large linear correlations (\( |r_P| \gtrsim 0.6 \)) for all pairs of parameters except for H I mass and color, that exhibit a correlation coefficient of medium size (\( |r_P| \sim 0.4 \)). This is specially true for the disattenuated matrix, which shows that the five physical quantities investigated are strongly correlated with the first principal component.
The latter is endowed with direction cosines of nearly equal absolute value (approximately $1/\sqrt{5}$) and has an associated eigenvalue $\lambda_1$ of 4.0, out of a maximum possible of 5.0, implying that about 80% of the total variance in the adopted 5-parameter space can be explained by this principal axis. The second principal component draws mostly from $M_{\text{HI}}$ and $(g-r)$, and accounts for an additional $\sim 12\%$ of the global variance, while a three degree-of-freedom manifold is enough to explain the practical totality ($\sim 99\%$) of it (this reduces somewhat to $\sim 93.5\%$ if we do not use disattenuated correlation estimates; see Table 4.2). The principal 4-plane brings the rms residuals down to the level of the adopted observational errors with the exception of $M_r$, whose variance cannot be fully accounted for. The latter might indicate either that there are still hidden parameters, like perhaps the current star formation rate, controlling the structure of disk galaxies, or that the observational error adopted for this variable has been too optimistic. Note also that the correlation matrix reveals a very strong negative relationship between the $r$-band magnitude and isophotal diameter, which is responsible for the nearly null variance attached to the last eigenvector.

Our result regarding that gas-rich galaxies form essentially a single parameter family (i.e., one principal component describes most of the variance of the five H I and optical properties investigated) is consistent with the findings of earlier PCA-based studies that have attempted to elucidate the degree of organization shown by the physical properties of disk-galaxies from a similar set of observational parameters (e.g. Brosche 1973; Bujarrabal et al. 1981; Conselice 2006), as well as in very good overall agreement with the recent work by Disney et al. (2008) from HIPASS and SDSS data. Regarding the possibility, suggested by these latter authors, that that H1-selected galaxies have colors made up of two components, one systematic, weakly correlated with the other variables and the single principal component, and a so-called rogue component, which is only aligned with itself, we reach a different conclusion. As stated above, we have found that the second principal component, which like in Disney et al. (2008) explains about 12–13% of the variance, is preferentially aligned with $M_{\text{HI}}$ and $(g-r)$. When PCA is carried on the attenuated $1/V_{\text{max}}$-weighted correlation matrix, the direction
4.3. Relationships among H I and optical properties

Table 4.2. Principal Component Analysis of H I and Optical Properties for Weighted Data

<table>
<thead>
<tr>
<th></th>
<th>(\log M_{\text{HI}})</th>
<th>(\log W_{50})</th>
<th>(\log D_{25,r})</th>
<th>(M_r)</th>
<th>(g - r)</th>
<th>(C_{59,r})</th>
<th>((b/a)_r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\log M_{\text{HI}})</td>
<td>1.00</td>
<td>0.65</td>
<td>0.74</td>
<td>−0.71</td>
<td>0.41</td>
<td>−0.17</td>
<td>−0.00</td>
</tr>
<tr>
<td>(\log W_{50})</td>
<td>1.00</td>
<td>0.62</td>
<td>−0.70</td>
<td>0.62</td>
<td>−0.31</td>
<td>−0.02</td>
<td></td>
</tr>
<tr>
<td>(\log D_{25,r})</td>
<td>1.00</td>
<td>0.93</td>
<td>0.55</td>
<td>−0.23</td>
<td>−0.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M_r)</td>
<td>1.00</td>
<td>−0.65</td>
<td>0.29</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_{59,r})</td>
<td>1.00</td>
<td>0.62</td>
<td>−0.70</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>((b/a)_r)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean (\(N_{\text{gal}} = 2796\))

|                | 8.93 | 2.23 | 0.60 | −17.83 | 0.29 | 0.44 | 0.53 |

Std. dev.

|                | 0.42 | 0.24 | 0.27 | 1.56 | 0.18 | 0.06 | 0.20 |

Observational error

|                | 0.03 | 0.04 | 0.05 | 0.11 | 0.11 | 0.01 | 0.09 |

√\(\mathbf{X}^T\mathbf{X}\)

|                | 1.00 | 0.99 | 0.98 | 1.00 | 0.80 | 0.98 | 0.91 |

Eigenvalues (%)

| \(\hat{\varepsilon}_1\) | 0.45 | 0.44 | 0.48 | −0.49 | 0.39 | \(\lambda_1 = 3.65\) (72.99) |
| \(\hat{\varepsilon}_2\) | 0.51 | −0.19 | 0.28 | −0.08 | −0.79 | \(\lambda_2 = 0.63\) (85.51) |
| \(\hat{\varepsilon}_3\) | −0.33 | −0.72 | 0.46 | −0.36 | 0.16 | \(\lambda_3 = 0.42\) (93.96) |
| \(\hat{\varepsilon}_4\) | −0.66 | 0.49 | 0.22 | −0.28 | −0.44 | \(\lambda_4 = 0.24\) (98.80) |
| \(\hat{\varepsilon}_5\) | 0.07 | −0.10 | −0.66 | −0.74 | −0.09 | \(\lambda_5 = 0.06\) (100.00) |

rms residuals:

|                | 0.24 | 0.13 | 0.11 | 0.52 | 0.12 |
| Principal Axis | 0.17 | 0.13 | 0.10 | 0.51 | 0.04 |
| Principal Plane | 0.14 | 0.06 | 0.05 | 0.35 | 0.04 |
| Princ. 3-Plane | 0.01 | 0.01 | 0.04 | 0.28 | 0.00 |
| Princ. 4-Plane | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Note. — PCA is carried only on the first 5 properties: \(\log(M_{\text{HI}}[h^{-2}\text{M}_\odot]), \log(W_{50}[\text{km s}^{-1}]), \log(D_{25,r}[h^{-1}\text{kpc}]), M_r - 5 \log h[\text{mag}]\), and \((g - r)[\text{mag}]\)

