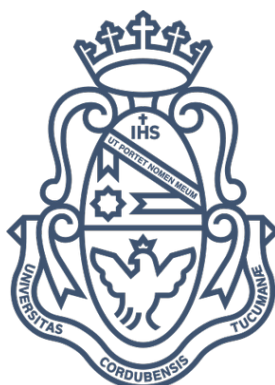


Universidad Nacional de Córdoba  
Facultad de  
Matemática, Astronomía, Física y Computación

## **Identificación de Grupos Compactos**

María Antonela Taverna



Presentado ante la Facultad de  
Matemática, Astronomía, Física y Computación  
como parte de los requerimientos para la obtención del grado de  
Doctora en Astronomía de la  
UNIVERSIDAD NACIONAL DE CÓRDOBA

©FaMAF - UNC

Director: Dr. Ariel Zandivarez

- Marzo 2020 -

Identificación de Grupos Compactos by Maria Antonela Taverna se distribuye  
bajo una Licencia Attribution-NonCommercial-NoDerivatives 4.0  
International.





## Resumen

En esta tesis doctoral se han adoptado como objeto de estudio los grupos compactos (CGs) de galaxias. Estos sistemas contienen galaxias en estrecha proximidad en una región muy limitada del espacio y se consideran aislados. Debido a su peculiar naturaleza, estos sistemas son de gran importancia para los estudios de formación y evolución de galaxias, ya que es un ambiente favorable para la interacción entre sus miembros sin influencia del entorno en gran escala.

A lo largo de esta tesis realizamos estudios estadísticos sobre CGs, los cuales fueron identificados principalmente en simulaciones numéricas combinadas con diferentes modelos semianalíticos de formación de galaxias, y también en catálogos observacionales. El objetivo de este trabajo fue tratar de aportar nueva información sobre temáticas relacionadas con la identificación de los CGs.

En primer lugar, se abordó el problema de la comparación de diferentes catálogos de CGs que recurrentemente surge en la literatura, sobre todo cuando las muestras surgen de la aplicación de diferentes algoritmos. Para ello nos concentramos en analizar el impacto de la banda fotométrica utilizada en la identificación de CGs utilizando un único algoritmo. Utilizamos un catálogo sintético de galaxias con evolución, e identificamos CGs en las bandas  $r$ ,  $K$  y  $u$ . Encontramos que algunas propiedades de los CGs cambian cuando los sistemas son identificados en una u otra banda. Además, existen CGs que pueden ser identificados en las 3 bandas al mismo tiempo, mientras que otros CGs sólo pueden ser identificados en una única banda. Esto implica que las muestras de CGs —y sus propiedades— dependen de la banda fotométrica utilizada para la identificación.

En segundo lugar, el estudio realizado se enfocó en el mejoramiento de los algoritmos de identificación de CGs en catálogos observacionales. El algoritmo propuesto respeta las restricciones establecidas por el reconocido criterio visual de Hickson para la identificación de CGs, pero adaptado para su automatización y para optimizar la selección de dichos sistemas. El nuevo algoritmo identifica aproximadamente el doble de CGs identificados que cuando se usa el algoritmo automático clásico que replica los pasos secuenciales de la inspección visual realizada por Hickson. El desarrollo de un nuevo algoritmo nos permitió aplicarlo al catálogo observacional de galaxias Sloan Digital Sky Survey Data Release 12. A partir de esta muestra identificamos 462 grupos compactos, siendo el catálogo con mayor número, hasta el momento, de grupos compactos identificados con información espectroscópica completa para todos sus miembros.

En tercer lugar, intentamos aportar nuevas pistas acerca de la naturaleza y frecuencia de ocurrencia de CGs. Este tema es de gran importancia y ha sido controversial a lo largo de los años. En esta oportunidad, realizamos un

estudio a gran escala de este tema analizando un gran número de muestras de CGs identificadas en catálogos sintéticos contruidos a partir de diferentes modelos semianalíticos de formación de galaxias (SAMs) y simulaciones numéricas con variados parámetros cosmológicos. Encontramos que la frecuencia y la naturaleza de los CGs depende en gran medida de los parámetros cosmológicos. Las diferencias intrínsecas en las recetas de los SAM también conducen a diferencias en la frecuencia y la naturaleza de los CGs, particularmente aquellas relacionadas con la forma en la que los SAMs tratan a sus galaxias huérfanas (galaxias limitadas por la resolución de las simulaciones utilizadas).

Finalmente, analizamos la pureza y la completitud de los catálogos de CGs identificados en el espacio de redshifts. Para eso, definimos una muestra de CGs identificados en el espacio real 3D que tuviesen las características esperadas de estos sistemas: galaxias muy próximas entre sí, tamaños pequeños y aislados. Esta muestra fue trasladada al espacio distorsionado de un catálogo sintético limitado en flujo y comparada con la muestra de CGs que se obtiene al aplicar el criterio de Hickson convencional. Nuestros resultados muestran que la muestra de CGs tipo Hickson es muy poco eficiente para recuperar CGs definidos en 3D, alcanzando una pureza del  $\sim 15\%$ . El resto de los sistemas que identifica el criterio de Hickson, que son considerados espurios, son sistemas que en su mayoría pueden estar asociados con sistemas físicamente densos en 3D pero que no cumplieron con el criterio de aislamiento.

### **Clasificación:**

95.35.+d Materia oscura (estelar, interestelar, galáctica y cosmológica).

98.62.Ve Estudios estadísticos y de las correlaciones de las propiedades.

98.65.-r Grupos de galaxias, cúmulos y supercúmulos; estructura a gran escala del Universo.

95.75.-z Observación y técnicas de reducción de datos; modelado y simulaciones computacionales.

98.65.Bv Grupos pequeños y compactos de galaxias.

**Palabras Clave:** Grupos Compactos de Galaxias - Análisis Estadísticos - Criterios de Identificación - Simulaciones cosmológicas - Catálogos de Galaxias



## Abstract

In this doctoral thesis, we studied the compact groups of galaxies (CGs). These groups are isolated systems that contain a few galaxies in an area very small. Due to their peculiar nature, these systems are very important to the study of formation and evolution of galaxies because they live in a favourable ambient for the interaction among their members.

We realised statistical studies about compact groups of galaxies which were identified in numerical simulations combined with different semi-analytic models of galaxy formation (SAM), as well as in observational catalogues. The objective of this work was to try to contribute new information on issues related to the identification of CGs.

In the first place, we compare different catalogues of CGs that recurrently appear in the literature, especially when the samples arise from the application of different algorithms. To do it, we study the impact of the photometric band used in the identification of CGs using a unique algorithm. We used a mock catalogue of galaxies and identify CGs in r, K, and u bands. Since we identify using the same catalogue, the selection function is maintained. We found that some properties of CGs change when identifying groups in a band or another. Besides, there are CGs that can be identified in a single band (r-pure, K-pure, u-pure) and other groups that can be identified in the three bands. This implies that CGs identified -and their properties- depend strongly on the band used for the identification.

In the second place, we improve the algorithm of identification of CGs using the Hickson criteria but with an optimisation to apply it in redshifts space. The new algorithm was tested in simulated catalogues build from the numerical simulation Millenium I and a semi-analytical model of galaxy formation. This improved algorithm identifies approximately twice of CGs than we use the classic algorithm. This allows us to implement the new algorithm to a observational catalogue of galaxies Sloan Digital Sky Survey Data Release 12. From this sample, we identify 462 compact groups. It is the most numerous catalogue of CGS identifies with full spectroscopic information until now.

In the third place, using the new algorithm of identification, we realised a comparative study to understand how the frequency and three-dimensional nature of CGs depend on semi-analytic modelling of galaxies and on cosmological parameters. We analyse nine lightcones of galaxies build from five different SAM, and we identify CGs like in the observational samples. We found the the frequency and the nature of CGs depend strongly on cosmological parameters. The intrinsic differences in the recipes of the SAM also lead to a differences in the frequency and nature of CGs, particularly those related to the way SAMs treat orphan galaxies.

Finally, we analyse the purity and completeness of the catalogues of CGs identified in redshifts space. To do this, we defined a sample of CGs identified in

real 3D space that had the expected characteristics of these systems: galaxies very close to each other, small and isolated sizes. This sample was moved to the distorted space of a mock catalogue limited in flux and compared with the sample of CGs obtained by applying the conventional Hickson criterion. Our results show that the Hickson-type CGs sample is very inefficient to recover defined 3D CGs, reaching a purity of  $\sim 15\%$ . The rest of the systems that identify the Hickson criterion, which are considered spurious, are systems that for the most part can be associated with systems physically dense in 3D but did not meet the isolation criteria.

**Clasification:**

95.35.+d Dark matter (stellar, interstellar, galactic, and cosmological).

98.62.Ve Statistical and correlative studies of properties.

98.65.-r Galaxy groups, clusters, and superclusters; large scale structure of the Universe.

95.75.-z Observation and data reduction techniques; computer modeling and simulation.

98.65.Bv Small and compact galaxy groups.

**Keys words:** Compact Groups of Galaxies - Statistical Analysis - Criteria of Identification - Cosmological Simulations - Galaxy Catalogues





# Índice general

<b>Resumen</b>	<b>I</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Introducción General . . . . .	1
1.2. Búsqueda Automática de CGs . . . . .	3
1.2.1. Criterio tipo Hickson . . . . .	5
1.2.2. Criterio tipo FoF . . . . .	6
1.3. Comparaciones entre catálogos y propiedades de CGs . . . . .	7
1.4. Estudios de CGs a partir de catálogos sintéticos . . . . .	9
1.5. Planificación del trabajo . . . . .	10
<b>2. Catálogos sintéticos</b>	<b>11</b>
2.1. Simulaciones numéricas . . . . .	11
2.2. Modelos semianalíticos de formación de galaxias . . . . .	13
2.3. Catálogos sintéticos en el espacio distorsionado . . . . .	22

<b>3. Identificación de CGs en diferentes bandas fotométricas</b>	<b>25</b>
3.1. Construcción de las muestras de CGs . . . . .	26
3.1.1. Catálogos sintéticos . . . . .	26
3.1.2. Identificación de CGs . . . . .	28
3.1.3. Propiedades de los CGs . . . . .	30
3.1.4. Muestras Restringidas . . . . .	36
3.2. Grupos compactos Puros y Comunes . . . . .	37
3.3. Resumen . . . . .	41
<b>4. Optimización del algoritmo de identificación de CGs</b>	<b>43</b>
4.1. Construcción del nuevo algoritmo para la identificación de CGs .	44
4.1.1. El algoritmo clásico . . . . .	45
4.1.2. El algoritmo modificado . . . . .	45
4.1.3. Comparación entre el algoritmo clásico y modificado . . . .	46
4.2. Aplicación del algoritmo modificado al SDSS . . . . .	49
4.2.1. El SDSS . . . . .	50
4.2.2. Identificación de CGs en el SDSS . . . . .	50
4.2.3. Comparación con otros catálogos de CGs seleccionados a partir del SDSS . . . . .	56
4.3. Resumen . . . . .	62

<b>5. Naturaleza y frecuencia de CGs en diferentes SAMs</b>	<b>64</b>
5.1. Los catálogos lightcones . . . . .	64
5.2. Muestras de CGs sintéticos . . . . .	67
5.3. Frecuencia y naturaleza de CGs . . . . .	70
5.3.1. Dependencia con parámetros cosmológicos . . . . .	73
5.3.2. Dependencia con la fracción de galaxias huérfanas . . . . .	76
5.3.3. Dependencias con otros parámetros . . . . .	77
5.4. Resumen . . . . .	81
<b>6. Naturaleza 3D de los CGs</b>	<b>82</b>
6.1. Criterio 3D para identificar CGs . . . . .	82
6.1.1. Compacticidad . . . . .	83
6.1.2. Aislamiento . . . . .	84
6.2. CGs en las simulaciones en espacio real . . . . .	88
6.3. ¿Puede el algoritmo de Hickson recuperar los CGs densos en 3D?	92
6.3.1. CGs 3D en el espacio de redshifts . . . . .	93
6.3.2. Identificación de CGs tipo Hickson en el espacio de redshifts	94
6.3.3. Pureza y Completitud de CGs . . . . .	98
6.3.4. Análisis de grupos espurios . . . . .	103
6.3.5. <i>Reals</i> vs CG-3D . . . . .	107
6.4. Resumen . . . . .	111

<b>7. Conclusiones</b>	<b>113</b>
7.1. Comparación de CGs identificados en diferentes bandas fotométricas . . . . .	113
7.2. Optimización del algoritmo de identificación de CGs . . . . .	115
7.3. Naturaleza y frecuencia de CGs . . . . .	117
7.4. CGs en el espacio 3D . . . . .	118
7.5. Discusión Final . . . . .	120
<b>A. Descarga de datos para los diferentes SAMs</b>	<b>122</b>
<b>B. Cálculo de propiedades de CGs</b>	<b>125</b>
<b>C. Tablas de galaxias del catálogo SDSS</b>	<b>128</b>
<b>Bibliografía</b>	<b>129</b>

# Capítulo 1

## Introducción

### 1.1. Introducción General

Los Grupos Compactos (CGs) son sistemas de galaxias muy estudiados en astronomía extragaláctica. Son sistemas pequeños aislados localmente de no más de 10 miembros donde sus galaxias se encuentran muy próximas entre si. Esto genera que los CGs sean sistemas sobre-densos y propensos a interacciones entre sus galaxias miembros. La alta tasa de fusiones propone un escenario perfecto para el estudio de la formación y evolución de estos objetos a partir de sus propiedades observacionales como también de sus miembros.

El primer CG observado fue el “Quinteto de Stephan” en 1877 (Stephan, 1877). Fue observado en proyección en el plano del cielo. Años posteriores se pudo medir el corrimiento al rojo (redshift) de las galaxias miembro, y descubrir que éste grupo es en realidad un grupo de 4 galaxias y el quinto miembro fue considerado como parte del grupo sólo por un efecto de proyección dada la falta de información espectroscópica. Mucho tiempo después, en 1948, Seyfert (1948) observó el segundo grupo compacto, también en proyección, y es conocido como el “Sexteto de Seyfert”, pero este grupo también tiene una galaxia con redshift discordante. (Ver figura 1.1)

Posteriormente, Shakhbazyan en 1957 observó una asociación de objetos rojos muy compacta, sistema que años mas tarde fue confirmado como grupo compacto por Robinson & Wampler (1973). Con la observación de éstos dos objetos, surge la iniciativa de construir catálogos observacionales de este tipo de sistemas para comenzar a estudiar su naturaleza. Durante varios años, en un trabajo en colaboración, Shakhbazyan (1973); Shakhbazyan & Petrosyan (1974), Baier et al. (1974), Petrosian (1974), Petrosyan (1978), Baier & Tiersch



**Figura 1.1:** Panel izquierdo: Quinteto de Stephan . Panel derecho: Sexteto de Seyfert. Ambas imágenes fueron tomadas por el Telescopio Espacial Hubble.

(1975), Baier & Tiersch (1976); Bayer & Tiersch (1976), Baier & Tiersch (1978) y Baier & Tiersch (1979); construyeron una lista de grupos compactos de galaxias observados hasta ese momento. Por otra parte, Rose (1977) realizó el primer catálogo de CGs con un único criterio de selección. Identificó 35 CGs en proyección de manera visual cubriendo solo un 7.5% del cielo. El autor utilizó un criterio de selección visual y por lo tanto no reproducible. Debido a esto, la muestra de CGs presentaba muchos sesgos.

Para reducir estos efectos, Hickson (1982) construyó el más famoso catálogo de CGs, donde seleccionaron sistemas de 4 o 5 galaxias cuyas magnitudes no eran más débiles que 3 magnitudes a partir de la magnitud de la galaxia brillante. Asumieron un criterio de compacticidad utilizando un brillo superficial límite y para evitar que estos sistemas sean el núcleo de sistemas más grandes, adoptaron un criterio de aislamiento utilizando el radio geométrico del grupo. Identificó 100 CGs en proyección cubriendo un 67% del cielo. Sin embargo, debido a que la identificación de estos grupos se realizaba mediante un procedimiento visual, la selección de estos grupos se encuentran sesgada por quién los ha seleccionado.

Con la posibilidad de obtener información espectroscópica de las galaxias miembros, Hickson midió sus redshifts y encontró que sólo 69 de los 100 CGs son sistemas con velocidad concordante de al menos 4 galaxias miembro y no alineaciones a lo largo de la línea de la visual (Hickson et al., 1992). Este catálogo fue y sigue siendo hasta el día de hoy, el catálogo más usado de CGs (HCGs).

## 1.2. Búsqueda Automática de CGs

Para no tener que lidiar con la problemática de las identificaciones visuales, surge la idea de construir nuevas técnicas automatizadas de identificación.

Prandoni et al. (1994) identificaron CGs usando un algoritmo automatizado adaptado al criterio de selección propuesto por Hickson (1982) en proyección. Utilizó el relevamiento COSMOS/UKST Southern e identificó 59 CGs. Una comparación entre esta nueva muestra y la de Hickson indica que esta última está incompleta en términos de criterios de selección, satisfaciendo restricciones más estrictas que las establecidas explícitamente. Como consecuencia de esto, las conclusiones obtenidas a partir de la muestra HCGs se deben, al menos parcialmente, a sesgos visuales.

Años más tarde, Barton et al. (1996) construyeron el primer catálogo de CGs utilizando algoritmos de búsqueda automática seleccionados a partir de una muestra completa en magnitud en el espacio de redshifts. Los autores mantuvieron la idea de la naturaleza de un CG adoptada por Hickson, pero utilizaron un criterio de identificación diferente basado en el algoritmo Friends-of-Friends (FoF, Huchra & Geller (1982), Ramella et al. (1994)). Seleccionaron grupos con más de 3 miembros, basados sólo en la separación proyectada y la distancia a lo largo de la línea de la visual (redshifts concordantes). Se identificaron 89 CGs con propiedades similares a los HCG, sin tener en cuenta las magnitudes ni el criterio de aislamiento en el procedimiento de selección. De los 89 grupos, sólo 30 CGs fueron identificados con 4 o más miembros, el resto eran tripletes, y la gran mayoría de los grupos estaban embebidos en entornos densos.

Con el advenimiento de nuevos relevamientos con mejor tecnología para seleccionar objetos, de manera más eficiente (CCD), y alcanzando redshifts más altos se comenzaron a construir nuevos catálogos de CGs. Siguiendo con la idea de una identificación utilizando el algoritmo FoF desarrollada por Barton et al. (1996), Allam & Tucker (2000) seleccionaron 253 CGs con 3 o más miembros (76 con más de 4 miembros), encontrando propiedades similares a las obtenidas por Hickson (1982) o Barton et al. (1996). Nuevamente llegaron a la conclusión que estos sistemas se encuentran en general embebidos en sistemas de galaxias más grandes. Por otro lado, Focardi & Kelm (2002) identificaron CGs utilizando un algoritmo FoF más relajado permitiendo distancias proyectadas más grandes. Seleccionó un total de 291 CGs con más de 3 miembros (69 con 4 o más miembros). Encontraron que los CGs con 3 miembros tienden a ser sistemas aislados mientras que los CGs más numerosos son parte de un sistema mayor.

Luego, durante los siguientes 10 años, se construyeron catálogos utilizando únicamente criterios tipo Hickson y en proyección. Iovino (2002) identificó un total de 121 CGs luego de realizar una inspección visual para evitar galaxias intrusas en el anillo de aislamiento, mejorando el porcentaje de grupos conta-

minados comparados con los HCG. Por otra parte, en un trabajo similar, Iovino et al. (2003), identificaron 84 CGs utilizando otro relevamiento (DPOSS, Djorgovski et al. (1998)), alcanzando redshift más altos ( $z \sim 0,2$ ), pero modificando los parámetros del criterio Hickson (1982), en el cual bajaron la contaminación a un 10%. Una vez que el relevamiento DPOSS contó con una mayor cobertura del área del cielo, de Carvalho et al. (2005) identificaron 459 CGs, utilizando el mismo procedimiento.

Lee et al. (2004), identificaron CGs en el Sloan Digital Sky Survey (SDSS) Early Data Release obteniendo 175 CGs. Como mencionamos anteriormente, estos catálogos de CGs fueron identificados sin utilizar información espectroscópica. De las 744 galaxias en CGs, sólo 158 tenían información del redshift. Utilizando el catálogo SDSS Data Release Sixth (DR6), McConnachie et al. (2009) construyeron el catálogo más numeroso de CGs hasta ese momento. En su muestra "A" identificaron 2297 CGs mientras que en su muestra "B" identificaron 74791 CGs. La principal diferencia entre estas dos muestras es la magnitud aparente límite utilizada (18.0 mags para la muestra A y 21.0 mags para muestra B).

Posteriormente, Díaz-Giménez et al. (2012) identificaron 230 grupos utilizando el relevamiento 2MASS (Skrutskie et al. (2006a)) y el criterio Hickson en proyección, sumando otra condición necesaria al criterio que tiene que ver con la restricción en la magnitud de la galaxia más brillante. Sólo algunas galaxias tienen redshift disponible, por lo tanto los autores pudieron encontrar que 85 CGs tienen velocidades concordantes con al menos 4 miembros.

Por otro lado Hernández-Fernández & Mendes de Oliveira (2015) construyeron un catálogo de grupos compactos basándose en fuentes ultravioletas (UV), para seleccionar CGs con formación estelar. Seleccionaron 280 CGs usando un algoritmo FoF con longitudes de percolción angulares proyectadas, pero solo en proyección ya que no se contaba con el redshift de todas las fuentes.

En los últimos años se han construido grandes relevamientos espectroscópicos de galaxias completos en flujo que cubren un área importante del cielo y en donde la profundidad de catálogo alcanza redshift mayores a  $z \sim 0,2$ . Esto ha permitido construir nuevos catálogos de CGs teniendo el redshift de las galaxias y de esta manera, disminuir la contaminación.

Sohn et al. (2015) midieron los redshifts de las galaxias de los CGs identificados por McConnachie et al. (2009) y encontraron que el 42% de estas galaxias son intrusas. La muestra final de CGs con al menos 3 miembros con redshifts concordantes es de 332 (192 con al menos 4 miembros con velocidades concordantes). Por otro lado, Sohn et al. (2016) construyeron un catálogo utilizando un criterio tipo FoF con información espectroscópica a partir del SDSS DR12. Usaron el mismo algoritmo utilizado por Barton et al. (1996) sin tener en cuenta ningún criterio de aislamiento. Seleccionaron 1588 CGs con 3 o más miembros, de los cuales 312 tienen 4 o más.



Catálogo	Criterio	Relevamiento	#CGs	Área [ $deg^2$ ]
Rose (1977)	V	POSSI	35	3100
Hickson (1982)	H-proy (V)	POSSI	100	27000
Hickson et al. (1992)	H-proy-z(V)	POSSI	69	27000
Prandoni et al. (1994)	H-proy	COSMOS/UKST	59	1300
Barton et al. (1996)	FoF-z	CfA2/SSRS	30	HN
Allam & Tucker (2000)	FoF-z	LCRS	76	700
Focardi & Kelm (2002)	FoF-z	3D UZC	69	HN
Iovino (2002)	H-proy	COSMOS/UKST	121	5200
Iovino et al. (2003)	H-proy	DPOSS	84	2000
Lee et al. (2004)	H-proy	SDSS	175	153
de Carvalho et al. (2005)	H-proy	DPOSS	459	6260
McConnachie et al. (2009) [A]	H-proy	SDSS DR6	2297	9583
McConnachie et al. (2009) [B]	H-proy	SDSS DR6	74791	9583
Díaz-Giménez et al. (2012)	H-proy	2MASS	230	23844
Díaz-Giménez et al. (2012)	H-proy-z	2MASS	85	23844
Hernández-Fern. et al. (2015)	FoF-p	GALEX	280	22000
Sohn et al. (2015)	H-proy-z	SDSS DR6-DR12	192	9583
Sohn et al. (2016)	FoF-z	SDSS DR12	312	14555

**Tabla 1.1:** Catálogos observacionales de CGs identificados en diferentes relevamientos de galaxias. Los criterios utilizado para cada identificación se detallan en la segunda columna. *V*: Inspección visual; *H-proy*: Hickson en proyección; *H-proy-z*: Hickson en proyección con un filtrado posterior de la galaxias con información del redshifts; *FoF-z*: FoF en el espacio de redshift; *FoF-p*: FoF en proyección.

En la tabla 1.1 se muestran los diferentes catálogos observacionales con sus características principales como el relevamiento que utilizaron para la identificación, el número de grupos compactos identificados, el área del cielo donde se identificaron, y el criterio utilizado.

En resumen, la construcción de catálogos de CGs es un tema recurrente en el área de astronomía extragaláctica y que continuará hasta que se logren obtener muestras grandes y confiables de CGs que permitan realizar estudios estadísticos importantes de estos sistemas peculiares. La construcción de nuevos relevamientos profundos y que cubren una gran área del cielo, propone un escenario muy favorable para poder identificar este tipo de sistemas y tener muestras estadísticamente significativas. Hasta el momento no existe un criterio unificado de selección, pero como ya vimos, las identificaciones de CGs han utilizado mayoritariamente el criterio tipo Hickson, que replica la idea original del Hickson (1982), y el criterio que utiliza el algoritmo de identificación FoF pero adaptado para buscar sistemas pequeños.

### 1.2.1. Criterio tipo Hickson

El criterio de Hickson (1982) está basado en tres propiedades globales: población, compacticidad y aislamiento. Estas características están representadas

por propiedades observacionales y en proyección. El criterio puede resumirse de la siguiente manera:

- Población:  $4 \leq N \leq 10$ ,  $(m - m_{\text{bri}} \leq 3)$
- Compacticidad:  $\mu_m \leq \mu_{\text{lim}}$  [mag/arcsec<sup>-2</sup>]
- Aislamiento:  $\Theta_N > 3\Theta_G$
- Límite en flujo:  $m_{\text{bri}} \leq m_{\text{lim}} - 3$

donde  $N$  es el número de galaxias cuyas magnitudes aparentes están en un rango de 3 magnitudes desde la galaxia brillante;  $m_{\text{bri}}$  es la magnitud aparente de la galaxia más brillante del grupo;  $m_{\text{lim}}$  es la magnitud aparente límite del catálogo;  $\mu_m$  el brillo superficial medio, promediado sobre el círculo más pequeño que circunscribe los centros de las galaxias;  $\Theta_G$  es el diámetro angular del círculo circunscrito más pequeño;  $\Theta_N$  es el diámetro angular del círculo concéntrico más grande que no contiene otras galaxias dentro del rango de magnitudes seleccionado o más brillantes.

A pesar que el criterio sobre el *Límite en flujo* está implícito en el criterio original de Hickson (1982), es necesario incluir esta restricción en la identificación de estos grupos para asegurar que se cumplan los criterios de población y compacticidad, pero sólo algunos autores lo han tenido en cuenta al momento de utilizar dicho criterio. Esta nueva restricción al criterio de Hickson fue incluida por primera vez por Díaz-Giménez & Mamon (2010).

Cuando la información espectroscópica estuvo disponible en los grandes relevamientos, el criterio fue mejorado agregando el *Filtro en velocidad* como una nueva condición.

- Filtro en velocidad:  $|v_i - \langle v \rangle| \leq 1000 \text{ km s}^{-1}$

donde  $v_i$  es la velocidad radial de cada galaxia miembro, y  $\langle v \rangle$  es la mediana de los redshift de las galaxias miembros.

### 1.2.2. Criterio tipo FoF

La identificación de CGs utilizando el criterio tipo FoF esta basada en el algoritmo de identificación Friends-of-Friends (Huchra & Geller (1982)), el cual usa una longitud de percolación en proyección ( $\Delta D$ ), y, si se cuenta con información espectroscópica, se le suma una condición de cercanía a lo largo de la línea de la visual ( $\Delta cz$ ), para seleccionar grupos en proyección o en el espacio de redshifts. Alguno autores han sumando una criterio para asegurar la compacticidad, y otros no lo han tenido en cuenta.

Estas cantidades son arbitrarias y dependen del autor que identifique los CGs, como también la inclusión o no del criterio de compactidad. Generalmente, los parámetros utilizados son los siguientes:

- Algoritmo FoF:
  - $\Delta D \leq 50 \text{ kpc } h^{-1}$
  - $\Delta cz \leq 1000 \text{ km } s^{-1}$
- Compactidad:  $\mu_m \leq \mu_{\text{lim}} [\text{mag/arcsec}^{-2}]$

De esta manera, a lo largo de muchos años, los CGs fueron identificados en proyección (con o sin el criterio de límite en flujo), o en el espacio de redshifts. Es importante tener en cuenta que las magnitudes límites y el brillo superficial límite dependen de la banda fotométrica que se utilizó para la identificación.

### 1.3. Comparaciones entre catálogos y propiedades de CGs

Debido a la amplia diversidad de catálogos, la identificación de CGs no ha sido homogénea. Los relevamientos tienen diferentes límites en magnitud aparente, diferentes bandas y pueden tener o no información espectroscópica. Las galaxias seleccionadas en diferentes bandas van a formar CGs con diferentes características físicas. Además, los límites impuestos en los criterios son arbitrarios y pueden cambiar dependiendo de quien lo identifique.

A causa de estas diferencias, las propiedades de los grupos identificados en cada catálogo se pueden ver afectadas y se hace muy difícil comparar las muestras. Sin embargo, se han realizado estudios comparativos de CGs identificados en diferentes catálogos con diferentes características.

Díaz-Giménez et al. (2012) compararon los grupos que identificaron en el relevamiento 2MASS en proyección, y los grupos HCG. Sólo el 25% de los 2MCG fueron identificados por otros autores. Además, de los 42 HCG que existen en la misma área que el 2MCG, solo 20 fueron identificados, el resto de los HCGs no fueron identificados probablemente por el límite en la banda K, ya que no todos los CGs fueron identificados con el criterio en límite en flujo que asegura la membresía. Hernández-Fernández & Mendes de Oliveira (2015) también realizaron un estudio comparativo entre sus grupos identificados en proyección. De las 463 galaxias de HCG solo 87 son fuentes FUV, y por lo tanto, solo 3 HCG tienen más de 4 miembros UV. A pesar de esto, no cumplen el criterio de aislamiento. Hernández-Fernández & Mendes de Oliveira (2015) también estudiaron y analizaron la relación entre el diámetro físico proyectado en función

del redshift, comparando los catálogos 2MCG, HCG y los grupos identificados por McConnachie et al. (2009). Las diferencias entre los CGs son probablemente debido al criterio específico de selección aplicado a cada muestra y a la muestra de galaxias de la cual fueron extraídos los grupos. Para estudiar la compacticidad los SFCG utilizaron  $\Delta D$  proyectada mientras que los catálogos con criterio tipo Hickson utilizaron el brillo superficial  $\mu$ . Esto genera claras dependencias en tamaños con el redshift.

Estos estudios han demostrado que la identificación de CGs no es una tarea sencilla y utilizar diferentes criterios y relevamientos puede afectar el análisis de las propiedades. Además, incluso identificando en la misma área del cielo, diferentes autores no encontraron los mismos CGs. Es principalmente por esta razón que aún se siguen realizando nuevos catálogos de CGs, intentando mejorar la identificación para poder estudiar la naturaleza de estos grupos en una muestra de CGs que no esté muy contaminada por CGs espurios.

Sin embargo, durante ya más de 40 años se han realizado numerosos estudios sobre CGs utilizando cada uno de los catálogos mencionados, para entender los procesos de formación y evolución de estos objetos. A modo de ejemplo, mencionaremos algunos estudios realizados sobre las galaxias miembros de los CGs. En Duplancic et al. (2013) utilizaron una muestra de tripletes (O'Mill et al. (2012)) y un muestra de CGs (McConnachie et al. (2009), catálogo A) y encontraron que las galaxias en tripletes se comportan de manera similar a las galaxias en CGs (población estelar, colores y tasa de formación estelar). Los autores sugieren que los tripletes al tener propiedades más similares a los CGs que a los pares de galaxias, y entornos parecidos (grupos aislados localmente), son una extensión natural de los CGs. Duplancic et al. (2018) realizaron un trabajo similar pero utilizando un criterio tipo Hickson para identificar los grupos (pares, tripletes y grupos pequeños (CGs)). Encontraron que mientras más numerosos son los sistemas, más rojas y con baja formación estelar son las galaxias que lo componen. Coenda et al. (2015) realizaron un trabajo comparativo utilizando los CGs (catálogo A), pero con grupos difusos (menos densos y mas numerosos) identificados por Zandivarez & Martínez (2011). Encontraron que galaxias en CGs tienen poblaciones mas viejas que los grupos difusos o galaxias en campo. Debido principalmente a este resultado, los autores concluyen que los CGs son ambientes extremadamente favorables para que se producen procesos que transformen galaxias con formación estelar en galaxias pasivas siendo esta transición más eficientemente y más rápida que en otros entornos. Por otra parte, Mendes de Oliveira et al. (2015) encontraron que el 95% de las galaxias en sus grupos (SFCGs) son galaxias con alto porcentaje de formación estelar (identificados con fuentes UV), pero sugieren que estas diferencias entre SFCGs y otras muestras seleccionadas en el óptico o infrarrojo pueden deberse a que los SFCGs se encuentran en un estado evolutivo diferente. Sohn et al. (2015) encontraron que el porcentaje de galaxias tipo temprano en sus CGs es de un 62% .

## 1.4. Estudios de CGs a partir de catálogos sintéticos

Uno podría recurrir a las simulaciones numéricas cosmológicas para inferir la naturaleza y frecuencia de los CGs. Muchos autores han extraído CGs usando modelos semianalíticos (SAM, sigla en inglés) de formación de galaxias y evolución para este fin. Los SAMs emplean recetas físicas para describir galaxias, usualmente asociadas con los subhalos de los halos de una simulación cosmológica. Dadas las posiciones de las galaxias en diferentes estadios temporales, uno puede construir catálogos sintéticos para generar muestras de galaxias en el espacio distorsionado (espacio de redshift). Luego puede correrse un algoritmo para extraer CGs de esos catálogos. Los CGs extraídos de esta manera cuentan con información en el espacio de redshifts como así también en el espacio 3D, y de esta manera se puede determinar si los CGs son físicamente densos o espurios. McConnachie et al. (2008) encontraron que el 30% de los CGs seleccionados en proyección son físicamente densos en el espacio 3D. Díaz-Giménez & Mamon (2010) (DGM10) analizaron las propiedades de los CGs identificados en proyección utilizando 3 SAMs diferentes, todos aplicados a la misma simulación, y encontró que la fracción de CGs físicamente densos varía entre un 60% a un 75% dependiendo del SAM. Estos autores también encontraron que la densidad espacial de CGs es 10 veces más alta en las simulaciones que en las observaciones, lo que sugiere que la muestra de HCG es extremadamente incompleta. Estudios subsecuentes de CGs extraídos de diferentes SAMs basados en físicas diferentes o corridos sobre simulaciones con diferentes parámetros cosmológicos han obtenido diferentes fracciones de grupos físicamente densos. Por lo que el tema de la verdadera naturaleza de los CGs no puede considerarse como un caso cerrado.

Por otro lado, Wiens et al. (2019) desarrollaron un catálogo 3D de grupos compactos para estudiar la evolución de estos grupos en simulaciones numéricas. Diseñaron un criterio de CGs en el espacio real que representa las propiedades físicas típicas de los HCG ( $z=0$ ) y rastrearon las galaxias hasta  $z=11$ . Estos autores encontraron que el 16% de las galaxias en CGs han sido miembros de CGs en algún momento de su historia. Curiosamente, la gran mayoría (96%) de los CGs a  $z=2$  sufrieron colisiones entre sus galaxias, fusionándose en una única galaxia a  $z=0$ . Farhang et al. (2017) estudiaron la historia evolutiva de CGs (identificados utilizando un criterio tipo Hickson) en simulaciones numéricas usando SAMs, desde  $z=1$  hasta  $z=0$ . Encontraron que el 55% de la masa del halo de CGs ha sido ensamblado a  $z \sim 1$ , sugiriendo que estos sistemas son relativamente jóvenes. Encontraron también que los CGs se vuelven más compactos con el tiempo, perdiendo su condición de compactidad a  $z \sim 0,3$ .

## 1.5. Planificación del trabajo

En esta tesis utilizaremos los catálogos sintéticos contruidos a partir de simulaciones numéricas en combinación con SAMs de formación de galaxias, para analizar en profundidad varios de los problemas que han venido caracterizando a los catálogos observacionales de CGs durante los últimos 40 años. La descripción de las herramientas utilizadas se encuentra en el capítulo 2.

En primer lugar, en el capítulo 3, abordaremos el tema de la banda fotométrica en la cual se identifican los CGs de manera de desentrañar si es factible realizar comparaciones directas entre los diferentes catálogos de CGs que se encuentran en la literatura.

En el capítulo 4 mostraremos una nueva manera de identificar CGs en catálogos observacionales siguiendo el criterio de Hickson pero optimizando el algoritmo para mejorar la completitud de la muestra resultante.

En tercer lugar, en el capítulo 5 se realizará un estudio muy completo tratando de entender la naturaleza de los CGs teniendo en cuenta un gran número de variables presentes cuando se utilizan catálogos sintéticos. Este trabajo involucrará la utilización de una amplia variedad de simulaciones y modelos semianalíticos.

Finalmente, en el capítulo 6, se intentará desentrañar la verdadera naturaleza de los CGs que surgen de aplicar el criterio tipo Hickson a un catálogo observacional. Para eso se definirá una muestra *ideal* de grupos físicamente densos y aislados en las 3 dimensiones del espacio real (no distorsionado por velocidades peculiares) y se analizará la capacidad del conocido criterio de Hickson para recuperar estos sistemas.

De lo resultados obtenidos en esta tesis esperamos poder contribuir a mejorar el entendimiento de estos sistemas tan peculiares en el universo y sus métodos de identificación, así como también poder cuantificar las limitaciones de las herramientas que se han venido usando en la literatura.

# Capítulo 2

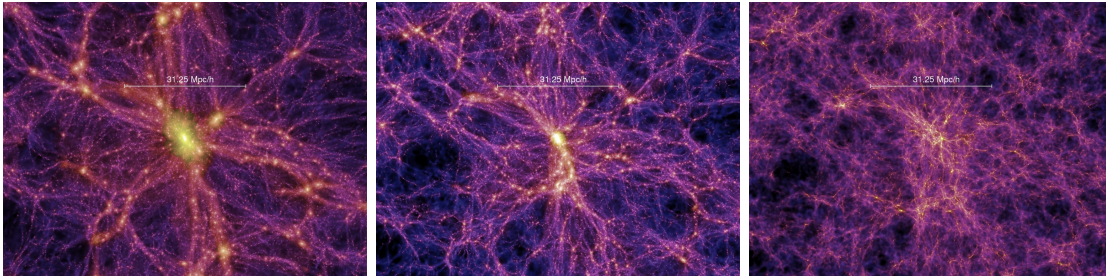
## Catálogos sintéticos

Una herramienta fundamental en el estudio de la formación y evolución de estructuras en el universo son las simulaciones, ya sean hidrodinámicas o de materia oscura. En particular, en las simulaciones cosmológicas de N-cuerpos, partículas no colisionales evolucionan gravitacionalmente recreando la evolución de las estructuras del universo en regiones cúbicas y periódicas, partiendo de una distribución de materia inicial hasta llegar a simular las distribuciones que se pueden detectar en el universo observable en el presente. En conjunto con modelos semianalíticos de formación de galaxias, se pueden recrear las distribuciones de toda la materia (oscura y bariónica) en el universo y estudiar la formación y evolución de galaxias en el universo simulado.

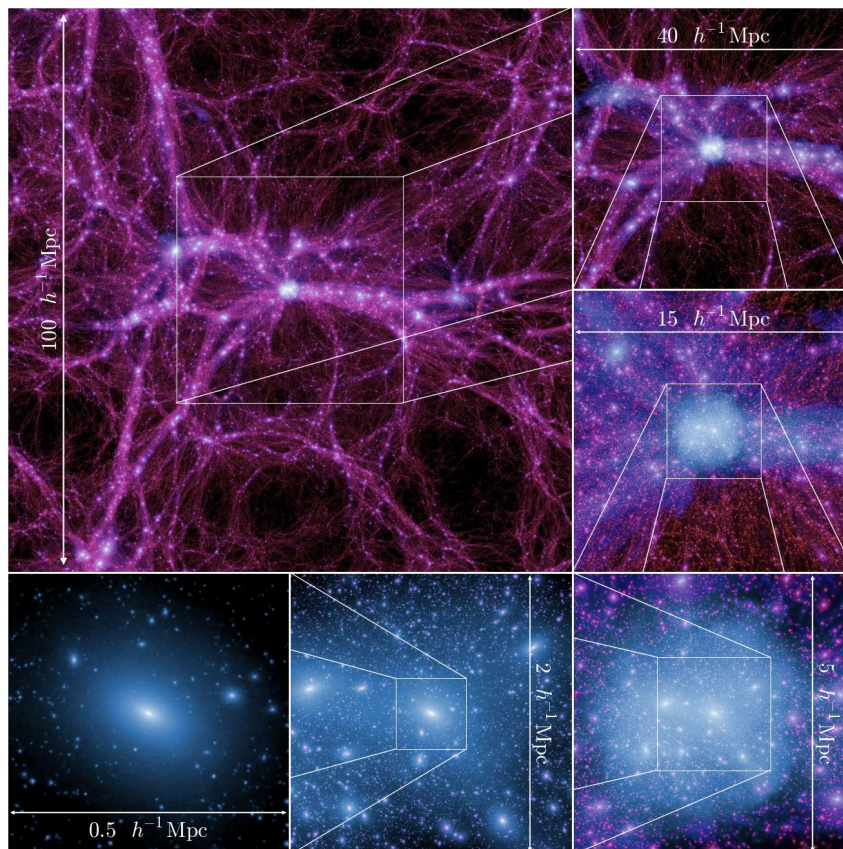
En este trabajo hemos utilizado diferentes simulaciones numéricas cosmológicas combinadas con distintos modelos semianalíticos de formación de galaxias. Todas las muestras utilizadas se encuentran disponibles públicamente para la comunidad científica. A continuación realizaremos una descripción de las simulaciones y los modelos utilizados, y detallaremos la forma en que las muestras fueron utilizadas en este trabajo de tesis.

### 2.1. Simulaciones numéricas

A lo largo de esta tesis hemos utilizado tres simulaciones cosmológicas de materia oscura: Millennium Run Simulation I (Springel et al. 2005a), Millennium Run Simulation II (Boylan-Kolchin et al. 2009) y MultiDark (Klypin et al. 2016). La evolución gravitacional de las partículas de materia oscura para las tres simulaciones fue realizada por el código GADGET2 (Springel et al., 2001). En las figuras 2.1, 2.2 y 2.3 se visualizan las partículas de materia oscura en el cubo periódico de las simulaciones Millennium I (MS), Millennium II (MSII) y Multidark (MDPL2), respectivamente.

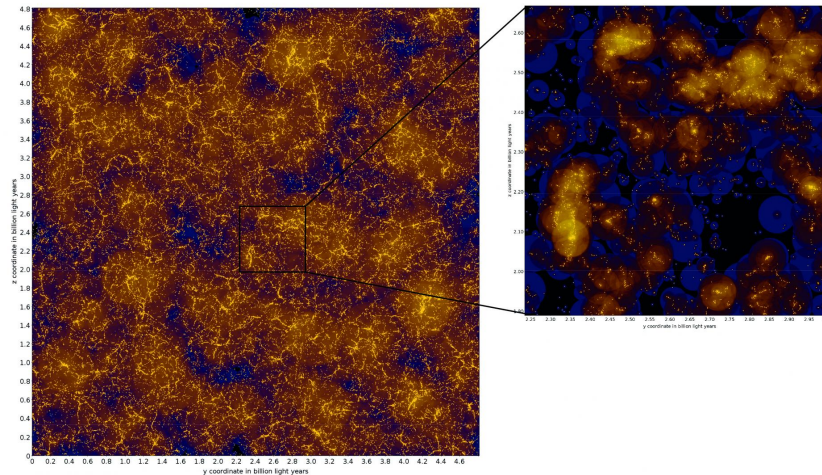


**Figura 2.1:** Evolución de la densidad de materia en la simulación Millennium I (Springel et al., 2005b) para tres estadios evolutivos diferentes. Panel izquierdo: redshift  $z = 0$  ( $t = 13,6$  Gyr), panel medio: redshift  $z = 1,4$  ( $t = 4,7$  Gyr), panel derecho: redshift  $z = 5,7$  ( $t = 1,0$  Gyr).



**Figura 2.2:** Campo de densidades de la simulación Millennium II Boylan-Kolchin et al. (2009) desde 100 hasta  $0.5 \text{ Mpc}/h$  dentro del halo más masivo de materia oscura a redshift cero.





**Figura 2.3:** Visualización de la distribución de materia oscura en la simulación MultiDark Klypin et al. (2016).

Las tres simulaciones fueron corridas en un marco cosmológico determinado por un universo plano con constante cosmológica ( $\Lambda$ CDM). Los parámetros cosmológicos que caracterizan la cosmología son los siguientes:

- $\Omega_m$ : densidad total de materia (materia oscura + bariones)
- $\Omega_b$ : densidad de bariones
- $\Omega_\Lambda$ : densidad de energía oscura
- $n$ : índice espectral del espectro de potencia primordial
- $\sigma_8$ : amplitud de las fluctuaciones en densidad en esferas de  $8 h^{-1}\text{Mpc}$  al presente
- $h$ : constante de Hubble al presente ( $H_0 = 100 h \text{ km s}^{-1}\text{Mpc}^{-1}$ )

En la tabla 2.1 se muestran los parámetros cosmológicos que corresponden a las cosmologías WMAP1 (Spergel et al., 2003), WMAP7 (Komatsu et al., 2011) y Planck (Planck Collaboration et al., 2015) junto con las simulaciones a los que fueron aplicados. La simulación MS fue realizada con dos de estas cosmologías, WMAP1 y WMAP7, y también se realizó un re-escalado de la simulación para adaptarla a la cosmología Planck. La MSII fue realizada con cosmología WMAP1 y re-escalada para la cosmología Planck. En el caso de la simulación MDPL2, la cosmología utilizada fue Planck.

## 2.2. Modelos semianalíticos de formación de galaxias

Los modelos semianalíticos de formación de galaxias (SAMs) proporcionan una herramienta de gran importancia en el estudio de las galaxias en el universo.

Cosmología	$\Omega_m$	$\Omega_b$	$\Omega_\Lambda$	$n$	h	$\sigma_8$	Simulación
WMAP1	0,250	0,043	0,750	1,00	0,730	0,900	MS, MSII
WMAP7	0,272	0,045	0,728	0,96	0,704	0,807	MS
Planck	0,307	0,048	0,693	0,96	0,677	0,823	MS, MSII, MDPL2

**Tabla 2.1:** Parámetros cosmológicos de las cosmologías utilizadas en las simulaciones.

Estos modelos simulan los principales procesos físicos que ocurren en el universo, principalmente en la formación y evolución de galaxias. Algunos de los procesos involucrados son la caída y el enfriamiento del gas, reionización de universo, formación estelar, crecimiento del agujero negro central, retroalimentación (feedback) debido a núcleos activos de galaxias (AGN) y supernovas (enriquecimiento químico), colisiones de galaxias, evolución fotométrica, entre otras. En la práctica, estas ecuaciones describen cómo los bariones se mueven entre los diferentes reservorios de masa (halos de materia oscura). A través de un tratamiento analítico, estos bariones están relacionados con los árboles de fusión (merger tree) de la materia oscura y se siguen a través de su evolución en el tiempo. Se ajusta un conjunto de parámetros para obtener resultados que se aproximen de la mejor manera posible a las principales propiedades observacionales de las poblaciones de galaxias del Universo local. En general, los parámetros utilizados para esta calibración son los utilizados para reproducir las distribuciones de color de las galaxias observacionales, las funciones de luminosidad y/o de masa estelar, a redshift  $z = 0$  o incluso a redshifts más altos.

Por lo tanto, los modelos SAMs representan un escenario ideal para estudiar el comportamiento de las galaxias y los sistemas de galaxias en el espacio real para realizar estudios evolutivos de formación y evolución de estructuras, como así también sirven para realizar predicciones respecto del comportamiento de las galaxias y los sistemas en el espacio distorsionado de las observaciones (espacio de redshifts).

Sin embargo, las propiedades de las galaxias que los SAMs intentan reproducir, pueden verse afectadas por las recetas físicas utilizadas en estos modelos. Por ejemplo, en los SAMs, el feedback de AGN reduce la luminosidad y la masa estelar de las galaxias más brillantes, mientras que las supernovas son más eficientes en remover el gas de las galaxias de baja masa, lo que disminuye su tasa de formación estelar. Al tomar en cuenta las AGN, las supernovas y otros procesos físicos, los SAMs han logrado reproducir muchas propiedades observables importantes de las galaxias.

En la tabla 2.2 se citan los autores de todos los modelos semianalíticos utilizados en esta tesis junto con las simulaciones a las que fueron aplicados. En el apéndice A detallamos los programas para descargar las salidas de las simulaciones desde las bases de datos públicas de los consorcios que desarrollaron las simulaciones.

SAMs	Simulación	Cosmología	$L_{cubo}$ [ $h^{-1}$ Mpc]	Acrónimo
De Lucia & Blaizot (2007)	MS	WMAP1	500	<b>DLB</b>
Guo et al. (2011)	MS	WMAP1	500	<b>G11</b>
Guo et al. (2011)	MSII	WMAP1	100	<b>GII</b>
Guo et al. (2013)	MS	WMAP7	500	<b>G13</b>
Henriques et al. (2015)	MS	Planck	480	<b>HrI</b>
Henriques et al. (2015)	MSII	Planck	96	<b>HrII</b>
Cora et al. (2018)	MDPL2	Planck	1000	<b>SAG</b>

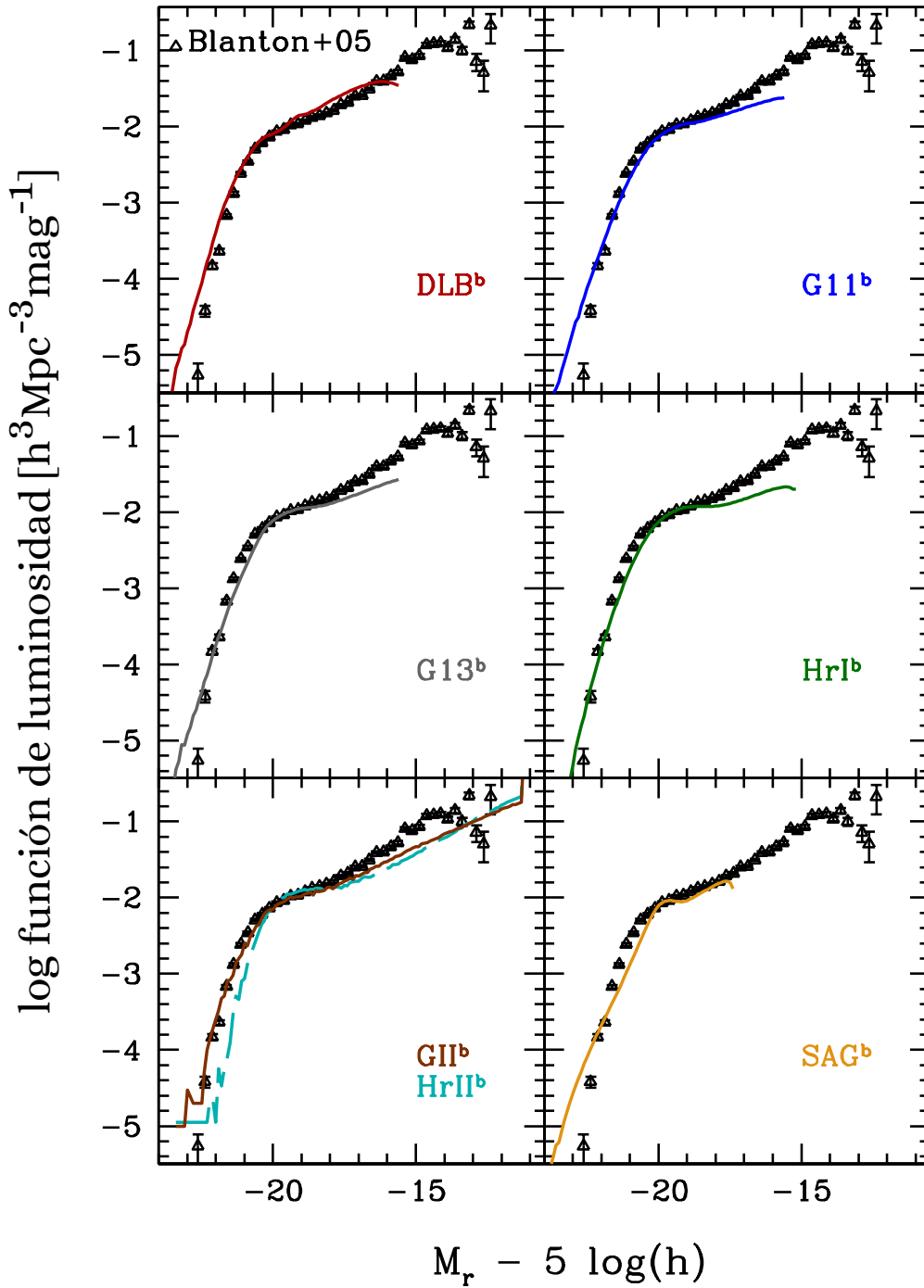
**Tabla 2.2:** Modelos semianalíticos de formación de galaxias utilizados en este trabajo.