The cosine of the second principal axis relative to the color is certainly the one with the largest (absolute) value (0.75), but after correcting the data from observational error this contribution significantly weakens and is no longer the dominant one (compare the corresponding eigenvectors in Tables 4.2 and 4.1, respectively). This leads us to conclude that a substantial part of the secondary independent contribution to the variance found by Disney et al. (2008) does not correspond to a true second degree of freedom related to the color of galaxies, but to the large observational scatter (compared to the intrinsic one) affecting this variable.
4.3.2 Standards of H I content

As stated in the former section, we are dealing with five observables that are strongly correlated. In this situation, any multiple regression model which attempts to describe the H I mass in terms of interrelated predictor variables (multicollinearity) would be associated with an ill-conditioned correlation matrix, i.e., a matrix whose inversion is numerically unstable (if there were one or more exact linear relationships among the variables the matrix would not be invertible). Thus, in the presence of multicollinearity the impact of the individual predictors on the response variable tends to be less precise than if the predictors were uncorrelated with one another. To remedy this problem, we have adopted a two-stage procedure known as Principal Component Regression (PCR; e.g. Cook 2007) that first carries out a PCA of all the predictor variables, and then uses the resulting principal components (which are independent, and hence associated with a correlation matrix of full rank) and the dependent variable in an ordinary least squares regression fit. Besides, one can take advantage of the initial PCA transformation to reduce the dimensionality of the data by keeping only those new variables most correlated with the H I mass (e.g. Jolliffe 1982).

We have applied the above procedure to different subsets of parameters, starting by finding the regression relations between the $M_{\text{H I}}$ and each one of the four remaining properties (see the plots above the diagonal in Figure 4.1), and then adding input variables progressively to seek for the combinations of regressors that best predict the H I content. This process stops when after adding a new predictor variable the rms residual of the multiple regression model increases or does not get reduced in an amount comparable or larger than the observational error in $M_{\text{H I}}$ (Table 4.2).

Not surprisingly, given the unidimensionality shown by the intrinsic parameter space of gaseous galaxies (Section 4.3.1), we find that the best predictions are those depending on a single regressor variable. Table 4.3 lists, ordered according to decreasing accuracy as given by the size of the rms residuals, the central values and associated 90% bootstrap confidence intervals of
the coefficients of the correlations

\[
\log \left( \frac{M_{\text{HI}}}{h^{-2}M_{\odot}} \right) = a + b \ X
\]  

(4.3.2)

for fits both with and without \(1/V_{\text{max}}\)-weighting carried on disattenuated data. It can be seen that the most precise predictor of the HI mass is \(D_{25,r}\), which is precisely the property most strongly correlated with it, followed by \(M_r\). Among the distance-independent observables, the best predictor of the HI content is the rotational width of the disk\(^4\), whereas the regression model using the \((g - r)\) color is the least accurate, as expected.

We have also verified that combinations of two or more distance-dependent input variables not only do not improve the prediction of the \(M_{\text{HI}}\), i.e., do not they diminish the value of the rms residuals, but do not lead to correlations less affected by distance uncertainties either. For illustrative purposes, we report in Table 4.3 the predictions obtained using the optical size and luminosity as diagnostic variables (cf. Haynes & Giovanelli 1984), which obey equations of the form

\[
\log \left( \frac{M_{\text{HI}}}{h^{-2}M_{\odot}} \right) = a + b \ X_1 + c \ X_2 .
\]  

(4.3.3)

Looking at the coefficients of these multiple regressions, it is clear that the right-hand side of the equations shows a dependence with distance that neither is null (the ratio \(b/c \neq -5\)), as would be expected if they were defining a surface magnitude, nor does it compensate the \(d^2\)-dependence of the HI mass.

4.3.3 Planar correlation diagrams

In the former section, we have been interested in deriving predictions for one global property, the HI mass, given a value for other four extensive parame-
Table 4.3. Coefficients of $M_{H_1}$ Predictions from Single and Multiple Linear Regression Models

<table>
<thead>
<tr>
<th>Weighting</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>log $D_{25,r}$</td>
<td>8.23$^{+0.06}_{-0.06}$</td>
<td>1.32$^{+0.10}_{-0.11}$</td>
<td></td>
<td></td>
<td>0.27</td>
</tr>
<tr>
<td>1/$V_{\text{max}}$</td>
<td>Mr</td>
<td>5.45$^{+0.44}_{-0.38}$</td>
<td>0.19$^{+0.02}_{-0.02}$</td>
<td></td>
<td></td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>log $W_{50}$</td>
<td>6.33$^{+0.29}_{-0.29}$</td>
<td>1.16$^{+0.13}_{-0.11}$</td>
<td></td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>$(g - r)$</td>
<td>8.49$^{+0.15}_{-0.15}$</td>
<td>1.50$^{+0.52}_{-0.51}$</td>
<td></td>
<td></td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>log $D_{25,r}$</td>
<td>Mr</td>
<td>6.74$^{+0.23}_{-0.20}$</td>
<td>0.64$^{+0.06}_{-0.06}$</td>
<td>$-0.10^{+0.01}_{-0.01}$</td>
<td>0.28</td>
</tr>
<tr>
<td>None</td>
<td>log $D_{25,r}$</td>
<td>8.41$^{+0.03}_{-0.03}$</td>
<td>1.35$^{+0.04}_{-0.03}$</td>
<td></td>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>Mr</td>
<td>5.74$^{+0.11}_{-0.11}$</td>
<td>0.19$^{+0.01}_{-0.01}$</td>
<td></td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>log $W_{50}$</td>
<td>7.00$^{+0.17}_{-0.17}$</td>
<td>1.03$^{+0.07}_{-0.07}$</td>
<td></td>
<td></td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>$(g - r)$</td>
<td>9.07$^{+0.04}_{-0.04}$</td>
<td>1.09$^{+0.10}_{-0.10}$</td>
<td></td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>log $D_{25,r}$</td>
<td>Mr</td>
<td>6.98$^{+0.07}_{-0.07}$</td>
<td>0.65$^{+0.02}_{-0.02}$</td>
<td>$-0.10^{+0.003}_{-0.003}$</td>
<td>0.22</td>
</tr>
</tbody>
</table>
4.3. Relationships among H\textsc{i} and optical properties