Las galaxias resultantes en cada SAM presentan diferencias entre un modelo u otro debido a las diferentes recetas sobre los procesos físicos que fueron tratados de diferentes maneras por diferentes autores. A continuación mostraremos los resultados de algunas de las propiedades de las galaxias en el cubo de las simulaciones a redshift  $z = 0$ , para los siete modelos SAMs. En particular veremos la función de luminosidad (FL), la función de masa estelar (SMF), la fracción de los distintos tipos morfológicos como función de la masa estelar, y la función de correlación de dos puntos (galaxia-galaxia).

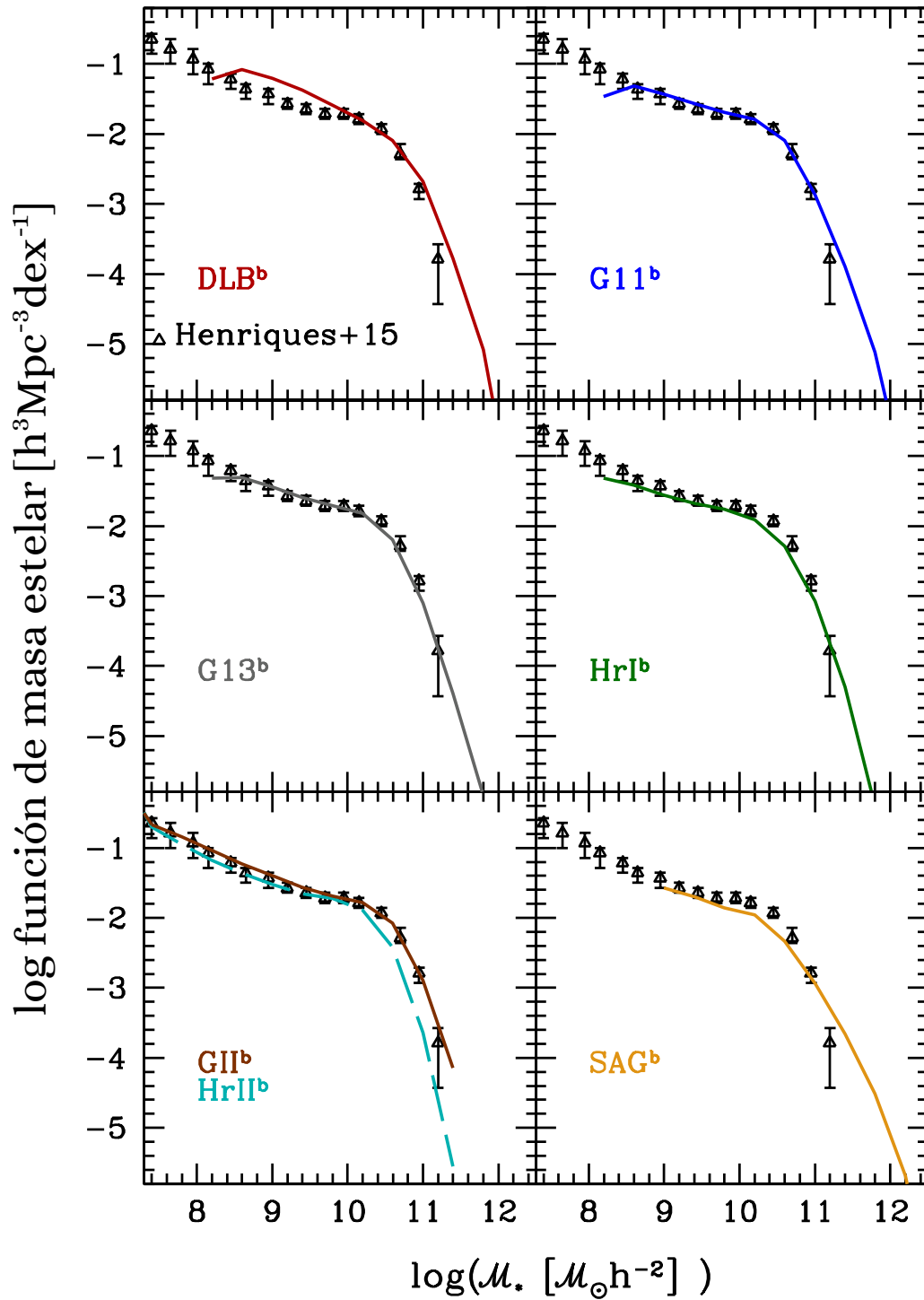
En la figura 2.4 comparamos la función de luminosidad de galaxias a  $z=0$  en la banda  $r$ -SDSS para los siete SAMs<sup>1</sup> con la muestra observacional de Blanton et al. (2005) obtenida usando el SDSS DR2 (SDSS DR2, Abazajian et al. 2004). Los SAMs, en general, reproducen bien la FL observacional. En particular, DLB<sup>b</sup> sobre predice el extremo brillante de la FL y también el rango de magnitudes más débiles que la  $M^*$ . Por otro lados, G11<sup>b</sup>, G13<sup>b</sup> y HrI<sup>b</sup> reproducen bien la FL observada alrededor de  $M^*$  y el extremo brillante, pero subestiman la FL a luminosidades bajas ( $M_r - 5 \log(h) > -18$ ). GII<sup>b</sup> muestra un buen acuerdo con las observaciones para galaxias mas luminosas que  $M_r = -17$ , mientras que HrII<sup>b</sup> subestima el extremo brillante de la FL. Estos últimos dos SAMs muestran una carencia de galaxias hacia el extremo débil de la FL. Finalmente, SAG<sup>b</sup> sobre predice el extremo brillante de la FL, pero recupera de buena manera las observaciones para galaxias con magnitudes más débiles que  $M^*$ .

En la figura 2.5 mostramos la función de masa estelar (SMF) de los siete SAMs comparados con las observaciones a partir de los catálogos SDSS y Galaxy and Mass Assembly (GAMA) realizada por Henriques et al. (2015). Se puede ver que DLB<sup>b</sup> sobre estima la cantidad de galaxias con bajas y altas masas estelares. G11<sup>b</sup> usó la SMF de las galaxias del SDSS para la calibración de su SAM por lo cual se obtiene un mejor acuerdo de la SMF con las observaciones, aunque sobre predice galaxias en el extremo de galaxias con masas estelares altas. G13<sup>b</sup> recupera muy bien la SMF en todo el rango de masas estelares,

<sup>1</sup>El superíndice “b” indicará cuando se utilizan galaxias en un único cubo de las simulaciones



**Figura 2.4:** Función de luminosidad de galaxias en el cubo de la simulación a  $z = 0$  en la banda  $r$ -SDSS para los diferentes SAMs (curvas, ver acrónimos en la tabla 2.2). El superíndice “b” indica que se tomaron galaxias en un cubo (box) de la simulación y para la muestra observacional SDSS DR2 (símbolos) obtenida por Blanton et al. (2005).



**Figura 2.5:** Función de masa estelar en el cubo de la simulación a  $z = 0$  para diferentes SAMs (*curvas*) comparadas con los datos observacionales de Henriques et al. 2015 a partir de la combinación de las muestras observacionales SDSS (Baldry et al., 2008; Li & White, 2009) y GAMA (Baldry et al., 2012) (*symbols*).

mientras que HrI<sup>b</sup> (quien también utilizó la SMF a  $z = 0$  para calibrar su SAM) muestra una densidad de galaxias levemente más baja cerca de la masa estelar característica. En la Millennium II, GII<sup>b</sup> recupera las observaciones en todo el rango de masas estelares, mientras que HrII<sup>b</sup> subestima el número de galaxias en el extremo de masas estelares altas. Por último, SAG<sup>b</sup> sobre estima galaxias en el extremo de masas estelares altas de la SMF, mientras que pierde galaxias para masas iguales o más bajas que la masa estelar característica.

Siguiendo con el análisis de las propiedades observacionales de las galaxias en comparación con las galaxias simuladas, en la figura 2.6 se muestra la fracción de los diferentes tipos morfológicos de las galaxias como función de la masa estelar en cada SAM comparada con la obtenida por Conselice (2006) para galaxias en el catálogo Third Reference Catalogue of Bright Galaxies (RC3, de Vaucouleurs et al. 1991). Los tipos morfológicos sintéticos han sido definidos acorde a la relación masa estelar del bulbo-masa estelar total de la siguiente manera:

**Galaxias elípticas:**  $\mathcal{M}_{\text{bulbo}}/\mathcal{M}^* \geq E_{\text{lim}}$ ,

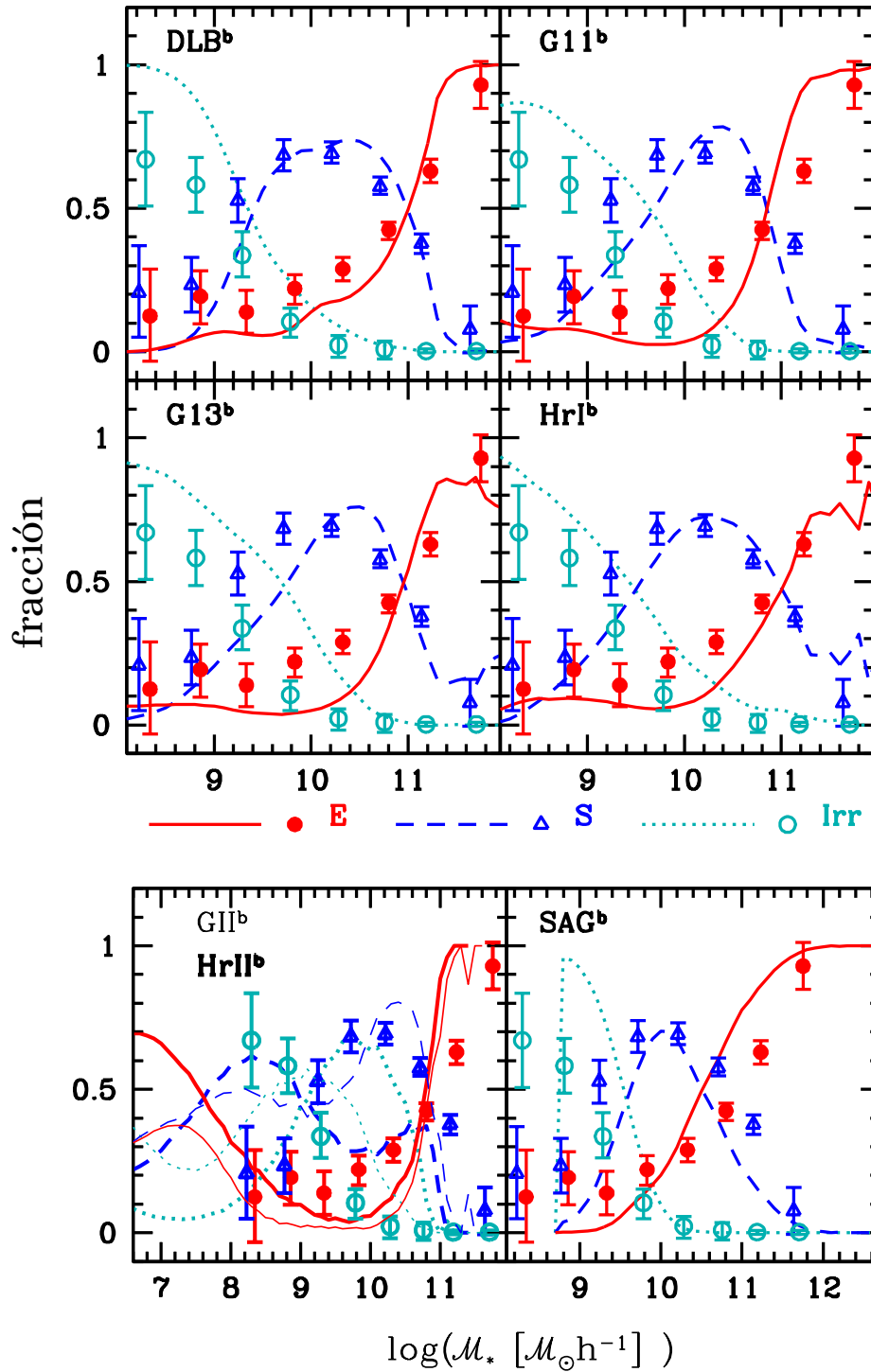
**Galaxias espirales:**  $S_{\text{lim}} < \mathcal{M}_{\text{bulbo}}/\mathcal{M}^* < E_{\text{lim}}$ , y

**Galaxias irregulares:**  $\mathcal{M}_{\text{bulbo}}/\mathcal{M}^* \leq S_{\text{lim}}$ ,

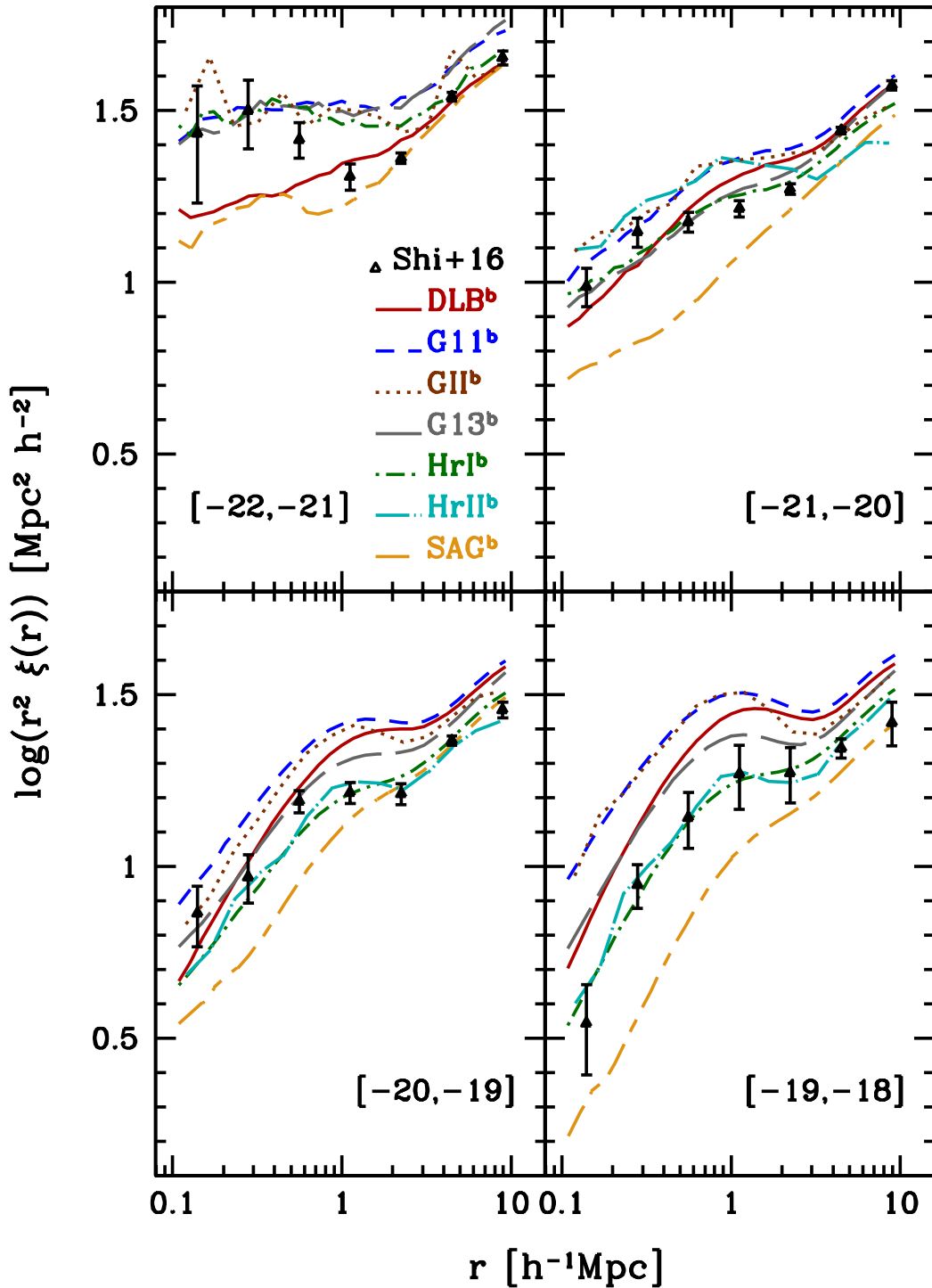
donde,  $E_{\text{lim}}$  es el umbral a partir del cual las galaxias son consideradas elípticas o espirales, y  $S_{\text{lim}}$  es el limite para el cual las galaxias dejan de ser espirales y pasan a ser consideradas irregulares. Siguiendo los trabajos publicados en la literatura, adoptamos los siguientes valores límites para obtener el mejor acuerdo con los datos observacionales:  $S_{\text{lim}} = 0,03$  y  $E_{\text{lim}} = 0,7$  para DLB<sup>b</sup>, G11<sup>b</sup>, G13<sup>b</sup>, GII<sup>b</sup> y HrII<sup>b</sup>s (Guo et al., 2011),  $S_{\text{lim}} = 0$  y  $E_{\text{lim}} = 0,7$  para HrI<sup>b</sup> (Bertone et al., 2007), mientras que  $S_{\text{lim}} = 0$  y  $E_{\text{lim}} = 0,85$  para SAG<sup>b</sup> (Cora et al., 2018).

Los SAMs reproducen bastante bien a distribución de los diferentes tipos morfológicos con respecto a las observaciones. En particular, los SAMs basados en la Millennium I son los que mejor recuperan las observaciones. Por otro lado, GII<sup>b</sup> y HrII<sup>b</sup> sobre estiman la fracción de galaxias espirales para masas estelares bajas (desde 8 hasta 9  $\mathcal{M}_{\odot} h^{-1}$ ), HrII<sup>b</sup> subestima la fracción de galaxias espirales a masas intermedias y SAG<sup>b</sup> subestima la fracción de galaxias espirales en el rango de masas estelares más alto y más bajo.

Finalmente, la figura 2.7 muestra la función de correlación de dos puntos en el espacio real de galaxias a  $z = 0$  obtenida para cada SAM y calculada usando el estimador desarrollado por Landy & Szalay (1993) usando la función `tpcf` del paquete `Halotools` (Hearin et al., 2017). Hemos dividido la muestra de galaxias en rangos de magnitud absoluta para comparar con las funciones de correlación de dos puntos en el espacio real reconstruidas por Shi et al. (2016) a partir de las galaxias extraídas del SDSS DR7 (Abazajian et al., 2009).

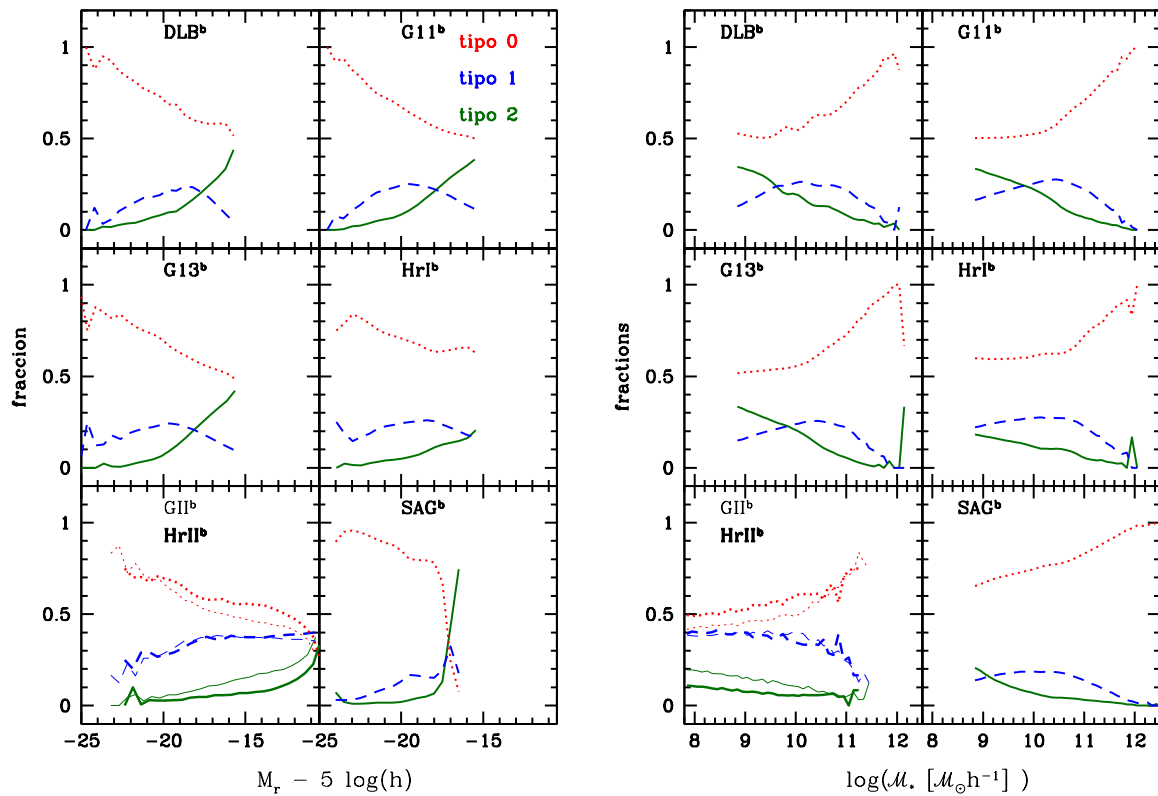


**Figura 2.6:** Fracción del tipo morfológico (elípticas *red*), espirales (*blue*), e irregulares (*cyan*), como función de la masa estelar en el cubo de la simulación a  $z = 0$ . *Lineas* muestran las fracciones obtenidas para los diferentes SAMs (ver acrónimos en la Tabla 2.2): para las elípticas (*continuas*), espirales (*discontinua*) e irregulares (*a puntos*). Los paneles inferiores, tiene los SAMs corridos en las simulaciones con mayor resolución por lo que tienen eje  $x$  es diferente. Las observaciones (*símbolos*) fueron calculadas por Conselice (2006) a partir del RC3.



**Figura 2.7:** Función de correlación de dos puntos en el espacio real a  $z=0$  en el cubo de la simulación para los diferentes SAMs (*lineas*). Cada muestra SAM fue restringida en magnitud absoluta para respetar los límites de la muestra observacional los cuales se muestran con corchetes en el gráfico. (En el panel superior izquierdo, la muestra HrlI<sup>b</sup> no fue incluida dado el bajo número de galaxias en ese rango de magnitudes). Los *símbolos* corresponden a una función de correlación en el espacio real reconstruida para las galaxias del SDSS DR7 determinada por Shi et al. (2016).





**Figura 2.8:** Fracción de galaxias de diferentes tipos como función de la magnitud absoluta (*izquierda*) y de la masa estelar (*derecha*). Las líneas punteadas rojas son las fracciones de galaxias centrales (tipo 0), las líneas azules a rayas son satélites (tipo 1), y las líneas sólidas verdes son las huérfanas (tipo 2).

Puede verse que  $\text{HrI}^b$  y  $\text{HrII}^b$  son los SAMs que mejor reproducen la función de correlación observacional en la mayoría de los rangos para todas las escalas. En contraste,  $\text{SAG}^b$  subestima el agrupamiento o *clustering* de galaxias a pequeñas escalas ( $r < 1 h^{-1} \text{ Mpc}$ ) en todos los rangos de magnitud.  $\text{G11}^b$  sobreestima el *clustering* en todas las escalas.  $\text{G13}^b$  y  $\text{DLB}^b$  recuperan de buena manera la función de correlación a diferentes escalas dependiendo del rango de magnitudes.

Es interesante notar que ninguno de los SAMs ha usado la función de correlación para calibrar sus resultados. Sin embargo, Pujol et al. (2017) demostraron que las diferencias en el tratamiento de las llamadas galaxias huérfanas en los diferentes SAMs y la fracción de éstas galaxias en cada cubo pueden tener un impacto directo en el clustering de las galaxias, principalmente a escalas pequeñas, que podría explicar las diferencias observadas en esta figura. Las galaxias huérfanas son aquellas cuyos subhalos caen por debajo de la resolución de la simulación, se dice que “pierden su halo de materia oscura” y por ello sus posiciones se asumen usando recetas analíticas para su evolución orbital que incluyen fricción dinámica, y por lo tanto esas posiciones estimadas tienen muchas incertezas.

En cada modelo SAM, las galaxias son clasificadas como “centrales” (tipo 0) que son las galaxias dominantes de un halo, “satélites” (tipo 1) que son galaxias no dominantes en los halos, y “huérfanas” (tipo 2) cuyos halos no pueden ser resueltos en la simulación. En general, las galaxias nacen como tipo 0, usualmente se convierten en tipo 1 cuando caen en un grupo o cúmulo, y pueden eventualmente convertirse en tipo 2, para finalmente fusionarse con la galaxia central de su halo. Cada modelo SAM adopta un tratamiento diferente para las galaxias huérfanas (ver Pujol et al., 2017 quienes hacen una comparación de los tratamientos de huérfanas en nueve modelos SAMs).

Si consideramos las galaxias en los cubos de las simulaciones a  $z = 0$  con masas estelares mayores que  $7 \times 10^8 h^{-1} M_{\odot}$  y magnitudes absolutas  $M_r \leq -16$ , las fracciones de galaxias huérfanas son 0.26, 0.25, 0.25, 0.14 y 0.12 para DLB<sup>b</sup>, G11<sup>b</sup>, G13<sup>b</sup>, HrI<sup>b</sup>, y SAG<sup>b</sup>, respectivamente. Tomando los cubos de galaxias en la MSII ( $7 \times 10^7 h^{-1} M_{\odot}$  y magnitud absoluta  $M_r \leq -11$ ), la fracción de huérfanas en estos cubos es más baja, aún cuando el rango de masas es más extenso: 0.16 en GI<sup>b</sup> y 0.09 en HrII<sup>b</sup>. En la figura 2.8 se muestra la contribución de galaxias de cada tipo en los cubos de las simulaciones en bins de magnitudes absolutas (paneles izquierdos) y de masas estelares (paneles derechos). Puede verse que la fracción de galaxias huérfanas se incrementa hacia magnitudes débiles y masas estelares chicas ya que es en esos rangos donde la simulación no logra resolver los halos más pequeños.

Dado que nuestros estudios estarán focalizados en las asociaciones de galaxias en espacios reducidos, este punto puede ser relevante en cuanto a los resultados sobre grupos compactos (pequeña escala) obtenidos a partir de cada SAM, por lo que será estudiado en detalle en otro capítulo más adelante.

### 2.3. Catálogos sintéticos en el espacio distorsionado

En los diferentes capítulos de esta tesis se utilizarán catálogos sintéticos contruidos con el objetivo de reproducir propiedades de distintos catálogos observacionales. En esta sección detallaremos los procedimientos utilizados para que los catálogos simulados de galaxias semianalíticas – que habitan el espacio real de coordenadas comóviles en los cubos de las simulaciones con sus propiedades absolutas, descritos en la sección anterior– sean transformados en catálogos sintéticos de galaxias en el espacio distorsionado de los redshifts y con propiedades observacionales.

Partiendo de que los cubos de las simulaciones tienen tamaños finitos, para construir catálogos sintéticos se debe replicar el cubo de la simulación tantas veces como sea necesario de manera de alcanzar la profundidad deseada, construyendo así un supercubo de galaxias en coordenadas comóviles. La periodicidad de los cubos de la simulación es la que permite que este procedimiento sea

factible. Uno puede construir catálogos sintéticos en el espacio de redshifts de dos formas: con evolución o sin evolución; la elección dependerá del estudio que se desee realizar, la profundidad de la muestra con la que se vaya a trabajar y la importancia de la evolución en las propiedades que se quiera analizar. Cuando el catálogo sintético que se desea construir no incluye evolución, un único cubo de la simulación (en general, el de la última salida que representa el estado evolutivo del universo al tiempo presente,  $z=0$ ) se repite colocándose uno al lado del otro (ver imagen izquierda de la figura 2.9). De manera similar, si se desea incluir la evolución gravitatoria de las estructuras y de las propiedades astrofísicas de las galaxias con el tiempo, en lugar de utilizar un único cubo de datos de la simulación, se utilizan las diferentes salidas de la simulación en diferentes estadios evolutivos para ir construyendo el supercubo comóvil, y se van seleccionando galaxias en cáscaras de diferentes salidas alejadas del origen de coordenadas una distancia correspondiente con el estadio evolutivo. Dada la evolución gravitatoria de las estructuras, se debe tener cuidado de no incluir galaxias repetidas en dos cáscaras consecutivas, ni tampoco de perder galaxias debido a sus movimientos propios. Las magnitudes absolutas de las galaxias se toman como una interpolación lineal entre dos salidas consecutivas. Una descripción completa de este procedimiento se encuentra detallada en la sección 3.3 de mi trabajo final de licenciatura (Taverna, 2015). Además, dependiendo de la cobertura angular del catálogo que se desee crear se pueden replicar los cubos en una única dirección (un octante de ángulo sólido  $\Omega = \pi/2$ , imagen derecha en la figura 2.9) o bien en todas las direcciones alrededor del punto de origen de las coordenadas (todo el cielo,  $\Omega = 4\pi$ , imagen izquierda en la figura 2.9).

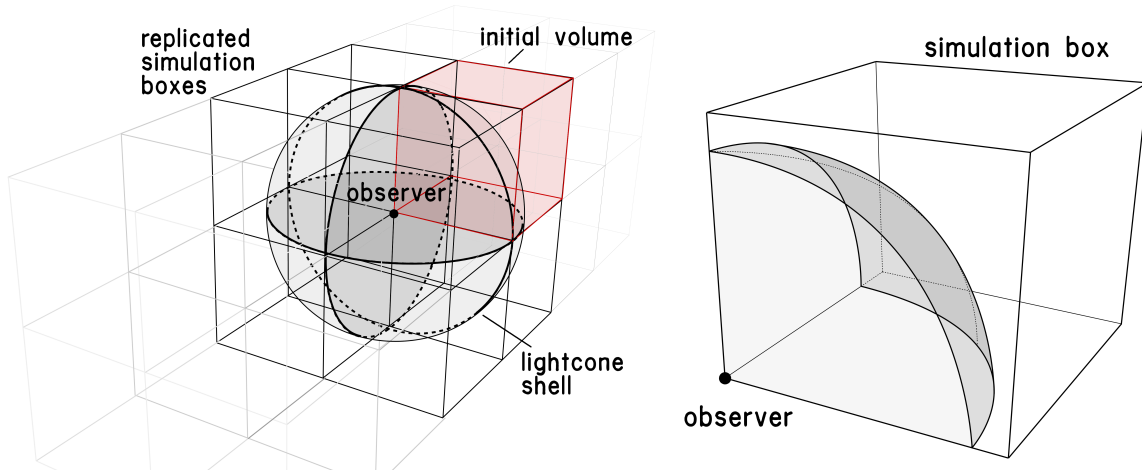
Una vez construido el supercubo de galaxias sintéticas, se pueden seleccionar galaxias dentro de una esfera o semiesfera hasta un determinado radio que se determinará en base al catálogo observacional que se desee reproducir, y para todas esas galaxias se le calcularán las propiedades observacionales de interés: posiciones angulares (coordenadas ecuatoriales), redshifts distorsionados, magnitudes aparentes observadas, etc.

A partir de las coordenadas comóviles en el supercubo,  $\vec{r} = (x, y, z)$ , se calcula el redshift cosmológico ( $z_c$ ) a través de la ecuación  $r = \frac{c}{H_0} \int_0^{z_c} \frac{dz'}{E(z')}$ , donde  $E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda}$ . Luego, con las velocidades peculiares de las partículas se obtiene el redshift distorsionado ( $z$ ) a través de

$$z = (1 + z_c) \left( 1 + \frac{v_p}{c} \right) - 1 \quad (2.1)$$

donde  $v_p = \frac{\vec{v} \cdot \vec{r}}{r}$  es la velocidad peculiar en la dirección radial y  $\vec{v}$  es el vector velocidad de las galaxias. Las coordenadas esféricas ecuatoriales absolutas ( $\alpha, \delta$ ) son calculadas a partir de las coordenadas cartesianas en el supercubo.

Dado que los SAMs proveen de las magnitudes absolutas de las galaxias, para



**Figura 2.9:** Descripción de la construcción de catálogos sintéticos a partir del cubo de una simulación (Figura 2.9 del artículo publicado por Hollowed (2019)). Panel izquierdo: caso donde el cubo de la simulación es replicado y adosado para generar un lightcone de todo el cielo ( $4\pi$  sr) en un supercubo. El cubo rojo hace referencia al cubo inicial. Panel derecho: cubo de la simulación a partir del cual se genera un catálogo sintético con un observador situado en una esquina del cubo. La cáscara gris representa el límite en distancia a la cual llega el catálogo construido. Notar que en este caso solo se tiene un octavo de la área total del cielo.

calcular las magnitudes aparentes en el sistema de referencia del observador es necesario afectar a la magnitud por corrección-K ( $K(z_i, \text{color})$ ), ya que las magnitudes aparentes observables están afectadas por la expansión del universo. Para esto, se utiliza un proceso iterativo que incluye correcciones-K calculadas por Chilingarian & Zolotukhin (2012) y que se encuentra detallado en uno de los artículos que publicamos durante el desarrollo de esta tesis (Díaz-Giménez et al., 2018b).

La magnitud aparente en el sistema de referencia del observador se calcula utilizando la fórmula del módulo de distancia:

$$m_i = M_i - 5 \log(h) + 25 + 5 \log(d_L) + K_i(z, \text{color}) \quad (2.2)$$

donde el subíndice  $i$  simboliza la banda fotométrica,  $d_L = s * (1 + z)$  es la distancia luminosidad de la galaxia en unidades de  $h^{-1}\text{Mpc}$ ,  $s$  es la distancia comóvil calculada a partir del redshift distorsionado, y  $K_i$  es la corrección K. Una vez calculada la magnitud aparente en la banda fotométrica elegida, se puede limitar el catálogo sintético en flujo para reproducir las muestras observacionales de interés.

## Capítulo 3

# Identificación de CGs en diferentes bandas fotométricas

*La gran variedad de relevamientos de galaxias construidos hasta el momento, ha proporcionado la posibilidad de hacer estudios comparativos entre los diferentes catálogos de CGs identificados en estos relevamientos. Una de las principales diferencias es que las galaxias fueron detectadas utilizando diferentes fotometrías. En particular, catálogos como Prandoni et al. (1994); McConnachie et al. (2009); Díaz-Giménez et al. (2012); Hernández-Fernández & Mendes de Oliveira (2015) y Díaz-Giménez et al. (2018b) son ejemplos de los relevamientos caracterizados por detección en diferentes bandas fotométricas: bandas B (Johnson & Morgan, 1953), R (Cousins, 1976),  $b_j$  (COSMOS-UKST, Yentis et al. 1992),  $r$  (Sloan Digital Sky Survey, SDSS, York & SDSS Collaboration 2000),  $K_s$  (Two Micron All-Sky Survey, 2MASS, Skrutskie et al. 2006a) y FUV (Galaxy Evolution Explorer, GALEX, Bianchi 2014).*

*Estos relevamientos cubren un amplio rango del espectro electromagnético, desde 1400 hasta 23000 Å, lo que implica que objetos con diferentes propiedades físicas pueden ser detectados dependiendo de la parte del espectro utilizada para su identificación. Otras diferencias entre estos relevamientos es que cuentan con diferentes funciones de selección como la magnitud aparente límite y el área que cubren en el cielo. Por otro lado, el algoritmo utilizado para identificar CGs también varía dependiendo del catálogo.*

*Por lo tanto, los catálogos de grupos compactos varían no solo en sus algoritmos de identificación y funciones de selección sino también en sus bandas fotométricas. Es por esto que es muy difícil establecer el impacto de la banda fotométrica en los catálogos de GCs.*

*Díaz-Giménez et al. (2012) compararon su propia muestra de CGs identificados en la banda K (algoritmo automático tipo Hickson mas límite en flujo) con la muestra identificada en la banda B por Focardi & Kelm (2002) (algoritmo tipo FoF), y en la banda R por Allam & Tucker (2000) (algoritmo FoF) y Hickson*

*et al. (1992) (inspección visual). Ellos encontraron que los CGs identificados en la banda R tienen tamaños proyectados más pequeños, menores separaciones proyectadas entre galaxias y menores tiempos de cruce comparados con los otros dos catálogos. También encontraron que la diferencia en magnitud entre la primera y segunda galaxia más brillante de la muestra identificada en la banda K es la más grande. Hernández-Fernández & Mendes de Oliveira (2015) también realizaron un estudio comparativo entre los grupos identificados por ellos usando fuentes ultravioletas (SFCG, algoritmo en el plano del cielo) con las muestras de CGs en la banda K (Díaz-Giménez et al., 2012), banda R (Hickson et al., 1992) y en la banda r (McConnachie et al. 2009, algoritmo automático tipo Hickson). Encontraron que los SFCGs presentan las dispersiones de velocidades más bajas (y masas viriales), mientras que los CGs en R presentan las separaciones proyectadas entre galaxias y tiempos de cruce más pequeños. Sin embargo, dados los distintos procesos de identificación, es difícil conocer cuál es la naturaleza de esas diferencias. Es por esta razón que en este capítulo realizaremos una comparación directa para discernir claramente cuál es la influencia de la identificación de CGs en diferentes bandas fotométricas.*

### **3.1. Construcción de las muestras de CGs**

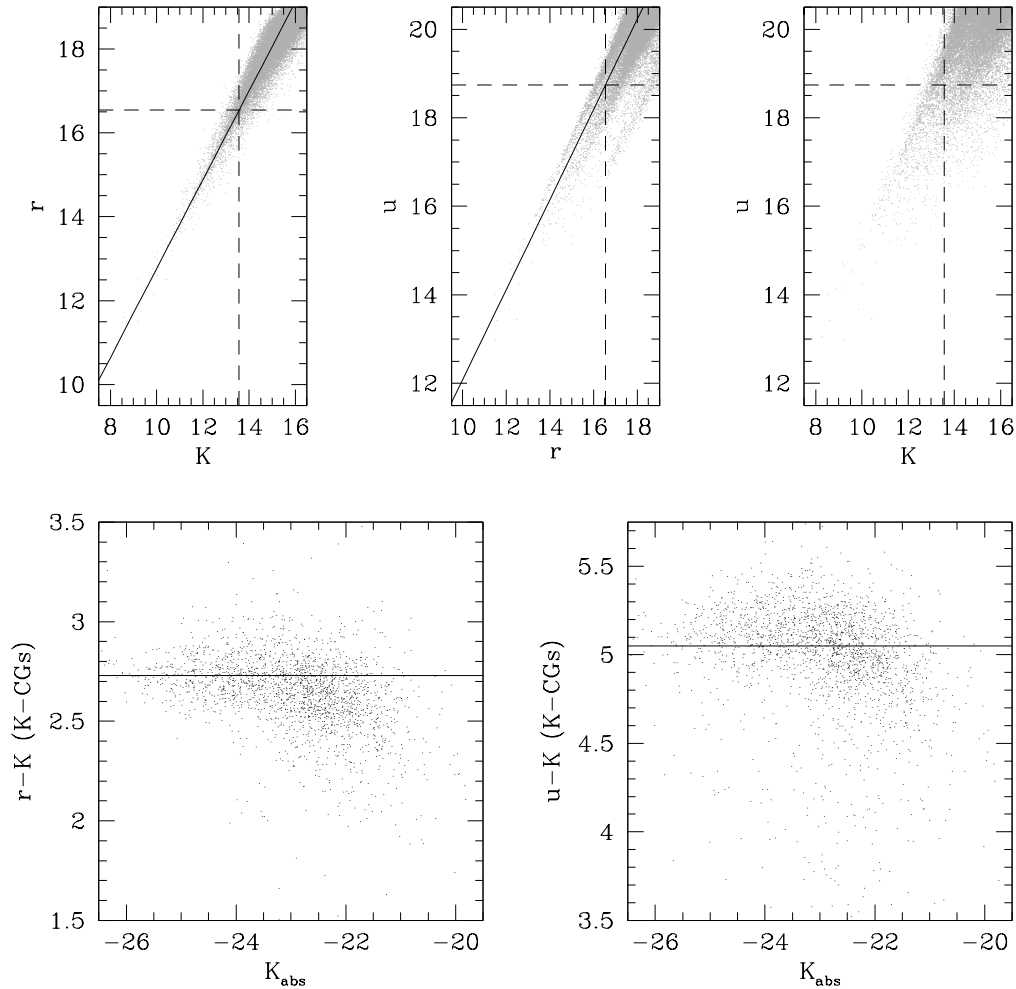
*Para emprender nuestro estudio sobre CGs identificados en diferentes bandas fotométricas, hemos usado un catálogo sintético de galaxias con evolución construido por Henriques et al. (2012) para identificar y comparar los CGs identificados en 3 bandas: K (2MASS), r (SDSS), y u (SDSS). Hemos aplicado la misma función selección (magnitud aparente límite, área del cielo) en las 3 bandas y el mismo algoritmo de identificación. Esto nos ha permitido estandarizar las condiciones de identificación y hemos podido estudiar si la causa de las diferencias observadas entre propiedades de CGs es la aplicación de diferentes algoritmos y restricciones al identificar, ó si los CGs identificados en diferentes bandas fotométricas son intrínsecamente diferentes.*

#### **3.1.1. Catálogos sintéticos**

*Identificamos grupos compactos en el catálogo sintético con evolución (“lightcone”) disponible públicamente construido por Henriques et al. (2012).<sup>1</sup> Estos autores construyeron dicho catálogo usando el SAM de formación de galaxias de Guo et al. (2011) aplicado a la simulación Millennium I, y seleccionaron galaxias de diferentes salidas de la simulación para incluir evolución de estructuras y propiedades de las galaxias dependientes del estado evolutivo, como se explicó en el capítulo 2.*

---

<sup>1</sup>Lightcone de galaxias sintéticas disponible como tabla wmap1.BC03.AllSky\_001 en <http://www.mpa-garching.mpg.de/millennium/>



**Figura 3.1:** Paneles superiores: magnitudes de las galaxias del catálogo lightcone principal (fueron graficados solo 0.05% de los puntos originales). Líneas continuas representan el ajuste lineal de los datos, mientras que las líneas discontinuas representan el límite en magnitud adoptado en cada banda. Paneles inferiores: diferencia entre las magnitudes aparentes en el sistema en reposo en dos bandas fotométricas diferentes de galaxias en CGs identificados en la banda  $K$  (color) vs la magnitud en la banda  $K$ . Las líneas sólidas representan los valores adoptados para determinar el corrimiento de cada banda para calcular el brillo superficial límite.

Los catálogos sintéticos originales de Henriques et al. (2012) son limitados en una magnitud aparente  $AB$  en el sistema en reposo de  $i < 21,0$  e incluye magnitudes aparentes para 9 filtros: SDSS  $u, g, r, i, z$  y VISTA  $Y, J, H, K_s$ . En este trabajo nos enfocaremos en tres bandas:  $K_s^2, r, y u$ . Para imitar las muestras observacionales de CGs identificados en trabajos anteriores en el catálogo 2MASS (Díaz-Giménez et al., 2012), hemos convertido la magnitud aparente en la banda  $K(AB)$  disponible en el catálogo sintético de galaxias desde el sistema  $AB$  al sistema Vega para que las magnitudes sean acordes al 2MASS:  $K(\text{Vega}) = K(AB) - 1,85$  (Cohen et al., 2003; Targett et al., 2012), y seleccionamos galaxias por debajo de una magnitud aparente límite de  $K = 13,57$ . La muestra contiene 701449 galaxias dentro de un ángulo sólido de  $4\pi$ . La densidad numérica de este catálogo sintético reproduce de muy buena manera la densidad numérica observada en el catálogo 2MASS.

También hemos seleccionado catálogos de galaxias sintéticas en la banda  $u$  y en la banda  $r$ . El límite en magnitud aparente impuesto en cada banda fue determinado con el fin de imitar el límite en la banda  $K$ , y de esta manera mantener la misma función de selección para los demás catálogos.

Para ello, hemos examinado la distribución de magnitudes en las tres bandas. En los paneles superiores de la figura 3.1 mostramos las distribuciones de las magnitudes de las galaxias:  $K$  vs.  $r$  (panel izquierdo),  $u$  vs.  $r$  (panel central), y  $K$  vs.  $u$  (panel derecho). Utilizando un ajuste lineal para  $K$  vs.  $r$ , hemos determinado que la magnitud aparente límite en la banda  $r$  correspondiente a la magnitud  $K_{\text{lim}} = 13,57$  es  $r_{\text{lim}} = 16,54$ . Para determinar el límite en la banda  $u$  hemos ajustado la distribución  $u$  vs.  $r$  ya que la distribución  $u$  vs.  $K$  tiene mayor dispersión, y nos ha permitido realizar un mejor ajuste. Para  $r_{\text{lim}} = 16,54$ , hemos encontrado  $u_{\text{lim}} = 18,73$ . La adopción de un límite de magnitud fijo en una  $u$  otra banda, inevitablemente conduce a un diferente muestreo de galaxias hacia las magnitudes más débiles. Este es un hecho que también está presente en los catálogos observacionales limitados por el flujo, como en muchos de los catálogos utilizados en la literatura para identificar CGs en diferentes bandas. El número de galaxias incluidas en cada catálogo sintético se muestra en la tabla 3.1.

### 3.1.2. Identificación de CGs

Hemos identificado CGs en un catálogo lightcone para 3 bandas fotométricas  $K, r$  y  $u$ , usando el criterio definido por Díaz-Giménez et al. (2012):

<sup>2</sup>De aquí en adelante nos referiremos a esta magnitud sólo con  $K$  en vez de  $K_s$ .



	K	r	u
$m_{\text{lim}}$	13,57	16,54	18,73
# galaxias	701449	677351	928367
$\langle z \rangle$	0,065	0,057	0,063
$m_{\text{bri}_{\text{lim}}}$	10,57	13,54	15,73
Corrimiento de banda	...	$r = K + 2,73$	$u = K + 5,05$
$\mu_{\text{lim}}$	23,6	26,33	28,65
$M_{\odot}^{(1)}$	3,29	4,65	6,44
$\alpha^{(2)}$	0,11	0,11	0,12
$\beta^{(2)}$	0,72	0,78	0,69
$\gamma^{(2)}$	0,09	0,12	0,11
$\mathcal{M}_0^{(2)}$	$1,57 \times 10^{10}$	$2,25 \times 10^{10}$	$1,70 \times 10^{10}$
$a^{(2)}$	$20,45 \times 10^{-3}$	$37,24 \times 10^{-3}$	$23,75 \times 10^{-3}$
$b^{(2)}$	0,22	0,20	0,23

**Tabla 3.1:** Principales parámetros usados para la selección de galaxias en el catálogo sintético, identificación de CGs en diferentes bandas fotométricas y asignación de tamaños para las partículas simuladas.

**Nota.** (1) [www.ucolick.org/~cnaw/sun.html](http://www.ucolick.org/~cnaw/sun.html); (2) El cálculo del radio a mitad de luz ( $R_i$ ) en función de la masa estelar ( $\mathcal{M}_i^*$ ) fue realizado utilizando las recetas de Lange et al. (2015). Galaxias elípticas:  $R_i = \gamma (\mathcal{M}_i^*)^\alpha (1 + \mathcal{M}_i^*/\mathcal{M}_0)^{(\beta-\alpha)}$ . Galaxias no elípticas:  $R_i = a(\mathcal{M}_i^*)^b$ .

**Población** :  $4 \leq N \leq 10$

**Compacticidad** :  $\mu \leq \mu_{\text{lim}} [\text{mag}/\text{arcsec}^{-2}]$

**Aislamiento** :  $\Theta_N > 3\Theta_G$

**Límite en flujo** :  $m_{\text{bri}} \leq m_{\text{lim}} - 3$

**Filtro en velocidad** :  $|v_i - \langle v \rangle| \leq 1000 \text{ km s}^{-1}$

donde  $N$  es el número de galaxias cuyas magnitudes en la banda seleccionada están dentro de un rango de 3 magnitudes desde la galaxia brillante;  $m_{\text{bri}}$  es la magnitud aparente de la galaxia más brillante del grupo;  $\mu$  es el brillo superficial medio en una dada banda, promediado sobre el círculo más pequeño que circunscribe los centros de las galaxias;  $\Theta_G$  es el diámetro angular del círculo circunscrito más pequeño;  $\Theta_N$  es el diámetro angular del círculo concéntrico más grande que no contiene otras galaxias dentro del rango de magnitudes seleccionado o más brillantes;  $v_i$  es la velocidad radial de cada galaxia miembro, y  $\langle v \rangle$  es la velocidad radial de las galaxias miembros. Vale la pena aclarar que el algoritmo de identificación busca los CGs en proyección (es decir que cumplan los primeros 4 ítems antes detallados) y luego corrobora si cumplen con el filtro en velocidad utilizando la información proveniente de los redshifts. Es decir, el algoritmo trata de imitar el procedimiento observacional realizado por Hickson.

En la tabla 3.1 se muestran los límites adoptados en cada banda para los criterios de compacticidad y límite en flujo. En la banda  $K$  hemos adoptado los valores descritos por Díaz-Giménez et al. (2012), y para determinar el valor límite del criterio de compacticidad en las otras dos bandas, al igual que Díaz-Giménez et al. (2012), hemos examinado la muestra resultante de CGs identificada en la banda  $K$  y determinamos el corrimiento medio en la banda:  $(r - K)^0 = 2,73$  y  $(u - K)^0 = 5,05$ . En las paneles inferiores de la figura 3.1, se observan las distribuciones de  $r - K$  y  $u - K$ , las cuales reflejan el corrimiento de las bandas  $r$  y  $u$  con respecto a la banda  $K$ . Estos corrimientos se muestran en líneas sólidas. Usando estos valores, hemos corregido el valor de  $\mu$  correspondiente a la banda  $K$  para obtener los valores correspondientes en las otras dos bandas.

Una vez identificados los CGs (siguiendo el criterio antes mencionado), estudiamos si estaban embebidos dentro un CGs más grande que también cumple con el criterio descrito, es decir, grupos compactos dentro de grupos compactos. Utilizando la misma metodología que Díaz-Giménez & Mamon (2010), para cada grupo, nos quedamos con el grupo más grande (para una descripción del algoritmo utilizado, ver figura 1 en el artículo realizado por Díaz-Giménez & Mamon (2010)).