eters: size, total magnitude, velocity width, and color. In this section, we turn our attention to the constraints that these five intrinsic variables put on the scaling relations among the global properties of spiral galaxies. This means that we now consider all physical variables on an equal footing and focus on the correlations arising from the same PCA technique used in Section 4.3.1 applied to pairs of these variables. The involved quantities are therefore treated symmetrically, which minimizes the inconsistencies that may result from the possible ‘non-commutativity’ of the inferred relationships. The medians and their associated 90% bootstrap confidence intervals of the $1/V_{\text{max}}$-weighted and unweighted orthogonal fits of the form $P_1 = a + b P_2$ between all 10 possible pairs of global galaxy properties $P_1$ and $P_2$ are presented in Table 4.4 for the disattenuated data, while the relations and their best

Figure 4.1: Empirical relations for pairs of properties. $1/V_{\text{max}}$-weighted (solid) and unweighted (dotted) direct regression fits to the joint distributions are shown above the diagonal, whereas orthogonal fits are shown below it. All correlations are corrected for attenuation (Equation 4.3.1).
linear fits can be visualized in the plots below the diagonal in Figure 4.1. Note that the strongest corrections associated with the flux limitation and the error disattenuation of the data occur in those correlations involving the color.

The study of scaling relations among the fundamental extensive properties of disks is central for constraining theories of galaxy formation and evolution. It has generated an abundant literature, whose detailed revision far exceeds the scope of the present work. Instead, we have decided to focus on the comparison of the values inferred here for the mean slopes of the fundamental correlations in the LRV space of luminosity or mass, size, and rotation speed, and those reported in similar studies that combine HI and optical observations (Haynes & Giovanelli 1984; Salpeter & Hoffman 1996), or that specifically investigate the scaling relations in spiral galaxies (Courteau et al. 2007). In comparing results allowance should be made not just for differences in sample size, but also in the waveband of the optical observations, the observables chosen to estimate the selected fundamental properties, the fitting method employed, and the dynamic ranges of the variables explored. Another complication that distorts the comparison of results is the statistical incompleteness of the databases used in all mentioned studies, which, with the exception of ours, are affected by intractable selection biases.

Given these various caveats in mind, we find that the agreement between the central values of the slopes (restricted to two significant digits) reported by the different studies listed in Table 4.5 is, in general, satisfactory. The largest discrepancies correspond to the correlation involving the luminosity versus the rotation speed, i.e., the Tully-Fisher (TF) relation, reflecting the fact that it is always problematic to find the slope of this empirical law accurately due to the low dynamic range in log $V$. The mean values of the log slope $\gamma$ listed in Table 4.5 range from $\sim 2.6$ (Haynes & Giovanelli 1984) and this work ($-6.50$ in magnitudes), to $\sim 3.4$ (Courteau et al. 2007) and $\sim 3.7$ (Salpeter & Hoffman 1996) (the first and the last ones being measured at blue wavelengths and the Courteau et al.’s for the I-band). At this point, it is important to realize that our dataset includes sources with a very broad range of axis-ratios and, hence, its good share of nearly face-on galaxies for
which the inclination corrections to the observed HI linewidths tend to be overestimated. As a result, we get a flatter slope than those typically inferred from samples specifically designed for TF studies that encompass a narrower range of inclinations. Indeed, by pruning away the galaxies with $i < 60^\circ$ that participate in our fits, we obtain a joint distribution with considerably less scatter than those represented in the corresponding panels of Figure 4.1 that, for the weighted data, has a slope of $-7.54^{+1.1}_{-0.86}$ in magnitude units. This result is fully consistent with the estimates reported in the subject-specific literature for bright spirals in the blue/near-IR band (e.g. Willick et al. 1997; Giovanelli et al. 1997b; Courteau 1997; Masters et al. 2006).