También hemos considerado el hecho de que las galaxias en el catálogo sintético son partículas puntuales y hemos adoptado un criterio para asignar un tamaño y discernir cuando dos galaxias se encuentran muy cerca en proyección en el plano del cielo de tal manera que serían observadas como una sola galaxia. Esto es necesario ya que modifica el número de objetos detectables cambiando la población de los grupos compactos. En función del tipo morfológico de cada galaxia —el cual se determina usando la relación masa estelar del bulbo y la masa estelar total (Bertone et al., 2007)—, hemos calculado su radio a mitad de luz en cada banda como función de la masa estelar de cada galaxia sintética siguiendo las prescripciones de Lange et al. (2015) (ver tabla 3.1). Finalmente, consideramos a dos galaxias como una si la separación angular entre las dos galaxias es más pequeña que la suma de sus radios angulares a mitad de luz.

Dentro del un ángulo sólido de  $4\pi$ , hemos identificado 447 CGs en la banda  $K$  ( $K$ -CGs), 406 en la banda  $r$  ( $r$ -CGs), y 276 en la banda  $u$  ( $u$ -CGs).

### 3.1.3. Propiedades de los CGs

Hemos medido varias propiedades de los grupos que las usaremos para comparar las tres muestras de CG: el diámetro angular del círculo mínimo que encierra todos los miembros ( $\theta_G$ ); el radio proyectado del círculo mínimo ( $r_p$ ); el radio virial proyectado de los CGs ( $R_{\text{vir}}$ ); la mediana de la separación proyectada entre galaxias ( $\langle d_{ij} \rangle$ ); la máxima separación proyectada entre las 4 galaxias más cercanas ( $s_{\perp}$ ); la máxima separación en la línea de la visual entre las 4 galaxias más cercanas ( $s_{\parallel}$ ); la máxima separación 3D interpartícula entre las 4 galaxias más cercanas ( $s_4$ ); la dispersión de velocidades radiales calculadas con el método

Muestras # CGs	Muestras totales			Muestras restringidas		
	K 447	r 406	u 276	$\tilde{K}$ 382	$\tilde{r}$ 289	$\tilde{u}$ 231
$\theta_G$ [arcmin]	3,2 (0,2)	3,4 (0,2)	3,5 (0,3)	3,3 (0,2)	3,7 (0,2)	3,5 (0,3)
$r_p$ [kpc h <sup>-1</sup> ]	67 (3)	72 (3)	66 (4)	65 (3)	73 (4)	62 (4)
$R_{\text{vir}}$ [kpc h <sup>-1</sup> ]	90 (5)	106 (5)	96 (7)	86 (5)	102 (6)	93 (7)
$\langle d_{ij} \rangle$ [kpc h <sup>-1</sup> ]	73 (4)	82 (4)	73 (5)	72 (4)	81 (4)	70 (4)
$s_{\perp}$ [kpc h <sup>-1</sup> ]	105 (6)	117 (6)	113 (8)	102 (6)	114 (8)	101 (7)
$s_{\parallel}$ [kpc h <sup>-1</sup> ]	103 (14)	119 (17)	156 (31)	99 (14)	117 (19)	140 (28)
$s_4$ [kpc h <sup>-1</sup> ]	168 (15)	182 (16)	205 (27)	162 (15)	180 (19)	187 (24)
$\sigma_v$ [km s <sup>-1</sup> ]	280 (14)	281 (15)	253 (21)	270 (15)	284 (18)	265 (22)
$H_0 t_{\text{cr}}$ [10 <sup>-2</sup> ]	2,7 (0,2)	2,9 (0,3)	2,9 (0,4)	2,7 (0,2)	2,8 (0,3)	2,6 (0,3)
$\mathcal{M}_{\text{vir}}$ [10 <sup>12</sup> $\mathcal{M}_{\odot}$ h <sup>-1</sup> ]	6,9 (0,9)	7,3 (1,1)	6,5 (1,4)	6,3 (0,9)	7,8 (1,4)	6,5 (1,4)
$K_{\text{bri}}$	10,0 (0,1)	10,2 (0,1)	9,9 (0,1)	9,9 (0,1)	9,9 (0,1)	9,8 (0,1)
$r_{\text{bri}}$	12,8 (0,1)	13,0 (0,1)	12,6 (0,1)	12,7 (0,1)	12,7 (0,1)	12,6 (0,1)
$u_{\text{bri}}$	15,2 (0,1)	15,3 (0,1)	15,0 (0,1)	15,1 (0,1)	15,0 (0,1)	14,9 (0,1)
$\Delta K_{12}$	1,3 (0,1)	1,2 (0,1)	1,1 (0,1)	1,3 (0,1)	1,2 (0,1)	1,1 (0,1)
$\Delta r_{12}$	1,3 (0,1)	1,2 (0,1)	1,0 (0,1)	1,3 (0,1)	1,2 (0,1)	1,1 (0,1)
$\Delta u_{12}$	1,3 (0,1)	1,2 (0,1)	1,1 (0,1)	1,3 (0,1)	1,2 (0,1)	1,1 (0,1)
$\mu_K$ [mag/arcsec <sup>2</sup> ]	21,9 (0,1)	22,2 (0,1)	22,1 (0,1)	21,9 (0,1)	22,2 (0,1)	21,8 (0,1)
$\mu_r$ [mag/arcsec <sup>2</sup> ]	24,8 (0,1)	25,1 (0,1)	24,9 (0,1)	24,7 (0,1)	25,1 (0,1)	24,7 (0,1)
$\mu_u$ [mag/arcsec <sup>2</sup> ]	27,1 (0,1)	27,4 (0,1)	27,2 (0,1)	27,1 (0,1)	27,4 (0,1)	27,0 (0,1)
$L_K$ [10 <sup>9</sup> $L_{\odot}$ h <sup>-2</sup> ]	231 (14)	219 (15)	194 (18)	227 (14)	224 (19)	207 (18)
$L_r$ [10 <sup>9</sup> $L_{\odot}$ h <sup>-2</sup> ]	68 (4)	66 (4)	59 (4)	67 (4)	67 (5)	61 (5)
$L_u$ [10 <sup>9</sup> $L_{\odot}$ h <sup>-2</sup> ]	41 (3)	41 (3)	44 (4)	41 (3)	42 (4)	41 (3)
$\mathcal{M}_v/L_K$ [h $\mathcal{M}_{\odot}/L_{\odot}$ ]	29 (3)	32 (4)	32 (6)	27 (3)	30 (5)	30 (5)
$\mathcal{M}_v/L_r$ [h $\mathcal{M}_{\odot}/L_{\odot}$ ]	100 (11)	109 (14)	106 (19)	94 (12)	104 (17)	103 (19)
$\mathcal{M}_v/L_u$ [h $\mathcal{M}_{\odot}/L_{\odot}$ ]	158 (20)	178 (24)	161 (27)	136 (20)	174 (30)	167 (30)
$T1_K$	0,56 (0,03)	0,56 (0,03)	0,75 (0,04)	0,54 (0,02)	0,54 (0,03)	0,65 (0,04)
$T2_K$	0,61 (0,02)	0,61 (0,02)	0,68 (0,03)	0,61 (0,02)	0,61 (0,03)	0,65 (0,03)
$T1_r$	0,56 (0,02)	0,57 (0,02)	0,72 (0,04)	0,54 (0,02)	0,56 (0,03)	0,64 (0,03)
$T2_r$	0,61 (0,02)	0,61 (0,02)	0,67 (0,03)	0,59 (0,02)	0,60 (0,03)	0,63 (0,03)
$T1_u$	0,63 (0,03)	0,66 (0,03)	0,71 (0,04)	0,61 (0,03)	0,65 (0,03)	0,68 (0,04)
$T2_u$	0,67 (0,02)	0,71 (0,03)	0,69 (0,03)	0,66 (0,03)	0,70 (0,03)	0,66 (0,03)
% <i>Reals</i>	55.6	52.0	44.4	57.9	52.2	48.9

**Tabla 3.2:** Propiedades de los CGs para las muestras totales y restringidas. En cada celda, el formato  $xx$  ( $ss$ ) contiene la mediana ( $xx$ ) y el error ( $ss$ ) correspondiente aproximadamente al 95% del intervalo de confianza,  $CI = xx \pm ss$ , excepto para los valores  $T_1$  y  $T_2$  donde la cantidad entre paréntesis son las barras de error computadas usando la técnica de remuestreo bootstrap.

*gapper* ( $\sigma_v$ ) (Beers et al., 1990); el tiempo de cruce adimensional ( $H_0 t_{cr}$ ); la masa virial  $M_{vir}$ ; las magnitudes en el sistema de referencia del observador de la galaxia más brillante en las 3 bandas ( $K_{bri}, r_{bri}, u_{bri}$ ); la diferencia en magnitud absoluta entre la primera y la segunda galaxias más brillantes en el sistema de referencia en reposo ( $\Delta K_{12}, \Delta r_{12}, \Delta u_{12}$ ); el brillo superficial del grupo en las 3 bandas ( $\mu_K, \mu_r, \mu_u$ ); la luminosidad total del grupo en las 3 bandas ( $L_K, L_r, L_u$ ); y la relación masa-luminosidad en las 3 bandas ( $M_v/L_K, M_v/L_r, M_v/L_u$ ). También hemos calculado para cada muestra y cada banda el estadístico de Tremaine-Richstone, T1 y T2 (Tremaine & Richstone, 1977). Grupos con la galaxia más brillante mucho más brillante que la segunda en brillo, tienen valores de T1 o T2 más bajos que la unidad. En concordancia con Mamon (1987), colisiones dentro del grupo toman valores de T1 y T2 menores a 0.7. Las fórmulas que se han utilizado para calcular todas las propiedades anteriores se encuentran en el apéndice B.

Finalmente, dividimos las muestras de CGs en grupos físicamente densos (*Reals*) y alineamientos aleatorios (CAs) siguiendo la clasificación 3D realizada por Díaz-Giménez & Mamon (2010). Esta clasificación utiliza la propiedad “ $s_4$ ” para definir el tamaño 3D y “ $s_{\perp}$ ” y “ $s_{\parallel}$ ” para definir la elongación de los grupos a lo largo de la línea de la visual ( $s_{\parallel}/s_{\perp}$ ). Recordar que estas distancias son mediciones en el espacio real.

Según esta definición, un grupo se considera ‘Real’ si :

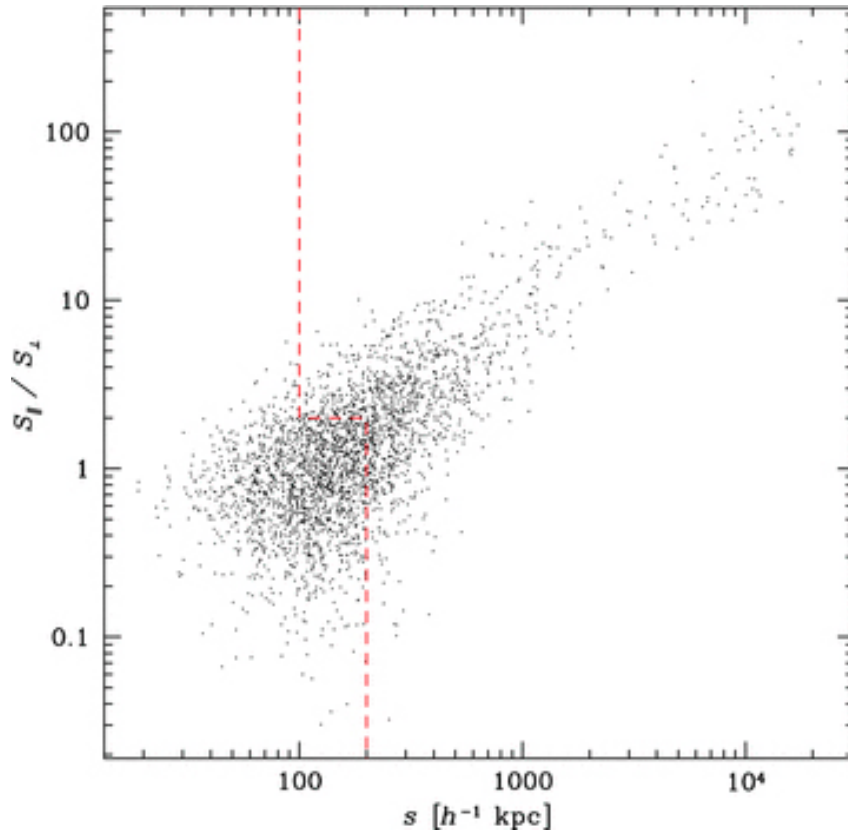
$$s_4 < 100 h^{-1}kpc, \quad \text{ó}$$

$$s_4 < 200 h^{-1}kpc \quad \text{y} \quad s_{\parallel}/s_{\perp} < 2.$$

A modo de ejemplo, en la figura 3.2 se muestra la región definida por este criterio tal cual se realizó en el trabajo de Díaz-Giménez & Mamon (2010). Es importante aclarar que esta manera de definir si un sistema puede ser considerado denso físicamente fue realizada por Díaz-Giménez & Mamon (2010) usando como base un catálogo sintético de CGs identificados con el criterio tipo Hickson en el espacio de redshifts.

La mediana de esas propiedades y los porcentajes de *Reals* se muestran en las primeras 3 columnas de la tabla 3.2. El intervalo de confianza (IC) del 95% para la mediana está dado por  $1,58 \times IQR/\sqrt{n}$ , donde  $n$  es el número de objetos en la muestra e *IQR* es el rango intercuartil (Krzywinski & Altman, 2014).

Por otro lado, usando el test no paramétrico Kolmogorov-Smirnov, hemos medido la probabilidad de que muestras puedan considerarse que provienen de la misma distribución. En la tabla 3.3, se muestran los “valores  $p$ ” obtenidos por el test cuando comparamos pares de muestras. Las celdas sombreadas remarcaban cuando la probabilidad es menor al valor crítico típico de 0.05, lo cual indica que existen diferencias estadísticas para las distribuciones de una dada propiedad entre las dos muestras. Le hemos llamado *K*-CGs a los CGs identificados en la banda *K*, y *r*-CGs y *u*-CGs a los CGs identificados en la banda *r* y en la banda



**Figura 3.2:** Elongación a lo largo de la línea de la visual  $s_{\parallel}/s_{\perp}$  en función de la separación máxima 3D entre las 4 galaxias más cercanas (llamada  $s$  en la figura, análogo al  $s_4$  utilizado en esta tesis) de los CGs extraídos del SAM De Lucia & Blaizot (2007). La zona delimitada a la izquierda de las líneas a rayas rojas marcan la región donde se encuentran los CGs considerados *Reals*. Figura 6 del artículo realizado por Díaz-Giménez & Mamon (2010).

*u* respectivamente. Analizando los valores de las medianas (tabla 3.2) y la comparación entre las propiedades en diferentes bandas fotométricas (tabla 3.3) hemos encontrado:

- Los *K*-CGs tienen tamaños proyectados ( $r_p$ ,  $R_{\text{vir}}$ ,  $d_{ij}$ ) estadísticamente más pequeños que los *r*-CGs.
- Los *K*-CGs muestran una separación 3D entre partículas ( $s_4$ ) más pequeña que los *u*-CGs.
- Los *u*-CGs presentan la dispersión de velocidades más baja (aunque la diferencia es estadísticamente insignificante)
- $\Delta M_{12}$  en las 3 bandas es más grande para los *K*-CGs que para los *u*-CGs.
- Los *K*-CGs muestran mayor compacticidad (bajo brillo superficial) en las 3 bandas que los *r*-CGs, y valores similares a los *u*-CGs (excepto para  $\mu_u$ ).

- la luminosidad en la banda- $u$  es similar para CGs identificados en cualquiera de las 3 bandas. Sin embargo, las luminosidades en la banda  $K$  y en la banda  $r$  de los  $K$ -CGs son más altas que las luminosidades de los  $u$ -CGs.
- las galaxias más brillantes de los  $u$ -CGs son más brillantes que las galaxias más brillantes de los  $K$ - o  $r$ -CGs en las tres bandas.
- Los valores  $T_1$  y  $T_2$  calculados en las bandas  $K$  y  $r$  muestran valores muy bajos para las muestras  $K$ -CGs y  $r$ -CGs lo cual significa valores muy diferentes entre la primer y segunda galaxia con mayor luminosidad, indicando señales claras de colisiones dentro de los grupos. En cambio los  $u$ -CGs muestran valores cercanos a 0.7.
- la muestra  $K$ -CG contiene más CGs *Reals* que las otras bandas fotométricas. El porcentaje más bajo de *Reals* es obtenido en la muestra  $u$ -CG.

Díaz-Giménez et al. (2012) y Hernández-Fernández & Mendes de Oliveira (2015) compararon CGs identificados en diferentes bandas y con diferentes algoritmos. Estudiando los tamaños proyectados, ambos estudios están en acuerdo en que los grupos identificados en la banda  $R$  (Hickson et al., 1992; Allam & Tucker, 2000) o la banda  $r$  (McConnachie et al., 2009) tienen tamaños proyectados más pequeños que los grupos identificados en la banda  $K$  (Díaz-Giménez et al., 2012), y por otro lado, Hernández-Fernández & Mendes de Oliveira reportaron que sus CGs identificados en la banda  $UV$  son los más pequeños. En este trabajo, hemos encontrado un resultado diferente:  $K$ -CGs y  $u$ -CGs son más pequeños que los  $r$ -CGs. Este resultado puede ser esperado debido a que las diferencias en los algoritmos usados para identificar las muestras observacionales que han sido comparadas introducen las diferencias en tamaños que se han reportado en la literatura. Usando algoritmos *friends-of-friends*, o criterios tipo Hickson basados en los tamaños proyectados e incluyendo o no las opciones CG-en-CG (grupos compactos dentro de un grupo compacto más grande) se introduce un sesgo en los tamaños proyectados de los grupos resultantes. En este trabajo, hemos evitado este problema utilizando el mismo algoritmo en las tres bandas fotométricas, por lo tanto los resultados reportados aquí solo se deben a diferencias intrínsecas entre bandas.

También, Díaz-Giménez et al. (2012) y Hernández-Fernández & Mendes de Oliveira (2015) encontraron que los tiempos de cruce de los grupos identificados en las bandas  $B$  y  $UV$  son más grandes que los  $K$ -CGs los cuales son más grandes que los  $R$ - o  $r$ -CGs. No hemos encontrado diferencias significativas entre los tiempos de cruce de los grupos identificados en las 3 bandas. Además, Díaz-Giménez et al. (2012) encontraron que la separación media proyectada interpartícula entre galaxias en  $R$ -CGs es más pequeña que en  $K$ -CGs, las cuales son más pequeñas que los  $B$ -CGs. Sin embargo, Hernández-Fernández & Mendes de Oliveira (2015) encontraron que la separación entre partículas más pequeña estaba en los  $UV$ -CG, y los  $r$ -CG presentaban las mayores separaciones. En este trabajo hemos encontrado que los  $r$ -CG son los que tiene la mayor

	Totales			Restringidas			Totales vs Restringidas		
	K - r	K - u	u - r	$\tilde{K} - \tilde{r}$	$\tilde{K} - \tilde{u}$	$\tilde{u} - \tilde{r}$	K - $\tilde{K}$	r - $\tilde{r}$	u - $\tilde{u}$
$\theta_G$	0,16	0,08	0,32	$4 \times 10^{-3}$	0,72	0,19	1,00	0,11	1,00
$r_p$	0,02	0,99	0,13	$7 \times 10^{-3}$	0,56	$1 \times 10^{-3}$	1,00	1,00	0,51
$R_{\text{vir}}$	$1 \times 10^{-4}$	0,06	0,19	$3 \times 10^{-5}$	0,18	0,05	0,97	1,00	0,66
$\langle d_{ij} \rangle$	0,01	0,42	0,12	$2 \times 10^{-3}$	0,47	$4 \times 10^{-3}$	0,98	1,00	0,31
$s_{\perp}$	0,02	0,23	0,87	0,04	0,57	0,14	0,99	1,00	0,29
$s_{\parallel}$	0,56	$5 \times 10^{-4}$	0,02	0,68	$9 \times 10^{-3}$	0,26	1,00	1,00	0,79
$s_4$	0,26	$3 \times 10^{-3}$	0,06	0,24	0,10	0,75	1,00	1,00	0,68
$\sigma_v$	1,00	0,23	0,24	0,72	0,94	0,66	0,99	1,00	1,00
$H_0 t_{\text{cr}}$	0,19	0,09	0,88	0,41	0,96	0,36	1,00	1,00	0,55
$\mathcal{M}_{\text{vir}}$	0,45	0,40	0,27	0,08	0,78	0,70	0,91	1,00	1,00
$K_{\text{bri}}$	$1 \times 10^{-5}$	0,01	$1 \times 10^{-4}$	0,93	0,29	0,26	0,10	$1 \times 10^{-7}$	0,03
$r_{\text{bri}}$	$1 \times 10^{-4}$	0,37	$1 \times 10^{-4}$	0,95	0,25	0,37	0,03	$1 \times 10^{-8}$	0,06
$u_{\text{bri}}$	$8 \times 10^{-3}$	$2 \times 10^{-3}$	$1 \times 10^{-5}$	0,99	0,20	0,60	0,01	$6 \times 10^{-8}$	1,00
$\Delta K_{12}$	0,04	$1 \times 10^{-3}$	0,18	0,24	0,03	0,71	1,00	0,96	0,96
$\Delta r_{12}$	0,08	$1 \times 10^{-3}$	0,25	0,35	0,01	0,64	1,00	1,00	1,00
$\Delta u_{12}$	0,11	0,03	0,99	0,35	0,01	0,64	1,00	1,00	1,00
$\mu_K$	$4 \times 10^{-3}$	0,10	0,20	$7 \times 10^{-3}$	0,72	$1 \times 10^{-3}$	0,97	1,00	0,19
$\mu_r$	$4 \times 10^{-3}$	0,19	0,13	$5 \times 10^{-3}$	0,39	$1 \times 10^{-3}$	0,94	1,00	0,23
$\mu_u$	0,01	0,50	0,04	0,01	0,26	$1 \times 10^{-3}$	0,86	1,00	0,59
$L_K$	0,73	$8 \times 10^{-3}$	0,14	0,87	0,24	0,56	1,00	1,00	0,75
$L_r$	0,75	0,03	0,17	0,89	0,50	0,51	1,00	1,00	1,00
$L_u$	1,00	0,79	0,73	0,79	0,99	0,80	1,00	1,00	0,96
$\mathcal{M}_v/L_K$	0,29	0,11	0,76	0,10	0,19	0,97	0,99	1,00	0,97
$\mathcal{M}_v/L_r$	0,34	0,20	0,91	0,12	0,29	1,00	0,99	1,00	1,00
$\mathcal{M}_v/L_u$	0,39	0,73	0,51	0,21	0,48	0,91	0,97	1,00	1,00

**Tabla 3.3:** “Valores p” del test Kolmogorov-Smirnov para dos muestras. Cada columna combina diferentes pares de muestras de CGs: las primeras 3 columnas comparan las muestras totales, las siguientes tres columnas comparan las muestras restringidas, mientras que las últimas tres columnas comparan las muestras totales contra las muestras restringidas en cada banda. Resaltamos en color gris las celdas donde los “valores p” son menores que 0.05.

separación entre partículas de acuerdo con Hernández-Fernández & Mendes de Oliveira (2015). Finalmente, Díaz-Giménez et al. (2012) encontraron que la única muestra que presenta valores de T1 y T2 por debajo de la unidad fueron los identificados en la banda K. Por el contrario, en este trabajo usando el mismo algoritmo en las tres bandas, encontramos que todos los valores están por debajo de la unidad, siendo más altos en la banda  $u$  y muy similares en las bandas  $K$  y  $r$ .

### 3.1.4. Muestras Restringidas

Dadas las diferencias que encontramos en la sección anterior surge la pregunta si estas diferencias podrían estar sesgadas debido al hecho de que cada muestra ha sido seleccionada con criterios levemente diferentes dependiendo de la banda ( $\mu_{\text{lim}}$  y  $m_{\text{bri}_{\text{lim}}}$ ). Por lo tanto, en esta sección normalizamos los criterios para evitar cualquier dependencia de la función de selección. Hemos definido las muestras “restringidas” ( $\tilde{K}$ -,  $\tilde{r}$ - y  $\tilde{u}$ -CGs) seleccionando los grupos en cada catálogo de tal manera que satisfagan los criterios de compactidad y flujo en las otras dos bandas. Por ejemplo,  $\tilde{K}$ -CGs son una submuestra de los  $K$ -CGs que también satisfacen que  $\mu_r < 26,33$ , y  $\mu_u < 28,65$ , y  $r_{\text{brightest}} < 13,54$ , y  $u_{\text{brightest}} < 15,73$ . De manera similar hemos seleccionado los  $\tilde{r}$ -CGs y  $\tilde{u}$ -CGs.

Encontramos que el 85.5% de los  $K$ -CGs sobreviven la triple restricción, mientras que el 71.2% de los  $r$ -CGs y 83.7% de los  $u$ -CGs conforman sus muestras restringidas. En la tabla 3.2 presentamos las medianas de las propiedades de estas muestras restringidas y sus intervalos de confianza del 95%, mientras que en la tabla 3.3 presentamos los valores  $p$  obtenidos de aplicar el test KS cuando se comparan las muestras restringidas.

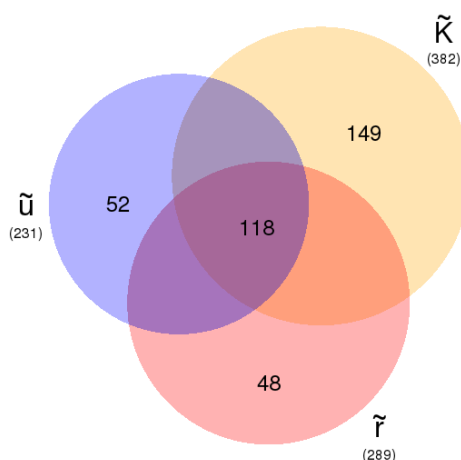
En general, las muestras restringidas solo difieren de la muestra original en esas propiedades que dependían explícitamente de los límites adoptados en las magnitudes (ver las últimas tres columnas de la tabla 3.3). Como consecuencia, las diferencias encontradas entre las muestras originales en las magnitudes de las galaxias brillantes y las luminosidades totales de los grupos ya no están presentes una vez que impusimos los mismos límites para los tres catálogos. Sin embargo, las diferencias en tamaños, gaps en magnitud y brillo superficial permanecen en las muestras restringidas:  $\tilde{u}$ -CGs son más pequeños en proyección, los  $K$ -CGs tienen la diferencia más grande en magnitud (aunque no significativos comparados con los  $\tilde{r}$ -CGs), mientras que los  $\tilde{r}$ -CGs tienen el brillo superficial más grande (menor compactidad) y este resultado es más importante cuando usamos muestras restringidas.

Tener muestras restringidas en las 3 bandas al mismo tiempo ha incrementado levemente la probabilidad de seleccionar grupos físicamente densos (*Reals*) en las bandas  $K$  y  $u$ . Esto es bueno ya que en una muestra observacional no podemos distinguir fácilmente entre alineamientos aleatorios y grupos físicamente densos.



### 3.2. Grupos compactos Puros y Comunes

A pesar de que los grupos compactos pertenecientes a las muestras restringidas cumplen con los criterios de compacticidad y de flujo en las tres bandas al mismo tiempo, esos grupos no son los mismos. En la figura 3.3 mostramos el diagrama de Venn de los tres conjuntos de grupos restringidos ( $\tilde{K}$ ,  $\tilde{r}$ ,  $\tilde{u}$ -CGs). En esta sección nos interesa analizar los CG que se pueden identificar en los tres bandas al mismo tiempo y aquellos CG que pueden identificarse solo en una banda en particular.



**Figura 3.3:** Diagrama de Venn esquemático que muestra las diferentes regiones superpuestas que fueron obtenidas cuando comparamos los grupos identificados en cada banda fotométrica. El número de grupos en las muestras restringidas están mostrados entre paréntesis. El número de grupos identificados solo en una dada banda y los que fueron identificados en todas las bandas están citados dentro de los círculos de colores.

Para seleccionar CGs que sean identificados en las tres bandas, primero hemos buscado CGs que fueran identificados en dos bandas al mismo tiempo, y luego, hemos controlado que estos grupos también sean identificados en la tercera banda. Cuando comparamos CGs en dos bandas, pueden ser CGs que pertenecen a ambos catálogos pero no necesariamente exactamente el mismo: pueden diferir levemente en el número de miembros en las dos bandas incluyendo o excluyendo una galaxia en una o otra banda. Por lo tanto, en lugar de hacer una comparación miembro a miembro, hemos adoptado un criterio simple para determinar cuando un CG que pertenece a dos catálogos diferentes es el mismo: si la distancia angular entre los centros de los círculos mínimos es menor que dos veces el radio angular del CG y la diferencia en velocidad radial de los centros es menor que  $1000 \text{ km s}^{-1}$ . Usando este criterio hemos encontrado 118 CGs en común en las 3 bandas (C-CGs). Estos grupos representan el 31 % de los  $\tilde{K}$ -CGs,

el 41% de los  $\tilde{r}$ -CGs y el 51% de los  $\tilde{u}$ -CGs. Por otro lado, CGs en cada banda que no tienen su contra parte en ninguno de los otros dos catálogos conforman la muestra de CG "puros" ( $|K|$ -,  $|r|$ -,  $|u|$ -CGs). Hemos encontrado que el  $\sim 39\%$  de los  $\tilde{K}$ -CGs solo pueden ser identificados cuando analizamos una muestra de galaxias seleccionadas en la banda  $K$ , el  $\sim 17\%$  de los  $\tilde{r}$ -CGs existen solo en la banda  $r$ , mientras que el  $\sim 23\%$  de los  $\tilde{u}$ -CGs solo existen en la banda  $u$ .

En la figura 3.4, usando diagramas de caja hemos comparado algunas propiedades de los CGs pertenecientes a cada submuestra. En estos diagramas, las muescas corresponden al intervalo de confianza de aproximadamente el 95%. En general, cuando las muescas no se superponen, las medianas pueden ser tomadas como significativamente diferentes (Krzywinski & Altman, 2014), pero el hecho de que se superpongan no descarta una diferencia significativa.

En la tabla 3.4 mostramos los valores de las medianas y sus respectivos intervalos de confianza para todas las propiedades de grupos utilizadas en este trabajo en cada una de esas muestras, mientras que los valores  $p$  del test KS entre pares de muestras en la tabla 3.5.

Del análisis de las propiedades (ver figura 3.4 y tabla 3.4) para las diferentes muestras de CGs en diferentes bandas fotométricas podemos mencionar lo siguiente:

- Los  $|K|$ -CGs presentan los radios viriales proyectados, masas viriales, brillo superficial más pequeños (más alta compacticidad), relación masa luminosidad más pequeña y las diferencias en magnitud más grandes, incluso más grandes que todas las muestras analizadas previamente. También producen los valores más bajos de  $T1$  y  $T2$ .
- Los  $|r|$ -CGs presentan los tamaños proyectados, masas viriales, brillo superficial más grandes (más baja compacticidad) y la relación masa luminosidad más grandes.
- Los  $|u|$ -CGs presentan la galaxia más brillante y las separaciones interpartículas 3D y en la línea de la visual más grandes (lo que genera que la muestra  $|u|$  presenta el porcentaje más bajo de grupos físicamente densos acorde a nuestro criterio basado en las separaciones 3D comóviles). En promedio, estos grupos muestran uno de los más grandes valores de  $T1$  y  $T2$ .
- Los  $C$ -CGs presentan tamaños comóviles 3D, en la línea de la visual y proyectados de las 4 galaxias más cercanas más pequeños (lo que implica tener mayor fracción de grupos *Reals*).

Comparando la tabla 3.5 y la tabla 3.3, podemos analizar los efectos de incluir o excluir grupos que pertenecen a una u otra banda. En general, las diferencias

Muestra		$ K $	$ r $	$ u $	$C$
# CGs		149	48	52	118
$\theta_G$	[ <i>arcmin</i> ]	2,7 (0,3)	4,0 (0,5)	4,0 (0,7)	3,5 (0,3)
$r_p$	[ <i>kpc h<sup>-1</sup></i> ]	63 (6)	87 (11)	64 (8)	65 (6)
$R_{\text{vir}}$	[ <i>kpc h<sup>-1</sup></i> ]	71 (8)	115 (19)	90 (17)	95 (8)
$\langle d_{ij} \rangle$	[ <i>kpc h<sup>-1</sup></i> ]	71 (7)	93 (13)	69 (10)	70 (6)
$s_{\perp}$	[ <i>kpc h<sup>-1</sup></i> ]	105 (10)	125 (19)	103 (14)	93 (10)
$s_{\parallel}$	[ <i>kpc h<sup>-1</sup></i> ]	123 (23)	128 (59)	220 (139)	108 (26)
$s_4$	[ <i>kpc h<sup>-1</sup></i> ]	173 (24)	214 (53)	229 (133)	153 (26)
$\sigma_v$	[ <i>km s<sup>-1</sup></i> ]	255 (27)	262 (60)	268 (47)	269 (24)
$H_0 t_{\text{cr}}$	[ $10^{-2}$ ]	2,7 (0,4)	2,9 (1,1)	2,5 (0,7)	2,5 (0,4)
$\mathcal{M}_{\text{vir}}$	[ $10^{12} M_{\odot} h^{-1}$ ]	5,0 (1,1)	8,7 (5,8)	7,6 (2,4)	6,1 (1,7)
$K_{\text{bri}}$		10,0 (0,1)	10,1 (0,2)	9,7 (0,3)	9,8 (0,1)
$r_{\text{bri}}$		12,8 (0,1)	12,9 (0,2)	12,5 (0,3)	12,6 (0,1)
$u_{\text{bri}}$		15,2 (0,2)	15,2 (0,2)	14,7 (0,3)	15,0 (0,1)
$\Delta K_{12}$		1,5 (0,1)	1,3 (0,2)	1,1 (0,2)	1,2 (0,1)
$\Delta r_{12}$		1,4 (0,1)	1,3 (0,2)	1,2 (0,1)	1,2 (0,1)
$\Delta u_{12}$		1,5 (0,2)	1,3 (0,3)	1,3 (0,2)	1,1 (0,1)
$\mu_K$	[ <i>mag/arcsec<sup>2</sup></i> ]	21,6 (0,2)	22,5 (0,2)	22,0 (0,3)	21,9 (0,2)
$\mu_r$	[ <i>mag/arcsec<sup>2</sup></i> ]	24,5 (0,2)	25,4 (0,2)	24,8 (0,2)	24,7 (0,2)
$\mu_u$	[ <i>mag/arcsec<sup>2</sup></i> ]	26,9 (0,2)	27,6 (0,2)	27,0 (0,2)	27,2 (0,2)
$L_K$	[ $10^9 L_{\odot} h^{-2}$ ]	231 (23)	197 (68)	196 (32)	224 (24)
$L_r$	[ $10^9 L_{\odot} h^{-2}$ ]	66 (6)	63 (18)	58 (9)	64 (7)
$L_u$	[ $10^9 L_{\odot} h^{-2}$ ]	41 (5)	40 (15)	41 (7)	41 (4)
$\mathcal{M}_v/L_K$	[ $M_{\odot}/L_{\odot} h$ ]	22 (4)	35 (17)	34 (10)	27 (5)
$\mathcal{M}_v/L_r$	[ $M_{\odot}/L_{\odot} h$ ]	73 (16)	126 (55)	117 (32)	98 (20)
$\mathcal{M}_v/L_u$	[ $M_{\odot}/L_{\odot} h$ ]	119 (26)	183 (93)	175 (58)	143 (34)
$T1_K$		0,53 (0,04)	0,53 (0,07)	0,67 (0,09)	0,55 (0,04)
$T2_K$		0,60 (0,04)	0,60 (0,07)	0,72 (0,08)	0,58 (0,04)
$T1_r$		0,50 (0,03)	0,56 (0,07)	0,62 (0,07)	0,57 (0,05)
$T2_r$		0,55 (0,03)	0,60 (0,06)	0,61 (0,07)	0,62 (0,05)
$T1_u$		0,56 (0,04)	0,71 (0,06)	0,72 (0,08)	0,58 (0,05)
$T2_u$		0,62 (0,04)	0,73 (0,08)	0,59 (0,07)	0,66 (0,05)
% Reals		53,7	45,8	40,4	59,3

**Tabla 3.4:** Propiedades de los CGs para muestras de grupos compactos puros y comunes. En cada celda, el formato  $xx$  ( $ss$ ) contiene la mediana ( $xx$ ) y el error ( $ss$ ) correspondiente aproximadamente al 95% del intervalo de confianza,  $CI = xx \pm ss$ , excepto para los valores  $T_1$  y  $T_2$  donde la cantidad entre paréntesis son las barras de error computadas usando la técnica de remuestreo bootstrap.

	$ K  -  r $	$ K  -  u $	$ u  -  r $	$ K  - C$	$ r  - C$	$ u  - C$
$\theta_G$	$2 \times 10^{-4}$	0,02	0,43	$3 \times 10^{-3}$	0,23	0,21
$r_p$	$2 \times 10^{-3}$	0,58	0,01	0,10	$7 \times 10^{-3}$	0,95
$R_{\text{vir}}$	$2 \times 10^{-4}$	0,12	0,09	$3 \times 10^{-5}$	0,20	0,14
$\langle d_{ij} \rangle$	$1 \times 10^{-3}$	0,35	$4 \times 10^{-3}$	0,17	$7 \times 10^{-4}$	1,00
$s_{\perp}$	0,04	0,62	0,20	0,60	$4 \times 10^{-3}$	0,36
$s_{\parallel}$	0,82	$4 \times 10^{-3}$	0,13	0,88	0,44	$5 \times 10^{-3}$
$s_4$	0,28	0,02	0,13	0,66	0,16	0,01
$\sigma_v$	0,77	0,73	0,93	0,38	0,28	0,30
$H_0 t_{\text{cr}}$	0,09	0,83	0,35	0,15	0,34	0,97
$\mathcal{M}_{\text{vir}}$	0,02	0,27	0,26	0,20	0,12	0,84
$K_{\text{bri}}$	0,15	0,13	0,02	0,20	0,03	0,39
$r_{\text{bri}}$	0,22	0,08	0,03	0,06	0,04	0,39
$u_{\text{bri}}$	0,84	0,03	0,04	0,28	0,59	0,08
$\Delta K_{12}$	0,12	0,07	0,20	$3 \times 10^{-3}$	0,49	0,97
$\Delta r_{12}$	0,18	0,01	0,66	$1 \times 10^{-4}$	0,22	0,42
$\Delta u_{12}$	0,18	0,01	0,66	$1 \times 10^{-4}$	0,22	0,42
$\mu_K$	$2 \times 10^{-5}$	0,03	$7 \times 10^{-3}$	0,07	$2 \times 10^{-4}$	0,34
$\mu_r$	$6 \times 10^{-5}$	0,06	0,01	0,15	$2 \times 10^{-4}$	0,34
$\mu_u$	$5 \times 10^{-4}$	0,22	$9 \times 10^{-3}$	0,10	$2 \times 10^{-3}$	0,79
$L_K$	0,39	0,39	0,83	0,69	0,11	0,42
$L_r$	0,53	0,56	0,68	0,73	0,26	0,31
$L_u$	0,30	0,80	0,46	0,66	0,08	0,60
$\mathcal{M}_v/L_K$	0,03	0,07	0,67	0,29	0,06	0,44
$\mathcal{M}_v/L_r$	0,04	0,11	0,67	0,29	0,21	0,59
$\mathcal{M}_v/L_u$	0,06	0,20	0,64	0,20	0,26	0,96

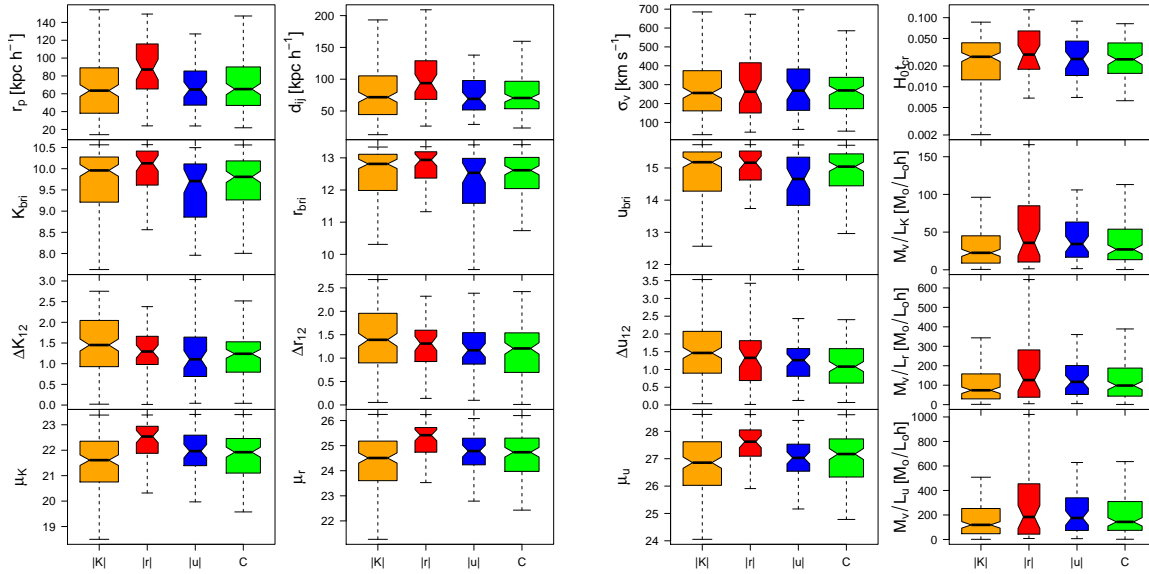
**Tabla 3.5:** “Valores p” del test Kolmogorov-Smirnov para dos muestras. Cada columna combina diferentes pares de muestras puras y comunes. Resaltamos en color gris las celdas donde los “valores p” son menores que 0.05.

encontradas entre las muestras restringidas permanecen, pero también aparecen otras diferencias cuando comparamos las muestras puras entre sí, las cuales no aparecían cuando las comparábamos con las muestras restringidas. Hemos encontrado que las masas viriales  $|K|$  son significativamente más chicas que las masas viriales  $|r|$ , y ésto se puede observar analizando la relación masa luminosidad. Además, las separaciones comoviles interpartículas 3D de los  $|u|$ -CGs son estadísticamente más grandes que para los  $|K|$ -CGs. También, las galaxias más brillantes (en cualquier banda) de los  $|r|$ -CGs son típicamente más brillantes que las galaxias más brillantes en los  $|u|$ -CGs.

Finalmente, cuando comparamos los CGs puros con los CGs comunes hemos encontrado algunas diferencias entre algunas muestras. Los tamaños proyectado de los C-CGs son bastante similares a los de la muestra  $|u|$ , sin embargo, las separaciones interpartículas 3D de los C-CGs son significativamente más pequeñas que en  $|u|$ . Las diferencias en magnitud en las tres bandas de la muestra  $|K|$  son típicamente más grandes que las diferencias en magnitud para los grupos comunes, es decir, los CGs comunes están conformados por galaxias vecinas más similares que los puros  $K$ . Los  $|r|$ -CGs son estadísticamente menos “compactos” ( $\mu$  más débiles) que los CGs comunes. Además, las galaxias más brillantes en C-CG son más brillantes que su contra parte en  $|r|$ -CGs. Curiosamente, los CG comunes presentan el mayor porcentaje de CG *Reals*. Este es un resultado novedoso que puede ayudar a seleccionar muestras observacionales con el porcentaje más bajo de alineaciones aleatorias, aunque es necesario tener datos en múltiples bandas fotométricas. Algunas de las diferencias que encontramos entre las muestras de CGs “puros” no se observan en las muestras restringidas dada la existencia de los CGs comunes que, o en algunos casos, representan un alto porcentaje de las muestras restringidas o que sus propiedades se colocan entre los valores de las muestras puras extremas, borroneando las diferencias en la muestra restringida.

### 3.3. Resumen

En resumen, hemos realizado un estudio comparativo entre CGs identificados en diferentes bandas fotométricas con el objetivo de detectar el impacto de la banda utilizada en la selección de los CGs. Utilizamos un único catálogo *lightcone* con información en las bandas  $K$ ,  $r$  y  $u$ , y usamos la misma función selección y el mismo algoritmo. Si bien el adoptar un SAM particular puede introducir dependencia de los resultados con los parámetros del propio SAM, más adelante analizaremos el efecto de usar uno u otro SAM. En este caso, el catálogo *lightcone* utilizado es uno de los pocos catálogos sintéticos libremente disponibles que proveen magnitudes aparentes en 9 bandas fotométricas, lo que genera una muestra ideal para realizar estudios comparativos entre las propiedades de los grupos compactos en diferentes bandas. Además, este SAM particular ha demostrado ser bastante eficiente al reproducir las abundancias observadas de



**Figura 3.4:** Diagramas de caja de algunas de las propiedades de los grupos compactos puros y comunes. Las muescas indican el intervalo de confianza del 95% para las medianas, mientras que el ancho de la caja es proporcional a la raíz cuadrada del número de CGs en cada muestra. La lista completa de valores para las medianas y sus intervalos de confianza para todas las propiedades se muestran en la tabla 3.4.

*galaxias a bajo redshift en un amplio rango de masas estelares y luminosidades, y también el clustering a gran escala como función de la masa estelar y los colores de las galaxias (Guo et al., 2011).*

*Identificamos CGs en 3 bandas fotométricas:  $K$ -CGs,  $r$ -CGs y  $u$ -CGs y encontramos diferencias entre sus propiedades. Analizando más en detalle encontramos que hay grupos que sólo se identificaron en una única banda (“puros”) y también grupos identificados en todas la bandas al mismo tiempo (“comunes”). Esto sugiere que las diferencias encontradas entre las muestras totales puede deberse a que estas muestras contienen grupos que se identificaron también en las otras bandas, mientras que los CGs puros representan las propiedades intrínsecas de cada banda.*

*Por lo tanto, si no se tiene especial cuidado en la selección e identificación de las muestras, el análisis de las propiedades de los CGs puede ser malinterpretado. Este trabajo fue publicado en una revista internacional con referato (Taverna et al. (2016)).*

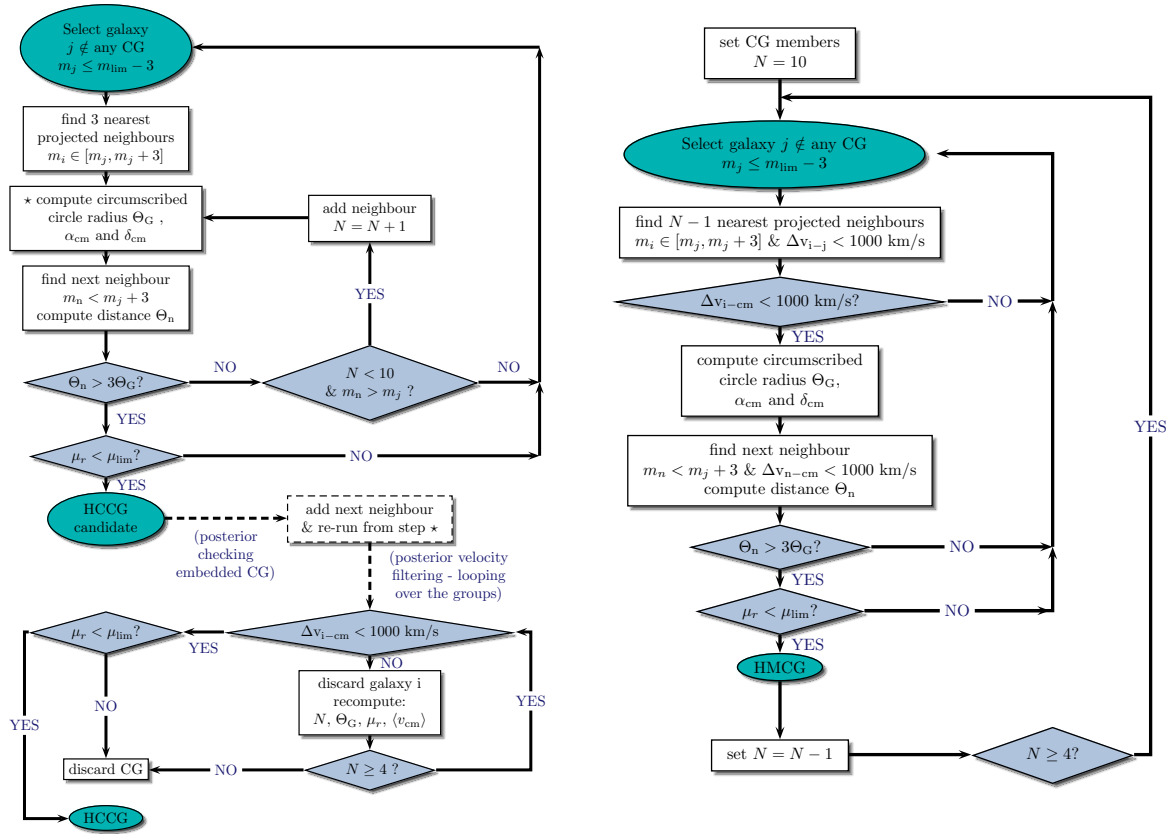
## Capítulo 4

# Optimización del algoritmo de identificación de CGs con el criterio de Hickson

*El criterio de identificación de CGs por Hickson (1982), como vimos en los capítulos anteriores, cuenta con un conjunto de propiedades globales para caracterizar a dichos grupos: compacticidad, aislamiento, y población. La gran mayoría de los trabajos que utilizan criterios tipo Hickson comparten la misma filosofía para la implementación del algoritmo: detectan primero los candidatos en proyección en el cielo, y si tienen información espectroscópica, eliminan intrusos a lo largo de la línea de la visual cuando los redshift no son concordantes. Esta forma de buscar CGs puede ser poco óptima porque puede descartar a los grupos que parecen no estar aislados en proyección, aunque sus redshift sean todos concordantes. También al seleccionar regiones del cielo como candidatos a CGs, puede descartar grupos que estén por delante o por detrás de la región seleccionada. El tipo de restricciones que surgen del propio procedimiento podría sesgar fácilmente estudios posteriores relacionados con el entorno de los CGs.*

*Utilizando el criterio tipo Hickson descrito en la sección 3.1.2 pueden aplicarse diferentes algoritmos para producir una muestra de CGs. El orden en el que se aplica el criterio puede modificar la identificación y afectar la completitud de la muestra resultante.*

*En este capítulo presentaremos una optimización del algoritmo de identificación de CGs basado en el criterio definido por Hickson que nos permitirá mejorar considerablemente la completitud de la muestra resultante. Además, aplicaremos dicho algoritmo al catálogo observacional del SDSS obteniendo un nuevo catálogo de CGs y lo compararemos con su antecesor obtenido a partir del mismo catálogo de galaxias.*



**Figura 4.1:** Diagramas que muestran la implementación automática de los criterios de Hickson para identificar CGs. A la izquierda, el diagrama muestra el algoritmo clásico, mientras que a la derecha se muestra el algoritmo modificado. (Figura 1 del artículo Díaz-Giménez et al. (2018b) ).

## 4.1. Construcción del nuevo algoritmo para la identificación de CGs

Para testear los algoritmos de identificación construimos un catálogo sintético de galaxias con evolución (*lightcone*). Usamos el SAM de galaxias construido por Henriques et al. (2015) corrido en la simulación de  $N$ -cuerpos Millennium run I (Springel et al., 2005b) re-escalada a la cosmología Planck I. Para la construcción del catálogo *lightcone* seguimos el procedimiento descrito por Zandivarez et al. (2014), el cual se encuentra detallado en el capítulo 2. El catálogo *lightcone* tiene una magnitud aparente límite en el sistema del observador que imita al límite del SDSS de  $r \leq 17,77$ . La muestra final contiene 3.139.409 galaxias dentro de un ángulo sólido de  $4 \pi \text{ sr}$ .