We note in passing that, since we are dealing with a noise-limited sample, this pruning has also the effect of discarding most of the slowest rotators with $W_{50} \lesssim 100$ km s$^{-1}$. From the expression of the signal-to-noise ratio $S/N$ of the ALFALFA detections (cf. Martin, Giovanelli, Haynes, Saintonge, Hoffman, Kent, & Stierwalt 2009)

\[ S/N = \frac{1000 F_{HI}}{W_{50}} \frac{w_{\text{smo}}^{1/2}}{\sigma_{\text{rms}}} , \]

(4.3.4)

where $W_{50}^{\text{obs}}$ is the observed velocity width of the source line profile corrected for instrumental broadening, $w_{\text{smo}}$ a smoothing width, and $\sigma_{\text{rms}}$ the rms noise figure measured in mJy at 10 km s$^{-1}$ resolution, its is straightforward to show that, by adopting the approximations $w_{\text{smo}} \approx W_{50}^{\text{obs}}/20$ and $\sigma_{\text{rms}} \approx 2$ mJy, the distribution of intrinsic velocity widths for galaxies with a minimum inclination of $60^\circ$ verifies

\[ W_{50} \gtrsim \left[ \frac{120}{(S/N)} \right]^{2} . \]

(4.3.5)

The orthogonal fits in which $W_{50}$ does not participate do not depend on the range of inclinations (or, equivalently, axis ratios) adopted.

Another result that deserves to be commented is the inferred relationship between the HI mass and the size of the stellar distribution of ALFALFA galaxies, represented in this work by their isophotal diameter in the $r$-band, $D_{25,r}$. Although the discrepancies with respect to the other estimates of this
## Table 4.4. Coefficients of Orthogonal Fits between pairs of Fundamental Properties

<table>
<thead>
<tr>
<th>Weighting</th>
<th>( P_1 )</th>
<th>( \log D_{25,r} )</th>
<th>( M_r )</th>
<th>( \log W_{50} )</th>
<th>( (g - r) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( a )</td>
<td>( b )</td>
<td>( a )</td>
<td>( b )</td>
</tr>
<tr>
<td>1/( V_{\text{max}} )</td>
<td>( \log M_{\text{HI}} )</td>
<td>8.05(^{+0.07}_{-0.08})</td>
<td>1.68(^{+0.09}_{-0.08})</td>
<td>4.08(^{+0.23}_{-0.24})</td>
<td>0.27(^{+0.01}_{-0.01})</td>
</tr>
<tr>
<td></td>
<td>( \log D_{25,r} )</td>
<td>( M_r )</td>
<td>( \log D_{25,r} )</td>
<td>( M_r )</td>
<td>( \log W_{50} )</td>
</tr>
<tr>
<td></td>
<td>( M_r )</td>
<td>( \log W_{50} )</td>
<td>( \log D_{25,r} )</td>
<td>( M_r )</td>
<td>( \log W_{50} )</td>
</tr>
<tr>
<td></td>
<td>( \log W_{50} )</td>
<td>( M_r )</td>
<td>( \log D_{25,r} )</td>
<td>( M_r )</td>
<td>( \log W_{50} )</td>
</tr>
<tr>
<td>None</td>
<td>( \log M_{\text{HI}} )</td>
<td>8.15(^{+0.03}_{-0.03})</td>
<td>1.68(^{+0.03}_{-0.04})</td>
<td>4.39(^{+0.11}_{-0.13})</td>
<td>0.26(^{+0.00}_{-0.01})</td>
</tr>
<tr>
<td></td>
<td>( \log D_{25,r} )</td>
<td>( M_r )</td>
<td>( \log D_{25,r} )</td>
<td>( M_r )</td>
<td>( \log W_{50} )</td>
</tr>
<tr>
<td></td>
<td>( M_r )</td>
<td>( \log W_{50} )</td>
<td>( \log D_{25,r} )</td>
<td>( M_r )</td>
<td>( \log W_{50} )</td>
</tr>
<tr>
<td></td>
<td>( \log W_{50} )</td>
<td>( M_r )</td>
<td>( \log D_{25,r} )</td>
<td>( M_r )</td>
<td>( \log W_{50} )</td>
</tr>
</tbody>
</table>
Table 4.5. Comparison of the Central Slopes of Fundamental Scaling Laws Reported by Different Authors

<table>
<thead>
<tr>
<th>Reference</th>
<th>$M_{\text{HI}} \sim R^\alpha$</th>
<th>$M_{\text{HI}} \sim L^\beta$</th>
<th>$L \sim V^\gamma$</th>
<th>$R \sim L^\delta$</th>
<th>$R \sim V^\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haynes &amp; Giovanelli (1984)</td>
<td>1.8</td>
<td>0.66</td>
<td>2.6</td>
<td>1.4</td>
<td>1.1</td>
</tr>
<tr>
<td>Salpeter &amp; Hoffman (1996)</td>
<td>2.0</td>
<td>0.74</td>
<td>3.7</td>
<td>0.37</td>
<td>1.1</td>
</tr>
<tr>
<td>Courteau et al. (2007)</td>
<td>…</td>
<td>3.4</td>
<td>3.4</td>
<td>0.32</td>
<td>1.1</td>
</tr>
<tr>
<td>This work</td>
<td>1.7</td>
<td>0.68</td>
<td>2.6–3.0 ($i \geq 60^\circ$)</td>
<td>0.40</td>
<td>1.1</td>
</tr>
</tbody>
</table>
4. Multivariate analysis of a flux-limited dataset

Quantity are relatively minor, it is obvious that our finding that $\alpha = 1.66^{+0.09}_{-0.07}$ does not support the idea that all the HⅠ-rich galaxies have the same global HⅠ surface density, as recently advocated by Garcia-Appadoo et al. (2009) from a sample of HIPASS galaxies and implied by correlations like the one found by Salpeter & Hoffman (1996) from optical size measurements in the $B$-band\textsuperscript{5}. As stated in Solanes et al. (1996), the use of datasets largely dominated by Sc galaxies, which are more prone to exhibit a nearly constant global HⅠ surface density, may explain the tendency to find such an aesthetically appealing conclusion in samples that, morphologically speaking, are not truly representative of the whole spiral population. In this respect, we want to stress that, in the case of Garcia-Appadoo et al. (2009), who deal with a sample representative of the whole range of galaxies between giant spirals and extreme dwarfs, the claimed constancy of the mean HⅠ column density emanates from a correlation between the HⅠ mass and the $g$-band radius $R_{50}$ with an actual central slope of $\sim 1.7$, fully consistent with ours.