### 4.1.1. El algoritmo clásico

Siguiendo la idea original de Hickson (1982), el algoritmo primero identifica CGs en proyección en el plano del cielo y después aplica el filtro en velocidad, en una segunda etapa. A este algoritmo nos referiremos como *HICKSON CLÁSICO (HC)*. En el panel izquierdo de la figura 4.1 se muestra el diagrama de flujo de la implementación del algoritmo, el cual comienza buscando los CGs más pequeños. Comienza con sólo 4 galaxias y chequea que cumplan con todos los criterios. Si el criterio de aislamiento no se cumple, eventualmente se incluye una nueva galaxia miembro o se descarta el grupo. Posteriormente, se examina si los CGs se encuentran embebidos en un grupo compacto más grande (CG-en-CG) y se mantiene el más grande. Finalmente, para esos grupos identificados en proyección, se aplica el filtro en velocidad utilizando un procedimiento iterativo que elimina los intrusos más lejanos y chequea compacticidad, aislamiento y población cada vez. Una vez que tenemos la muestra de CGs, como hemos trabajado con un catálogo sintético, tenemos que considerar el hecho de que las galaxias son objetos puntuales, y no objetos extendidos como en un catálogo observacional. De la misma manera que lo hicimos en el capítulo anterior, utilizamos el radio angular a mitad de luz en la banda  $r$  (calculado como función de la masa estelar) y consideramos que dos galaxias son indivisibles si la separación angular entre ellas es más pequeña que la suma de sus radios angulares. Una consecuencia directa es la pérdida de miembros en los CGs seleccionados, por lo que los grupos que quedan con menos de 4 miembros son descartados. Este procedimiento se realiza cuando los grupos proyectados son identificados y nuevamente cuando se le aplica el filtro en velocidad.

Aplicando del algoritmo clásico a la muestra se identificaron 680 grupos compactos Hickson Clásico ( $mHCCGs$ ) (ver tabla 4.1).

### 4.1.2. El algoritmo modificado

Nuestra implementación del algoritmo clásico fue inspirada en la idea original seguida por Hickson en un momento donde los relevamientos con redshift no estaban disponibles. Diferentes autores construyeron nuevas muestras siguiendo la misma metodología automática aplicada en diferentes catálogos de galaxias (McConnachie et al., 2009; Díaz-Giménez et al., 2012).

Sin embargo, la disponibilidad de grandes relevamientos con información espectroscópica puede permitir que el algoritmo de identificación de CGs se simplifique y mejore. Dado que los criterios de búsqueda de CGs incluyen el conteo de galaxias dentro de un área proyectada en el cielo, no tener en cuenta la posición de las galaxias a lo largo de la línea de visual, lleva a perder grupos al incluir intrusos, ya sea como miembros o como violaciones del aislamiento.

En nuestro algoritmo modificado (de aquí en adelante lo llamaremos *HICKSON MODIFICADO, HM*), hemos considerado los redshifts de las galaxias desde el

Muestras	Algoritmo	
	Clásico	Modificado
# de CGs sintéticos	680	1287
% CGs <i>Reals</i>	38	35
# de CGs SDSS	218	462

**Tabla 4.1:** Número de CGs identificados usando el algoritmo clásico y el modificado.

principio de la implementación. Las galaxias consideradas vecinas (para membresía y aislamiento) se seleccionan en un cilindro en el espacio de redshifts alrededor del punto donde los criterios están siendo evaluados. Por lo tanto, el filtro en velocidad se aplica al mismo tiempo que las otras restricciones. En la implementación del nuevo algoritmo, también se introduce otro cambio para obtener automáticamente el grupo compacto más grande posible en una sola corrida. En la implementación clásica, esto se verificaba en un paso posterior. Por lo tanto, comenzamos a buscar los CG más grandes, y luego seleccionamos el CG más pequeño que no se encuentra en ningún CG más grande que se haya encontrado anteriormente. Esto se logra al identificar los CGs con la mayor cantidad de miembros al principio ( $N = 10$ ) y luego buscar grupos con menos miembros. En el panel derecho de la figura 4.1, se esquematiza el algoritmo Hickson Modificado.

Una vez que el algoritmo es aplicado al catálogo lightcone, se procede a la corrección de la población de galaxias miembros por trabajar con catálogos sintéticos con objetos puntuales. De esta manera se descartan aquellos grupos cuya membresía, luego de chequear la superposición angular de galaxias, no cumple con el número mínimo de miembros.

La muestra identificada usando el algoritmo modificado contiene 1287 *mHMCs* (ver tabla 4.1).

### 4.1.3. Comparación entre el algoritmo clásico y modificado

A continuación, hemos comparado las dos muestras de CGs identificados usando los diferentes algoritmos. En primer lugar, de la tabla 4.1, puede verse que el algoritmo modificado produce una muestra que es 89% más grande que la muestra clásica. Hemos realizado un análisis cruzado entre las dos muestras para ver si existen grupos en común en ambas muestras. Para ello, hemos considerado que un grupo de una muestra ha sido recuperado por la otra muestra si la distancia angular entre sus centros es menor que la suma de los radios angulares de los grupos. Encontramos que el 95% de la muestra clásica (647 *mHCCs*) está incluida en la muestra modificada. Esto significa que el nuevo algoritmo puede recuperar la mayor parte de la muestra clásica, pero también

identifica casi dos veces el número de grupos, mejorando la estadística significativamente.

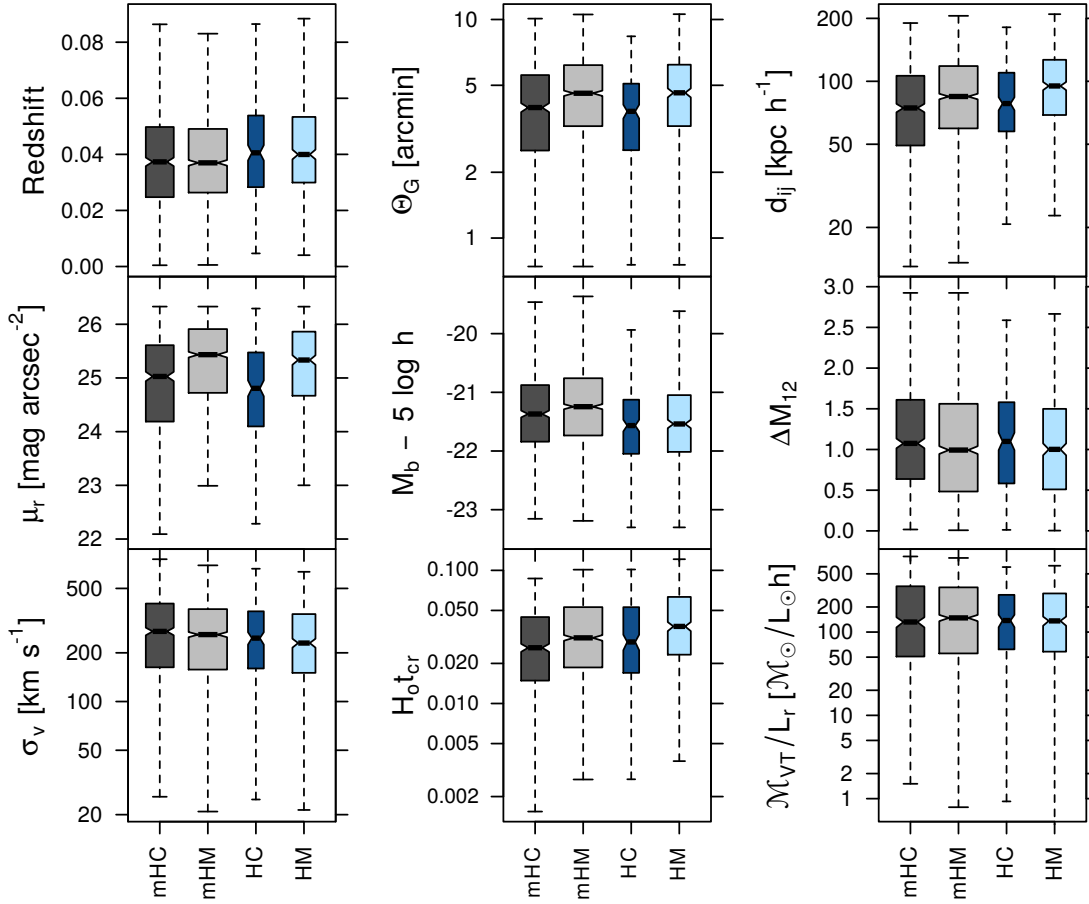
Luego, analizamos las razones de por qué el algoritmo clásico pierde casi la mitad de los  $mHMCGs$ . Encontramos que el 90% de los grupos perdidos son eliminados por el algoritmo clásico porque existen galaxias con velocidades **discordantes** contaminando el anillo de aislamiento en proyección. En el 10% restante, la membresía se vio alterada, algunos de los miembros verdaderos del grupo se identificaron en proyección, pero otros se perdieron, y en cambio se incluyeron nuevas galaxias con redshifts discordantes como miembros del grupo. Estos grupos no pasaron el filtrado de velocidad posterior.

Dado que existe una fracción no despreciable de CGs sintéticos que podrían considerarse como proyecciones aleatorias a lo largo de la línea de visual (Mamon, 1986; McConnachie et al., 2008; Díaz-Giménez & Mamon, 2010), es interesante examinar si los diferentes algoritmos podrían llevar a diferentes porcentajes de contaminación. Al igual que en el capítulo anterior, siguiendo la clasificación en espacio real 3D propuesta por Díaz-Giménez & Mamon (2010), dividimos la muestra en CGs *Reals* y alineamientos aleatorias a lo largo de la línea de la visual.

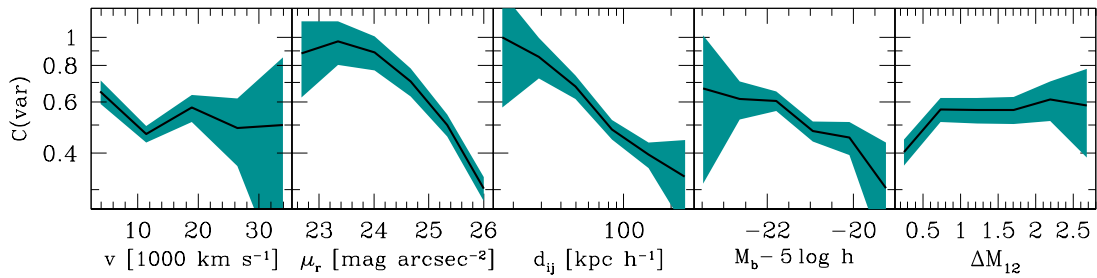
Encontramos que el  $(38 \pm 3)\%$  de los  $mHCCGs$  (256) son *Reals* (es decir, físicamente densos), mientras que para los  $mHMCGs$  ese porcentaje llega al  $(35 \pm 2)\%$  (Tabla 4.1). Por lo tanto, el porcentaje de grupos físicamente densos no se ve afectado por el algoritmo, es decir, es inmutable frente a las diferentes aplicaciones de los algoritmos. Este porcentaje de CGs *Reals* en este trabajo es significativamente menor que el citado en trabajos anteriores. En el capítulo 5 realizaremos un análisis más detallado sobre este tema.

En la figura. 4.2, mostramos las distribuciones de varias de las propiedades de los CGs obtenidos con ambos algoritmos usando diagramas de caja (ver apéndice B para el cálculo de las propiedades). Las primeras dos distribuciones en cada gráfico corresponden a las muestras simuladas  $mHCCGs$  y  $mHMCGs$ . En esta figura, las muescas (relacionadas con el intervalo de confianza) no se superponen para la mediana de la separación proyectada interpartícula ( $d_{ij}$ ), el tiempo de cruce adimensional ( $H_0 t_{cr}$ ), el diámetro angular del círculo mínimo ( $\Theta_G$ ), la magnitud absoluta de la galaxia más brillante ( $M_{bri}$ ), y el brillo superficial del grupo ( $\mu_r$ ), lo que indica que las medianas de esas distribuciones son significativamente diferentes (McGill et al., 1978; Krzywinski & Altman, 2014). La muestra modificada incluye CGs con separación proyectada entre galaxias más grande, tiempos de cruce más largos y tamaños angulares más grandes. La galaxia más brillante de los CGs son más débiles para los HM y el brillo superficial también es más débil (menos compactos).

Es posible calcular la completitud de la muestra clásica en función de varias variables: velocidad radial del grupo, brillo superficial del grupo, mediana de la separación proyectada entre galaxias, magnitud absoluta de la galaxia más brillante, y gap de magnitud absoluta entre las dos galaxias más brillantes. Nues-



**Figura 4.2:** Diagrama de caja de las propiedades de CGs para los grupos identificados usando el algoritmo clásico y el algoritmo modificado. Las muescas en los cuadros indican los intervalos de confianza  $\sim 95\%$  para las medianas, mientras que los anchos son proporcionales a las raíces cuadradas del número de objetos en cada muestra (Krzywinski & Altman, 2014). Las distribuciones del lado izquierdo del gráfico son las muestras obtenidas utilizando los algoritmos clásico y modificado en el catálogo sintético (mHC y mHM), mientras que los del lado derecho muestran los resultados de las muestras obtenidas aplicando ambos algoritmos en el catálogo de SDSS (HC y HM).



**Figura 4.3:** Cociente entre número de CGs identificados con el algoritmo clásico con respecto al modificado para una propiedad particular. Las áreas sombreadas corresponden a los errores calculados utilizando la fórmula de propagación de errores para el cociente (cada número en la relación tiene un error de Poisson).

tra intención es analizar si existe algún tipo de sesgo en el algoritmo o si simplemente los grupos que no pueden ser identificados por el algoritmo clásico están homogéneamente distribuidos en cuando sus propiedades. Hemos definido la completitud de la muestra clásica respecto de la muestra modificada, como el cociente entre el número de compactos clásicos y el número de compactos modificados por bin de cada variable ( $var$ ):

$$C(var) = N_{mHCCG}(var)/N_{mHMCg}(var)$$

Los resultados se muestran en la figura 4.3. Hay una pequeña dependencia de la completitud de la muestra clásica con la velocidad radial de los grupos, pero el algoritmo clásico falla en encontrar CGs en todas las distancias por igual. Por otro lado, el algoritmo clásico puede identificar sin problemas los CGs con brillos superficial más brillantes (valores bajos de  $\mu_r$ ) pero va perdiendo CGs gradualmente hacia el límite de compacticidad del criterio. Lo mismo sucede analizando la mediana de la separación proyectada entre galaxias: los CGs con mayor separación no son detectados por el algoritmo clásico. Además, este algoritmo también pierde más CGs a medida que se debilita la galaxia más brillante del grupo, aunque tampoco está completo en el extremo brillante. Finalmente, el algoritmo clásico pierde grupos en todo el rango de diferencias en magnitud entre las dos galaxias más brillantes, y la incompletitud es más notoria para los CGs con las dos galaxias más brillantes del grupo similares en brillo.

## 4.2. Aplicación del algoritmo modificado al catálogo observacional SDSS

Hemos implementado el algoritmo modificado a las galaxias del Sloan Digital Sky Survey (SDSS) para construir un nuevo catálogo de CGs que cumple estrictamente con los criterios de selección.

tamente con todos los criterios de Hickson.

### 4.2.1. El SDSS

Utilizamos el catálogo de galaxias construido por Tempel et al. (2017)<sup>1</sup> a partir del relevamiento Sloan Digital Sky Survey (SDSS) Data Release 12 (Eisenstein et al., 2011; Alam et al., 2015). Fue construido usando galaxias con información espectroscópica extraídas del área principal continua del relevamiento (the Legacy Survey) complementada con redshifts del Two-degree Field Galaxy Redshift Survey (Colless et al., 2001, 2003), el Two Micron All Sky Survey Redshift Survey (Jarrett et al., 2003; Skrutskie et al., 2006b; Huchra et al., 2012), y el Third Reference Catalogue of Bright Galaxies (de Vaucouleurs et al., 1991; Corwin et al., 1994). La muestra extendida de galaxias contiene 584 449 galaxias con magnitudes petrosianas en el sistema de referencia del observador  $r \leq 17,77$  y redshifts corregidos al sistema en reposo del Fondo Cósmico de Microondas (CMB)  $z_{\text{CMB}} \leq 0,2$  dentro de un ángulo sólido de 6828 grados cuadrados.

Hemos adoptado las magnitudes *model* como la magnitudes aparentes principales para este relevamiento. Seleccionamos galaxias con magnitudes *model* en el sistema en reposo  $r$  menores que 17,77 y colores en el sistema de referencia con respecto al observador  $g - r \leq 3$  para eliminar estrellas. Nuestro catálogo final esta formado por 557 517 galaxias. Para calcular las magnitudes absolutas en el sistema en reposo utilizamos las correcciones  $K$  para las magnitudes  $r$  y  $g$  que fueron calculadas siguiendo las prescripciones de Chilingarian & Zolotukhin (2012). Los parámetros cosmológicos usados fueron obtenidos por la Planck Collaboration et al. (2015).

### 4.2.2. Identificación de CGs en el SDSS

Hemos identificado CGs en la muestra de galaxias SDSS utilizando el criterio de Hickson usando el algoritmo *modificado*. La muestra seleccionada tiene 476 CGs.

Conociendo las diferentes limitaciones observacionales que tiene el catálogo SDSS y con el objetivo de obtener una muestra de grupos compactos confiable, realizamos una inspección visual de todos los CGs usando datos del relevamiento SDSS DR14 a través de la herramienta *Image List Tool*<sup>2</sup>. Esta inspección de los miembros reveló que 21 miembros de nuestros CGs fueron erróneamente clasificados como galaxias en la muestra extendida de Tempel et al. cuando en realidad eran partes de galaxias (PofG, por la sigla en inglés). El número de identificación de estos objetos en el SDSS (ObjIDs) están citados en la tabla C.1 en el apéndice C.

<sup>1</sup><http://cosmodb.to.ee>

<sup>2</sup><http://skyserver.sdss.org/dr14/en/tools/chart/listinfo.aspx>

También tuvimos en cuenta las incompletitudes que sufre el relevamiento espectroscópico SDSS relacionadas con la saturación de galaxias brillantes y/o colisión de fibras debido a pares de galaxias muy cercanos en proyección. Estas cuestiones pueden tener alto impacto en la identificación de CGs, principalmente en la aplicación del criterio de población y aislamiento. Por esta razón, hemos usado información fotométrica de las galaxias extraídas del SDSS DR14 (Abolfathi et al., 2018) para buscar potenciales galaxias que están en la vecindad de cada grupo identificado y que no fueron detectadas por el relevamiento espectroscópico ni tampoco incorporadas por Tempel et al. (2017). En el apéndice C, hemos descrito el pedido utilizado para recuperar galaxias sin espectroscopia a partir del SDSS DR14 Casjobs<sup>3</sup>. A partir de la muestra de galaxias en SDSS sin redshifts, hemos seleccionado aquellas que están en el plano del cielo dentro de  $3\Theta_G$  del centro de cada grupo identificado y donde sus magnitudes aparentes *model r* estén dentro de un rango de tres magnitudes desde la galaxia más brillante del CGs ( $r \leq r_b + 3$ ). Obtuvimos así una lista de 303 objetos sin información espectroscópica que pueden contaminar nuestra muestra de CGs. Para estos objetos, buscamos sus redshifts en la literatura usando la fuente de datos NASA/IPAC Extragalactic Database (NED)<sup>4</sup>. Encontramos que 63 de esas galaxias tienen sus redshifts determinados por diferentes fuentes (ver tabla C.2 en el apéndice C).

Los objetos clasificados como galaxias cuando en realidad no lo son, o galaxias no identificadas por problemas del relevamiento pueden cambiar la identificación de CGs. Es por esta razón que se corrigió el catálogo inicial de galaxias eliminando los objetos clasificados como PofGs y agregando los objetos con redshifts disponible. El algoritmo modificado fue corrido nuevamente en esta muestra corregida de 557 559 galaxias espectroscópicas. Obtuvimos una nueva muestra de 462 HMCG en el SDSS que constituye nuestra catálogo final (ver tabla 4.1). Nuevamente, se estudió el entorno de estos grupos compactos seleccionando 290 galaxias fotométricas sin redshift que están en el plano del cielo alrededor de los HMCGs de la misma manera como describimos anteriormente (proximidad angular y rango de magnitud). Examinamos estos objetos para clasificarlos como potenciales fuentes de contaminación o galaxias no contaminantes. Esta clasificación es de gran utilidad para determinar si los CGs identificados cumplen con los criterios de población y aislamiento.

De las 290 galaxias sin redshifts en el SDSS y localizadas cerca de los HMCGs, hemos encontrado que 52 galaxias tienen redshifts espectroscópico y ya habían sido incorporadas a la muestra extendida de Tempel et al.; 53 galaxias tienen redshifts fotométricos que son claramente discrepantes con la mediana del redshift espectroscópico del HMCG (Beck et al. 2016, galaxias con  $|z_{\text{phot}} - z_{\text{cm}}|/(1 + z_{\text{cm}}) > 0,06$  pueden ser descartadas de manera confiable por ser intrusos).

De las 185 galaxias fotométricas restantes consideradas como fuentes potenciales de contaminación de los HMCGs, hemos comparado las propiedades fotométricas

<sup>3</sup><http://skyserver.sdss.org/casjobs/>

<sup>4</sup><https://ned.ipac.caltech.edu/>

cas de esas galaxias con aquellas galaxias espectroscópicas que se encuentran alrededor de los mismos HMCGs que las galaxias intrusas podrían estar contaminando, y definimos un criterio que nos ha permitido seleccionar cuáles de esas galaxias no son potenciales contaminantes.

Para cada grupo con una potencial galaxia contaminante, hemos seleccionado galaxias espectroscópicas a partir del SDSS DR14 que se encuentran:

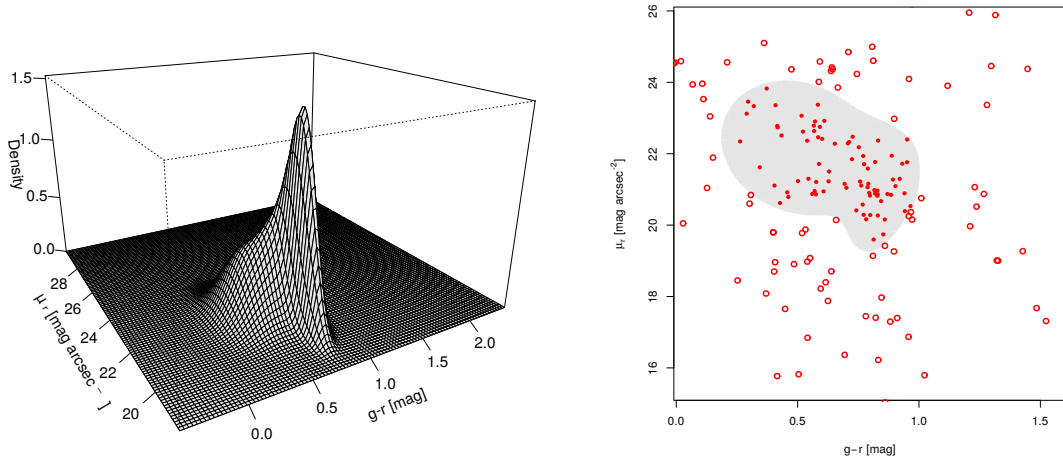
1. dentro de un círculo de 5 grados de radio desde el centro del HMCG,
2. dentro de un rango de 3 magnitudes desde la galaxia más brillante del HMCG,
3. y dentro de  $1000 \text{ kms}^{-1}$  con respecto a la mediana de la velocidad radial de los HMCG.

Para cada galaxia espectroscópica y cada potencial contaminante fotométrica, le medimos el brillo superficial ( $\mu_r$ ) y el color en el sistema de referencia del observador ( $g - r$ ). El brillo superficial de cada galaxia fue calculado usando el radio petrosiano proporcionado por el SDSS pipeline. Después de analizar la relación brillo superficial – color, hemos usado el paquete *Mclust* (Scrucca et al., 2017) del software R para generar una estimación de la densidad basada en el modelo de mezcla de gaussianas (finite Gaussian mixture modelling). De esta aplicación a nuestros datos, en el plano brillo superficial – color, observamos que de acuerdo con el criterio de información bayesiana, se puede obtener una representación de densidad adecuada utilizando una mezcla de cuatro funciones gaussianas para un modelo con distribución elipsoidal que permita variables tal como el volumen, forma y orientación. Los mejores parámetros ajustados para el modelado de la mezcla gaussiana obtenida se citan en la tabla 4.2, mientras que la gráfica de perspectiva de la densidad superficial resultante se muestra en el panel izquierdo de la figura 4.4.

Usando esta densidad superficial, hemos definido el contorno de isodensidad que encierra el 95% de las galaxias espectroscópicas en el plano brillo superficial – color. Esta región esta representada como el área gris en el panel derecho de la figura 4.4. También mostramos en el gráfico las posibles fuentes de contaminación (puntos rojos). Las galaxias fotométricas que permanecen fuera de la región gris son consideradas como galaxias no contaminantes (puntos rojos vacíos). Este criterio permite descartar 104 galaxias de las 185 originales potenciales contaminantes.

Finalmente, después de la inspección visual de los 81 objetos remanentes, hemos confirmado que 13 eran objetos mal clasificados como galaxias por el SDSS pipeline. Por lo tanto, sólo 68 galaxias son tenidas en cuenta como potenciales fuentes de contaminación. Los números de identificación de estas 68 galaxias son mostrados en la tabla C.3 en el apéndice C.