4.4 Summary and conclusions

We have sought for correlations between a large number of independent extensive 21-cm and optical properties of non-clustered, gas-healthy galaxies with the aim of identifying the combinations of intrinsic variables that make up the best diagnostic tools for their HⅠ content. The observations on which this study is based are provided by a complete 21-cm flux-limited subset of near 3000 galaxies that are members of the Low Density Environment HⅠ galaxy sample of the ALFALFA survey defined in Chapter 3.

The examination of the correlation structure of the many HⅠ and optical measurements available for the objects under investigation has shown that the isophotal $r$-band diameter, the total $r$-band luminosity, and the $W_{50}$ linewidth are the physical parameters showing the tightest correlations with the total neutral gas mass. On the other hand, we have found that this

\textsuperscript{5}Implicit in this conclusion is the assumption that HⅠ and stellar disk sizes are roughly proportional as shown by (Broeils & Rhee 1997).
latter quantity is at most only marginally aligned with crude but quantitative Hubble-type indicators like the \((g − r)\) color or the Sérsic index.

The principal component analysis of the space defined by the variables, \(\log M_{\text{HI}}, \log W_{50}, \log D_{25,r},\) and \(M_r\), as well as the \((g − r)\) color, has demonstrated that gas-healthy galaxies behave essentially as a one-dimensional family (i.e., a single principal component describes most of the variance associated with these five independent properties). This finding is in very good agreement with the conclusions from earlier investigations of the degree of organization of the physical properties of disk galaxies and from recent studies that, like ours, combine 21-cm and optical data, such as the one carried out by Disney et al. (2008). Regarding the possibility, indicated by the latter authors, that the color might constitute a weak, albeit nevertheless significant, second principal component, our results, based on a dataset about 14 times larger, suggest that this apparent second degree of freedom has not a physical origin, but it is attributable to the large relative measurement error of this observable.

In accordance with the principal component analysis results, we have also found that the best predictions for the most probable value of the \(\text{HI}\) mass of a galaxy, assuming that the other properties are precisely known, are those depending on a single regressor variable. Not surprisingly, we find that the most precise predictor of \(M_{\text{HI}}\) is \(D_{25,r}\), the parameter most strongly correlated with it, followed in decreasing order of accuracy by \(M_r, W_{50}\), and the \((g − r)\) color. More specifically, our analysis shows that the neutral gas content of galaxies free of environmental influences is essentially independent of their color (and of other morphological proxies like the Sérsic index), since this latter kind of predictors leads to rms residuals between models and data roughly equal to the total observed standard deviation associated with \(M_{\text{HI}}\). This confirms that galaxy morphology plays a secondary role in the determination of the neutral gas content, as already inferred by Haynes & Giovanelli (1984) and Solanes et al. (1996). On the other hand, the fact that the tightest correlations observed show only a mild dependence on distance indicates that distance uncertainties should not be increasing significantly their scatter.
We have also explored the joint distributions between all possible pairs of the independent observables mentioned above and the constraints they impose on the most firmly established empirical scaling laws among the global properties of disk galaxies. From orthogonal linear regressions that deal with logarithmic quantities, we find the following median $r$-band relations

$$M_{\text{HI}} \propto R^{1.7}, \quad M_{\text{HI}} \propto L^{0.68}, \quad L \propto V^{2.6}, \quad R \propto L^{0.40}, \quad R \propto V^{1.1},$$

(4.4.1)

where $L$ is the total luminosity from the (old) stellar disk, and $R$ and $V$ are its size and rotation speed, respectively. The fact that we have obtained a log slope of $2.6^{+0.27}_{-0.36}$ (90% confidence interval) for the empirical $LV$ relation, which is marginally inconsistent with the typical values reported in the TF literature at optical/near-IR wavelengths, is not surprising, but simply the consequence of not imposing any trimming on the range of inclinations included in our dataset. When we do so and prune away disks with $i < 60^\circ$, the central value of the estimated slope raises to 3.0, in full agreement with the expectations. Another remarkable consequence than can be drawn from the inferred scalings and associated confidence intervals is that the ratio between $M_{\text{HI}}$ and optical radius squared decreases slightly with galaxy size. As stated by Solanes et al. (1996), who also arrived at the same conclusion, claims in favor of the near universality of the hybrid global HI surface density for the entire spiral population typically rely on datasets predominantly dominated by galaxies of the latest Hubble types (mostly Sc and Irr) for which the constancy of this intensive property would be a relatively reasonable approximation.

Multidimensional statistical studies of the integral attributes of galaxies have had to contend to date with heterogeneous samples of modest size, largely incomplete, and affected by important selection effects that cannot be corrected. It is therefore obvious that disregarding any of these factors when they are in fact present may result in inconsistent estimation. In this respect, efforts like the present one based on the cross-correlation of wide-area, objective databases at different wavelengths and with controlled selection biases should mark the way forward in the study of galaxy scaling relations.
Conclusions and future perspectives

This thesis is devoted to the study of 21-cm line detections of galaxies in two markedly different density environments. In the first part, we carry out an individual and detailed analysis of the HI distribution and kinematics of several gas-poor objects in the high-galactic density region of the Virgo Cluster. In the second part, we deal with the statistical analysis of thousands of galaxies, with the aim of selecting gas-healthy objects which can constitute a reference sample for the HI content of galaxies. Both parts, in spite of their different aims, have the overall goal of improving our understanding on the nature of extragalactic HI sources. In the following paragraphs, we summarize the main conclusions we reached in each part and highlight some aspects that should deserve further investigation.

Our VLA observations of gas-poor objects in the Virgo Cluster Region were aimed at disentangling the origin of their HI-deficiency. The selected objects were apparently located, according to Tully-Fisher estimates, at the cluster background, a location that played against a galaxy-cluster interaction as the cause behind their HI deficiency. For the strongest HI-deficient objects, the obvious truncated gas disks they show are compatible with a previous ISM-ICM interaction, albeit the VLA data by themselves are not conclusive on the mechanism(s) that could have caused the deficiency. Two of our main targets, which had a mild HI deficiency according to previous
single-dish measurements, have shown normal HI gas contents according to our observations. In all cases, the performance of Tully-Fisher when used to derive distances to our targets has been analyzed, concluding that the uncertainties are much larger than previously assumed: HI truncated disks are certainly not the best targets for such studies (whether due to the inefficient estimate of the dynamical mass from rotation curves that do not exhibit a flat part, or because the likely affected luminosities of such systems); neither are the two almost face-on spirals, whose rotation speeds carry large uncertainties from inclination corrections. As a result, the large uncertainties in the location along the line of sight prevent us from unambiguously determining whether these galaxies belong to the main body of the Virgo cluster or to the infall region.

A possible way to gain more insight in the question of how far the impact of the cluster environment is felt by the late-type population of galaxies, is to estimate the timescales in which our HI-deficient objects have suffered gas ablation. This can be done by means of a multi-frequency study that collects the necessary information to fit the star formation histories of our VLA sources (e.g., Gavazzi et al. 2002). This study could be extended to those HI outliers in the sample of Sanchis et al. (2004) with confirmed gas deficiency by Chung et al. (2009), as well as other objects showing either a mild deficiency or being gas-healthy for comparison. Essentially, the idea is to constrain as accurately as possible the time at which the star formation was quenched in the outer disks, in a manner similar to what Crowl & Kenney did for about ten stripped spirals in Virgo (Crowl & Kenney 2006, 2008).

In the second part of the thesis has been focused on the derivation of standards of neutral gas content of galaxies from the combination of ALFALFA detections and SDSS data. Chapter 3 has taken care of the definition of a suitable control sample of galaxies. A reference sample, consisting on more than five thousand objects selected for lying in Low Density Environments (LDEs), has been compared with two smaller datasets gathered by applying different isolation criteria. The properties of the three control samples have been inspected in order to understand the nature of these extragalactic HI sources: large fractions (∼85-90%) of our atomic gas detections in
low density environments are characterized by having blue colors, low Sérsic index, and spiral GZ1 morphologies, whereas ALFALFA non-detections in the same environments show larger fractions of objects structurally similar to early-type galaxies. Our results confirm that blue, star forming late-type systems tend to have larger H\textsubscript{I} content than red, passive objects of S0/Sa or earlier morphologies. This trend is emphasized with increasing distance, whenever an effective flux limitation is present: the color distribution of the gas-healthy, blue population detected by ALFALFA is essentially luminosity independent, whereas the color distribution of galaxies not detected in H\textsubscript{I} changes according to the well-known color-luminosity correlation.

As detailed in Appendix A, we have also paid attention to the small subset of galaxies in LDEs that, although having a relatively gas-healthy status, show optical early-type characteristics. Their visual inspection has revealed that most of them are edge-on disk objects affected by dust reddening. However, we have also identified some twenty face-on objects that simultaneously show red colors, large Sérsic indexes, and high H\textsubscript{I} content, and that, given these puzzling properties, deserve further investigation.

Further improvement on the definition of our H\textsubscript{I} control sample could come from:

- expanding local density measurements to regions not surveyed by the SDSS, for instance by cross-correlating ALFALFA with the Two Micron All Sky Survey (2MASS). That would allow one to obtain a control sample of H\textsubscript{I} detections from the whole ALFALFA footprint. The subsequent multivariate analysis of this sample would provide new standards of H\textsubscript{I} content from the 2MASS $JHK_s$-bands. While the cross-correlation of these two surveys is hampered by the lack of measured redshifts for 2MASS sources out of the SDSS footprint (e.g., Blanton et al. 2005, for a cross-correlation of SDSS and 2MASS), the latter difficulty will be soon overcome thanks to the ongoing 2MASS Redshift Survey (2MRS, Huchra et al. 2005a,b), which has already obtained redshifts for 2MASS sources up to $K=$11.25 mag.
• Using the most recent catalogs of isolated galaxies, such as the 2MASS-selected Isolated Galaxies catalog Karachentseva et al. (2009), with about 3000 objects, or the sample of extremely isolated galaxies in 3D redshift space selected by Spector & Brosch (2009).

The modest sizes of the isolated samples and the potential biases affecting the optical properties of their members, have lent support to the use of the LDE sample in the subsequent multivariate analysis, from which we have selected a flux-limited subset of about three thousand galaxies to carry out a multivariate analysis of their H I and optical properties (Chapter 4). With the aim of seeking for the combinations of intrinsic properties that make the best diagnostic tool for the H I content of galaxies. The correlation structure of all the parameters has been inspected by means of a PCA, finding that the tightest correlations with neutral gas mass correspond, in decreasing order, to the $r$-band isophotal diameter, $D_{25,r}$, the $r$-band absolute magnitude, $M_r$, and the $W_{50}$ linewidth, whereas proxies for morphology such as $(g-r)$ color and Sérsic index show mild correlations with HI mass. This analysis has also shown, in good agreement with the most recent study on the subject carried out by Disney et al. (2008) from HIPASS data, that all the analyzed properties form, essentially, a one-dimensional space. The search for the best predictors for the H I mass of a galaxy from optical parameters has revealed that the lowest residuals are obtained by using a single regressor variable, being the $D_{25,r}$ diameter the most accurate predictor. This is in accordance with the previous determinations of the standards of H I normalcy in terms of the optical diameter established by Haynes & Giovanelli (1984) and Solanes et al. (1996), as well as with their conclusion on the secondary role played by morphology.

The PCA technique has also been used to explore the orthogonal linear regressions between all possible pairs of the investigated parameters, finding that the inferred slopes are in reasonable agreement with the empirical scaling relations that have been established to date among the global properties of disk galaxies (luminosity, size, rotation speed and HI mass).

As it has already been mentioned, the multivariate study of the flux-
limited sample we have just summarized could be expanded to the near-infrared properties of galaxies by analyzing 2MASS data. Apart from that, there are some other aspects that might be worth the trouble to address in the near future, such as imposing a weighting scheme in the PCA analysis that explicitly takes into account that ALFALFA is essentially a noise limited dataset. Note that our multivariate analysis has been carried on a subsample of galaxies in LDE with $F_{\text{H}1} \geq 1 \text{ Jy km s}^{-1}$, for which ALFALFA behaves as a flux-limited sample (see Appendix B). Therefore, our weighting scheme is based on the assumption that, for a galaxy of a certain H I mass, its probability of detection by ALFALFA is proportional to the volume defined by the adopted flux limit (Equation 1.1.4). Improved weights for our galaxies from the noise-limit approach could soon be obtained from the ongoing study of the H I mass function from ALFALFA data by Martin et al. (2010, in preparation).

As for the future research that could benefit from the new standards of H I content we have obtained, environmental studies on the H I properties of galaxies would be in first place. In this respect, it is important to mention the ongoing Arecibo Galaxy Environment Survey (AGES, Auld et al. 2006), that will map thirteen areas of the sky with large integration times and covering the full range of environments (from the Local Void through isolated galaxies, galaxy pairs and galaxy groups, to the Virgo Cluster).

Finally, one of the most obvious things to do is to reevaluate the H I-deficiency of the spirals in the Virgo region that have been analyzed in Part I, as well as of the rest of outliers in the sample selected by Sanchis et al. (2004), with respect to the new standards of H I content derived from ALFALFA. It must be taken into account that the standards of H I normalcy that were used to determine the H I deficiency of these galaxies were based on samples of optically-selected targets with non-homogeneous H I observations, and that applied statistical techniques different from ours. It is therefore a must to check if there are variations in their status of H I-deficient objects when one uses the new standards derived in this thesis.
5. Conclusions and future perspectives
In search of red, concentrated, gas-healthy galaxies in Low Density Environments

The fact that we have detected small but non-negligible fractions of HI sources, which are assumed to be late-type, showing either structural characteristics, or colors, or both typical of earlier objects deserves further attention. The obvious explanation would be that the adopted classification schemes are highly but not one-hundred percent reliable. Yet it is worth the trouble to try to find out whether there are truly early-type galaxies in our dataset that, nonetheless, are also gas-rich. For this reason, we have searched the LDE HI galaxy sample for galaxies that meet the following criteria: \((u - r) > 2.22\) mag\(^1\), \(n > 4.0\), and \(b/a < 0.5\) or \(b/a > 0.85\) (the latter two parameters measured in the \(r\)-band). In agreement with Maller et al. (2009), visual inspection of the SDSS images shows that the red concentrated galaxies in our dataset that are highly-inclined have a clear ‘disky’ morphology. Since, in addition, most of these objects are intrinsically blue, this strongly suggests that we are essentially witnessing gas-healthy spirals that owe their apparent red color to large amounts of dust reddening.

Some twenty galaxies in our LDE HI sample are face-on 21-cm sources

\(^1\)The selection of the galaxies is hardly affected by the adopted divider and the use of SDSS Petrosian or model magnitudes.
with early-type optical characteristics. Among them we have been able to identify at least one elliptical-looking galaxy that appears to have a dust lane (AGC 191407; see the top SDSS image in Figure A.1), which belongs to the minority of galaxies of this kind that have amounts of cold gas ($M_{\text{HI}} = 8.79 \log h^{-2} M_\odot$) comparable to those found in large spirals. Nevertheless, the most frequent situation corresponds to relatively HI-rich sources ($M_{\text{HI}} \gtrsim 9.5 \log h^{-2} M_\odot$) with an unmistakable disk morphology, but with big red central regions, harboring bright bulges and strong bars that dominate the light emission biasing the color and the Sérsic index towards large values (in a few cases, the reddening gets accentuated by contamination by starlight from a nearby star). These galaxies tend to have extended, very blue, but comparatively faint disks concentrating most of the star formation and where HI gas is likely plentiful. In some cases the disks of these galaxies are so faint that they are barely perceptible in the optical images, which explains why these sources are frequently assigned a lenticular or Sa morphology (middle pictures in Figure A.1). On the other hand, when the disks are visible it is not uncommon to find evidences of peculiar structures such as outer shells or inner rings (bottom pictures in Figure A.1), as well as occasionally spiral arms showing signs of possible tidal disturbances. Possible scenarios for the origin of these non-conventional objects are discussed in Section 3.5.
Figure A.1 SDSS images of red \((u-r) > 2.22\), concentrated \((n > 4)\), and face-on \((b/a > 0.85)\) objects from our LDE H I galaxy sample. Top: Nearby \((v_{\text{hel}} = 3641 \text{ km s}^{-1})\) E-like galaxy with a dust lane and a relatively important gaseous content \((\log M_{\text{HI}} = 8.79)\). Middle left: Spiral galaxy with very faint blue arms and disk that are not accounted for in the calculation of its global color made in the SDSS. Middle right: Gas-rich disk galaxy \((\log M_{\text{HI}} = 9.61)\) lacking spiral arms and showing a large red central region. Bottom: Two examples of Sa galaxies with strong red bars and rings, as well as blue outer disks, that are rather rich in neutral gas (the Sa on the left has \(\log M_{\text{HI}} = 9.73\), whereas the one on the right has \(\log M_{\text{HI}} = 9.46\)).
We display in this Appendix the distribution of galaxies as a function of the integrated 21-cm-line flux in Jy km s$^{-1}$ is provided to test the differential completeness of the data gathered by the ALFALFA survey, i.e., to search for possible variations with sample depth of the fraction of H I sources recovered with respect to the expectations from a survey of limited sensitivity. This is done by comparing the observed number counts in bins of $F_{\text{HI}}$ with the relation expected for a flux-limited, uniform distribution of sources, $dN_{\text{gal}}(F_{\text{HI}}) \propto F_{\text{HI}}^{-5/2} dF_{\text{HI}}$, scaled vertically so as to fit the right-hand side of the histogram. As shown in the top panel of Figure B.1, the deviation from the theoretical curve becomes apparent at $\log F_{\text{HI}} \lesssim 0$. Consistent results are obtained in the bottom panel of this figure, where we apply this test in integral form by comparing the cumulative number counts in $m_{\text{HI}} = -2.5 \log F_{\text{HI}}$ with the predicted linear growth $\log N_{\text{gal}}(\leq m_{\text{HI}}) = 0.6 m_{\text{HI}} + \text{const.}$ also depicted in the insert is the cumulative distribution of the SDSS DR7 spectroscopic targets that fall within the ALFALFA region before (filled symbols) and after (empty symbols) removing objects with velocities above 18,000 km s$^{-1}$. Both forms of this classical test that, it must not be forgotten, is sensitive to the effects of both large-scale structure and subsampling in redshift bins, suggest a completeness limit of 1 Jy km s$^{-1}$ ($m_{\text{HI}} = 0$ mag). This is an effective limiting value of $F_{\text{HI}}$ for the entire ALFALFA catalog that can be adopted independently of the observed velocity width and S/N.
Figure B.1 Top: ALFALFA source counts against integrated H\textsc{i} flux. The dashed curve represents \( N_{\text{gal}}(F_{\text{H}\textsc{i}}) \propto F_{\text{H}\textsc{i}}^{-5/2} \), as expected for a homogeneous flux-limited survey. Bottom: Cumulative distribution as a function of the H\textsc{i} magnitude, \( m_{\text{H}\textsc{i}} = -2.5 \log F_{\text{H}\textsc{i}} \), for galaxies in the ALFALFA catalog. The dashed line is the theoretical prediction expected for a flux-limited, uniform distribution of sources. The insert in the right shows the same Hubble test applied to the subset of \( r \)-band SDSS DR7 spectroscopic observations located in the region of the sky surveyed by ALFALFA both before (solid squares) and after (open squares) removing galaxies with velocities above 18,000 km s\(^{-1}\).

of sources. A more thorough investigation of the completeness of ALFALFA observations will be presented elsewhere (Martin et al. 2010, in preparation).
Bibliography


Bock, R. D., & Petersen, A. C. 1975, Biometrika, 62, 673


Boselli, A., & Gavazzi, G. 2006, Publ. of the ASP, 118, 517


Briggs, D. S. 1995, PhD thesis, New Mexico Institute of Mining and Technology, New Mexico, Mexico


Cook, R. D. 2007, Statistical Science, 22, 1


Deng, X., & Zou, S. 2009, Astroparticle Physics, 32, 129


Ekholm T., Lanoix, P., Teerikorpi,


Freedman, W. L., Madore, B. F., Gibson, B. K., Ferrarese, L., Kelson, D. D.,

Fisher, K. B., Huchra, J. P., Strauss, M. A., Davis, M., Yahil, A., & Schlegel,


Fujita, Y. 2004, Publ. of the ASJ, 56, 29

Association, 73, 99

Fumagalli, M., Krumholz, M. R., Prochaska, J. X., Gavazzi, G., & Boselli,

Garcia-Appadoo, D. A., West, A. A., Dalcanton, J. J., Cortese, L., & Disney,

Gavazzi, G., Bonfanti, C., Sanvito, G., Boselli, A., & Scodeglio, M. 2002,

Gavazzi, G., Boselli, A., Donati, A., Franzetti, P., & Scodeglio, M. 2003,
ArXiv Astrophysics e-prints

439


Hamilton & Tegmark 2004, MNRAS 349, 115-128


Jacoby, G. H., Branch, D., Ciardullo, R., Davies, R. L., Harris, W. E. et al. 1992, Publ. of the ASP, 104, 599


Karachentseva, V. E. 1973, Soobshcheniya Spetsial’noj Astrofizicheskoj Observatorii, 8, 3


BIBLIOGRAPHY


Spearman, C. 1904, American Journal of Psychology, 15, 72

Spector, O., & Brosch, N. 2009, ArXiv e-prints


Warmels, R. H. 1986, HI properties of spiral galaxies in the Virgo cluster, ed. Warmels, R. H.