**Figura 4.4:** Panel izquierdo: Gráfico de perspectiva del estimador de densidad bi variable en el plano brillo superficial vs color para galaxias espectroscópicas en el SDSS DR14 que están en la misma región (mismo redshift y rango de magnitud) que aquellos HMCs que contienen galaxias fotométricas potencialmente contaminantes. Panel derecho: El área gris encierra el 95 % de las galaxias espectroscópicas. Los puntos rojos rellenos son galaxias fotométricas en el mismo rango de magnitud que sus grupos anfitriones y son consideradas como fuentes potenciales de contaminación (dentro de la región gris), mientras que los puntos rojos vacíos son galaxias clasificadas como no contaminantes (fuera de la región gris).

Gaussiana	MP	Vector media		Matriz de Covariancia		
		$\bar{x}$	$\bar{y}$	$\sigma_{xx}$	$\sigma_{yy}$	$\sigma_{xy} = \sigma_{yx}$
1	0.220	0.745	21.776	0.0132	0.2884	0.0046
2	0.451	0.500	22.179	0.0239	0.8453	-0.0209
3	0.262	0.827	21.024	0.0044	0.2565	0.0011
4	0.067	0.812	20.216	0.0024	0.2212	0.0051

**Tabla 4.2:** *Mclust* mejor ajuste de paámetros obtenidos para el estimador bi variable obtenido a partir de cuatro funciones gaussianas en el plano brillo superficial – color. Los valores MP mostrados en la segunda columna son las proporciones de mezcla para cada componente del modelo de mezcla. Para los vectores medios y las matrices de covarianza, la variable  $x$  corresponde al color de la galaxia  $g - r$ , mientras que la variable  $y$  es el brillo superficial  $\mu_r$ .

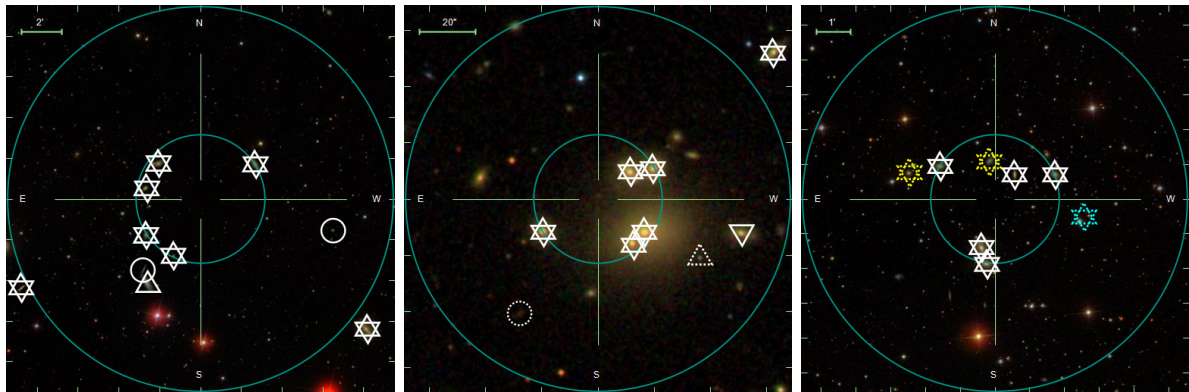
Por lo tanto, alrededor de los 462 CGs identificados existen 68 galaxias que son objetos potencialmente contaminantes, que de hecho están alrededor de 56 CGs (algunos grupos tienen más de una potencial contaminante). En la figura 4.5, mostramos tres ejemplos de las regiones alrededor de *HMC*Gs estudiadas. Las imágenes fueron obtenidas del SDSS DR14 Image Tool. Las galaxias miembro se muestran como triángulos apuntando hacia arriba combinadas con triángulos apuntando hacia abajo (estrellas) dentro o en el borde del círculo interior. La imagen de la izquierda muestra un *HMC*G sin ningún objeto con redshift desconocido dentro del radio del grupo o el anillo de aislación. La imagen del centro muestra un ejemplo donde dos objetos sin redshift están dentro del anillo de aislación ( $3 \Theta_G$ ), pero uno de esos objetos está fuera del rango de magnitud de los miembros del grupo (círculos abiertos), mientras que los otros dos tienen redshifts fotométricos claramente discordantes con respecto al grupo y por lo tanto son clasificadas como galaxias no contaminantes (triángulos apuntando hacia abajo); los demás objetos que se pueden ver en la imagen que no tienen símbolos son galaxias más débiles que la magnitud aparente límite de la muestra SDSS espectroscópica o son estrellas. Finalmente, la imagen de la derecha muestra un CG que está rodeado de tres objetos sin redshift conocido, y en el mismo rango de magnitudes que los miembros del grupo. Una galaxia en el anillo de aislación ha sido clasificada como no contaminante basados en el método descrito en el apéndice 4.2.2 (estrellas celestes con líneas discontinuas), y los otros dos, uno dentro del radio del grupo y otro dentro del anillo de aislación, son posibles contaminantes (estrellas amarillas). Por lo tanto, el único grupo con potenciales fuentes de contaminación es el grupo de la derecha.

Analizando los objetos fotométricos potencialmente contaminantes del SDSS DR14, 18% son espectroscópicamente confirmados con espectroscopía externa, fueron incluidos en nuestro catálogo principal, 59% son claros intrusos o objetos no clasificados, mientras que el 23% son inciertos. Por lo tanto, de la muestra de 462 *HMC*Gs, 406 de ellos pueden ser considerados limpios (sin fuentes de contaminación), mientras que 56 *HMC*Gs necesitan más información de las galaxias a su alrededor para eliminar potenciales fuentes contaminantes.

En el artículo Díaz-Giménez et al. (2018b) hemos presentado un nuevo catálogo de *HMC*Gs. Los grupos potencialmente contaminados no han sido eliminados de la muestra, pero presentan al final de la tabla una bandera que indica si necesita más información espectroscópica de fuentes potenciales de contaminación (ver tabla C.3 en el apéndice C).

### **Algoritmo clásico vs. modificado en el SDSS**

Con el fin de comparar con los resultados encontrados en los catálogos simulados, aplicamos el algoritmo clásico a la muestra extendida y corregida de galaxias de SDSS, y obtuvimos una muestra de 218 *HCC*Gs (ver tabla 4.1). De manera similar a los resultados encontrados en el catálogo simulado, el algoritmo modificado pudo recuperar 95% de estos grupos, mientras que casi duplica



**Figura 4.5:** Imágenes de tres HMCs en SDSS. El círculo interior denota el círculo mínimo que encierra todos los miembros del grupo ( $\Theta_G$ ), mientras que el círculo exterior representa el radio de aislamiento ( $3\Theta_G$ ). Las galaxias con información espectroscópica y magnitud aparente  $r \leq 17,77$  se muestran en símbolos con líneas continuas, mientras que galaxias  $r > 17,77$  se muestran con líneas discontinuas. El triángulo apuntando hacia arriba son galaxias en el mismo rango de magnitud que los miembros de los CGs; los triángulos apuntando hacia abajo están en el mismo rango de redshift que el CG; los círculos son galaxias fuera del rango de magnitud del CG. Los símbolos celestes son galaxias fotométricas consideradas no contaminantes mientras que símbolos amarillos son objetos potencialmente contaminantes. Los objetos sin símbolos son galaxias más débiles que  $r_{\text{lim}} = 17,77$  o estrellas. **Izquierda:** El ejemplo de un HMC no tiene ninguna fuente potencial de contaminación. (Grupo  $ID = 55$ ) **Centro:** Un CG con galaxias no contaminantes dentro del disco de aislamiento cuyos redshifts fotométricos están claramente fuera del rango de redshift del CG. (Grupo  $ID = 210$ ). **Derecha:** Un CG con dos potenciales fuentes de contaminación – una dentro del radio del grupo y otra en el anillo de aislamiento – que no se pueden descartar, mientras otra galaxia fotométrica en el anillo de aislamiento ha sido considerada como no contaminante. El catálogo de grupos HMCs se encuentra públicamente disponible para su descarga en VizieR (Diaz-Gimenez et al. (2018a), <http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=J/A+A/618/A157>).

la cantidad de CGs identificados. Además, el algoritmo clásico pierde casi la mitad de HMCGs debido a que las galaxias con velocidad discordante contaminan el anillo de aislamiento. En la figura 4.2, mostramos los diagramas de caja obtenidos para las propiedades de los CGs identificados usando ambos algoritmos, donde las muestras son llamadas como Hickson clásico (HC) y Hickson modificado (HM). Estudiando esta comparación, observamos tendencias similares a las de las muestras obtenidas del catálogo sintético: los valores de las medianas de  $d_{ij}$ ,  $H_0 t_{cr}$ ,  $\Theta_G$  y  $\mu_r$  son significativamente más altos con el algoritmo modificado que con el clásico.

Además, analizando la completitud en función de las diferentes propiedades han llegado a los mismos resultados encontrados para las muestras simuladas. Por lo tanto, la completitud de la muestra de CGs observados se ha mejorado con la nueva implementación del algoritmo.

### 4.2.3. Comparación con otros catálogos de CGs seleccionados a partir del SDSS

Varias muestras de CGs han sido extraídas del SDSS construidas por diferentes autores. Particularmente, McConnachie et al. (2009) (de aquí en adelante *McC09*) identificó CGs en proyección en el SDSS Data Release 6 (DR6). Es importante mencionar que la identificación realizada en proyección por *McC09* utilizó como magnitud límite una magnitud más brillante que la magnitud límite del catálogo ( $r_{lim} = 17,77$ ). Una de sus muestras (catálogo A) es identificada con una magnitud límite de  $r \leq 18$ , mientras que la otra muestra (catálogo B) es con  $r \leq 21$ . Por otro lado, *McC09* no tiene en cuenta el límite en flujo en su criterio ( $r_b \leq r_{lim} - 3$ ) y por esta razón, el aislamiento y la población no fueron estrictamente chequeados (como lo menciona *McC09*). Además, la falta del criterio de límite en flujo puede introducir una correlación artificial entre el tamaño y el redshift del grupo. Finalmente, la identificación de los CGs fue realizada sobre una submuestra del SDSS DR6 con solo galaxias más débiles que  $r_{inf} = 14,5$ . Esta restricción implica que la muestra pierda CGs a bajo redshift.

Años posteriores a la construcción del catálogo, Sohn et al. (2015) busco información de la velocidad radial a partir de diferentes fuentes de las galaxias miembros de *McC09* CGs, y luego aplicó el filtro en velocidad (CGs *SOHN15*). En el mejor de los casos, la muestra resultante puede ser equivalente a nuestra implementación del algoritmo clásico.

Existe otro catálogo el cual fue extraído del SDSS DR12 por Sohn et al. (2016). Ellos usaron un algoritmo de percolación con dos longitudes de percolación (en proyección y en la línea de la visual) y buscaron sistemas con alto contraste de sobredensidad (pequeña separación entre galaxias). Los autores han ignorado explícitamente el criterio de aislamiento, que podría llevar a una muestra de pequeñas estructuras dentro de sistemas más grandes. Este catálogo será examinado más adelante en la siguiente sección.

	McC09		SOHN15	SOHNCG	HMCG
	A	B			
SDSS	DR6	DR6	DR12 <sup>+</sup>	DR12 <sup>+</sup>	DR12 <sup>+</sup>
criterio	PRO	PRO	P+VF	FoF	MOD
$r_{\text{lim}}$	18	21	17.77	17.77	17.77
$r_{\text{inf}}$	14.5	14.5	--	--	--
$\mu_{\text{lim}}$	26	26	26	26	26.33
Aislamiento	SI	SI	SI	NO	SI
Total	2297	74791	332	1588	462
$N \geq 4$	2297	74788	192	312	462
+ $N_z = N$	153	55	140	312	462
+ $\Delta v_{i,\text{cm}} \leq 1000$	44	13	140	312	462
+ $r_b \leq 14,77$	2	2	11	142	462
+ $\mu_r \leq 26$	2	2	11	142	383
$v [10^3 \text{ km s}^{-1}]$	--	--	19.6(2.7)	10.6(0.7)	12.2(0.6)
$r_p [h^{-1} \text{ kpc}]$	--	--	40(8)	38(2)	73(3)
$\sigma_v [\text{ km s}^{-1}]$	--	--	215(71)	323(37)	236(16)
$\Delta M_{12}$	--	--	1.4(0.4)	0.8(0.1)	1.0(0.1)

**Tabla 4.3:** Catálogos de grupos compactos a partir del SDSS: McC09=McConnachie et al. (2009), SOHN15=Sohn et al. (2015), SOHNCG=Sohn et al. (2016). SDSS: el su-  
praiñdice + significa que los datos fueron complementados con información del redshift de otras fuentes diferentes. Criterio: PRO = en proyección, P+VF= en proyección más un posterior filtro en velocidad, FoF = algoritmo friends-of-friends, MOD = algoritmo modificado (este trabajo).  $N_z$  es el número de miembros con información espectroscópica mientras que el signo + al principio de cada fila indica que se incluye la restricciones previas. En las cuatro filas de abajo, el formato xx (ss) contiene las medias de las propiedades (xx) y el corrimiento (ss) para construir el intervalo de confianza del 95%,  $CI = xx \pm ss$ , donde  $ss = 1,58 \times \text{IQR}/\sqrt{n}$ ,  $n$  es el número de objetos en la muestra y IQR es el rango intercuartil (Krzywinski & Altman, 2014).

En la tabla 4.3, mostramos las principales características de estos grupos compactos extraídos de diferentes realizaciones del SDSS<sup>5</sup>. Las primeras filas se refieren al criterio de selección de cada muestra. Las filas del medio contienen el número de CGs en cada muestra. Cada fila en esta parte de la tabla es obtenida aplicando la restricción indicada en la primer columna e incluyendo las restricciones de las filas previas. Finalmente, las últimas cuatro filas muestran la mediana de alguna propiedades de las muestras de CGs resultantes después de aplicar todas las restricciones. Eso es necesario para que las muestras puedan ser comparables. Dado el pequeño número de CGs en la muestra de CGs McC09 restringida, estos valores no pueden ser incluidos.

<sup>5</sup>Las muestras de CGs usadas en este trabajo fueron descargadas desde <http://vizier.u-strasbg.fr/viz-bin/VizieR>

### HMCGs vs SOHNCGs

Como mencionamos anteriormente, es difícil comparar muestras de CGs identificados por diferentes autores dadas las diferencias entre los catálogos, la banda fotométrica, los algoritmos, las definiciones de qué es un CG, etc. Sin embargo, hemos realizado una detallada comparación entre la nueva muestra observacional *HMCG* descrita en la sección. 4.2 y la muestra de CGs identificada por Sohn et al. (2016).

Similar a nuestro trabajo, Sohn et al. usaron un catálogo espectroscópico de galaxias del SDSS DR12 (Alam et al., 2015), pero ellos complementaron el relevamiento agregando información espectroscópica de otros relevamientos como el de Hwang et al. (2010) y NED, especialmente para galaxias brillantes ( $r < 14,5$ ). La muestra final incluye 654 066 galaxias con magnitudes  $r < 17,77$  en el rango de redshift  $0,01 < z < 0,20$ .

A diferencia de nosotros, la identificación realizada por Sohn et al. (2016) fue utilizando el algoritmo FoF similar al utilizado por Barton et al. (1996). Su algoritmo FoF esta basado en dos longitudes de percolación que restringen la distancia espacial física proyectada de las galaxias ( $\Delta D \leq 50 \text{ kpc h}^{-1}$ ) y la separación en velocidad en la línea de la visual en el sistema en reposo ( $\Delta V \leq 1000 \text{ km/s}$ ) de cada par de vecinos. Además ellos incluyen la población y el criterio de compactidad de los algoritmos tipo Hickson ( $N(\Delta r \leq 3) \geq 3$  y  $\mu_r \leq 26 \text{ mag arcsec}^{-2}$ ). Sin embargo, no aplican ningún criterio de aislamiento a sus CGs. Obtuvieron una muestra de 1588 CGs<sup>6</sup>.

Para realizar una comparación apropiada con nuestra muestra, necesitamos incluir varias restricciones. Para empezar, nosotros seleccionamos solo grupos con cuatro o mas miembros. Por lo tanto la muestra de grupos identificados por los otros autores con cuatro miembros o mas queda restringida a 312 CGs. Además, seleccionamos grupos que cumplen con el criterio en flujo ( $r_b \leq r_{\text{lim}} - 3$ ) por lo que el criterio de población ha sido bien chequeado, dejando una muestra de 142 CGs. Finalmente aplicamos la máscara angular del catálogo de Tempel et al. (2017) para tener la misma área de cobertura. La muestra final contiene 109 grupos (de aquí en adelante *SOHNCG*). Realizando una correlación angular cruzada entre esta muestra y la nuestra *HMCGs*, encontramos que 44% de los 109 *SOHNCGs* están incluidos en nuestra muestra. Las razones de porque no pudimos recuperar los 56% restantes serán discutidas más adelante.

Por otro lado, para comparar ambas muestras nosotros también tenemos que restringir nuestra muestra para imitar el brillo superficial límite que impusieron Sohn et al.. Los dos límites son diferentes dado que ellos mantuvieron el límite original del criterio de Hickson en la banda  $R$  ( $\mu_R < 26$ ), mientras que nosotros convertimos el límite a la banda  $r$  ( $\mu_r < 26,33$ ). Con  $\mu_r < 26,0$ , nuestra muestra paso de tener 462 grupos a 383 *MHCGs*. Ahora estas muestras si son comparables,

<sup>6</sup>Disponible públicamente como tablas en <http://vizier.u-strasbg.fr/viz-bin/VizieR-3?-source=J/ApJS/225/23> en VizieR

y podemos notar que nuestra muestra *HMCG* es un factor 3.5 más grande que los *SOHNCG*.

En la figura 4.6, hemos comparado las propiedades de estas dos muestras de CGs extraídos del SDSS. Utilizamos un diagrama (“beanplot”) que nos permite comparar las distribuciones de dos muestras en un mismo gráfico y permite una mejor visualización de los datos. En los beanplots se muestra la forma de la densidad promediada con una normalización adecuada (ver Kampstra 2008) para una propiedad dada, lo que nos permite hacer una comparación directa entre las muestras. Las líneas continuas horizontales representan la mediana de la propiedad dada, mientras que las líneas discontinuas representan sus intervalos de confianza.

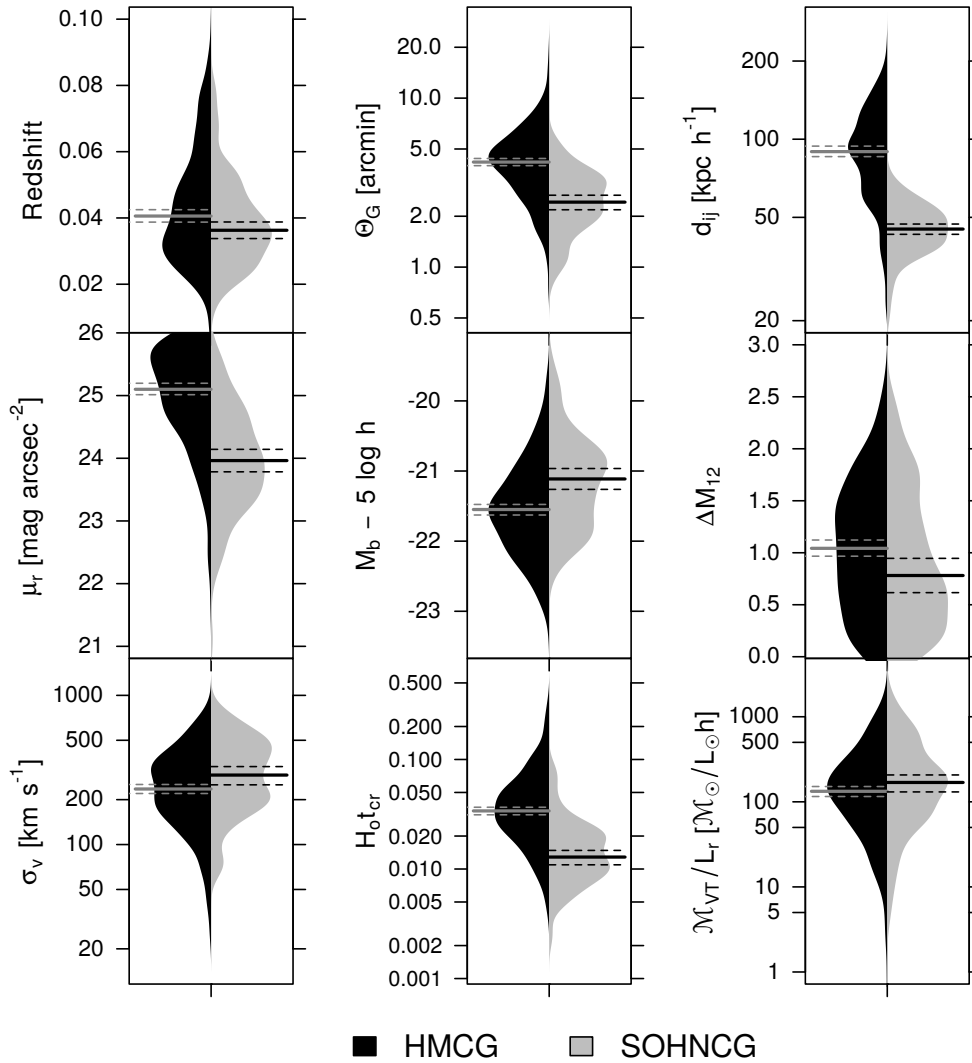
De esta comparación observamos que la principal diferencia es que los *SOHNCGs* son más pequeños, las distribuciones de  $\Theta_G$ ,  $d_{ij}$ ,  $\mu_r$ , y  $H_0 t_{cr}$  están corridas hacia valores más bajos comparados con nuestros *HMCG*. Estos resultados pueden estar relacionados con el procedimiento de identificación: las longitudes de percolación pequeñas usadas en el algoritmo FoF por Sohn et al. (2016) lleva a identificar grupos más pequeños en proyección que aquellos identificados con algoritmos tipo Hickson. Además, observamos que *HMCGs* muestran que las galaxias más brillantes del grupo y la diferencia en magnitud entre las dos galaxias más brillantes ( $\Delta M_{12}$ ) son estadísticamente más brillantes y con gap más grande que sus contraparte *SOHNCGs*. Por lo tanto, nuestra muestra está dominada por galaxias más brillantes.

Además, aunque la mediana en redshift para nuestra muestra de *HMCG* esté desplazada ligeramente hacia valores más altos (una diferencia de  $\sim 0,005$  en redshift ó  $\sim 1500 \text{ km s}^{-1}$ ) en comparación con los *SOHNCG*, esta diferencia no es estadísticamente significativa (ambas se encuentran dentro del intervalo de confianza del 95 % de la otra mediana) y las dos distribuciones de densidad son muy similares.

Sohn et al. (2016) encontró que la distribución de redshift de su muestra de CGs fue aproximadamente  $9000 \text{ km s}^{-1}$  más baja (más cercana) que la de la muestra identificada por McConnachie et al. (2009) en el SDSS DR6 (con filtrado de velocidad realizado posteriormente por Sohn et al. 2015). Sohn et al. (2016) atribuyó esta diferencia al criterio de aislamiento incorporado por McConnachie et al. (2009). Sin embargo, nuestros CGs clásico y modificado tienen una mediana en redshift similar a los valores de Sohn et al. (2016).

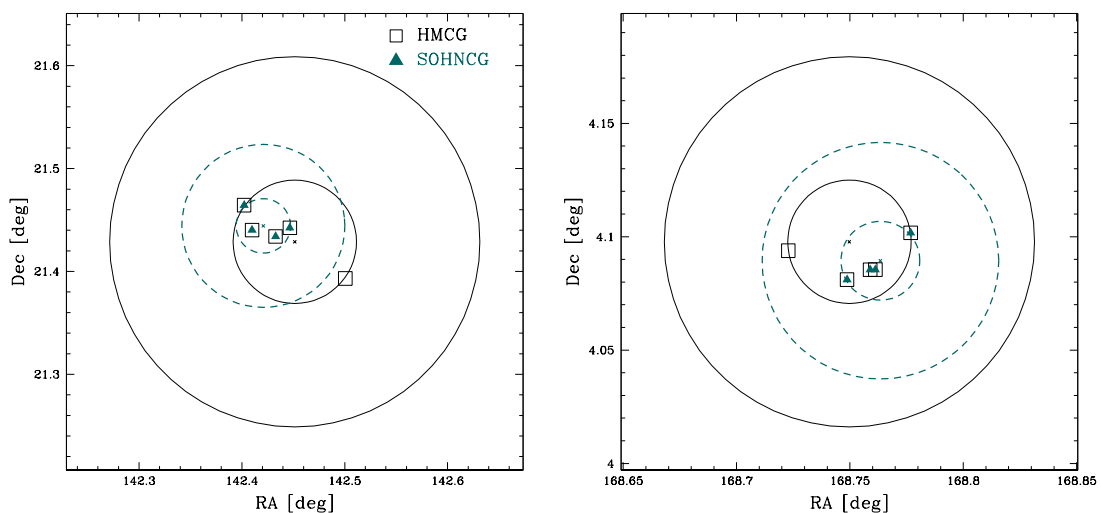
Es importante señalar que Sohn et al. (2016) realizó una comparación entre muestras restringidas a grupos cuya galaxia más brillante tenía un valor menor a 14,5 porque McConnachie et al. (2009) identificó sus CGs con esta restricción. Esta omisión de galaxias muy brillantes podría ser una mejor explicación de los diferentes rangos de desplazamiento al rojo de las muestras McConnachie et al. (2009) y Sohn et al. (2016) que la inclusión o no del criterio de aislamiento.

Finalmente, para profundizar nuestra comprensión de las diferencias encon-



**Figura 4.6:** Comparación de propiedades de CGs usando beanplots asimétricos (Kampstra, 2008). La región negra representa la distribución de densidad para los CGs identificados con el algoritmo modificado (HMCs), mientras que las regiones grises corresponden a las distribuciones de los CGs SOHNCGs. Ambas muestras fueron restringidas para realizar una comparación adecuada. Las líneas horizontales muestran los valores de las medianas para cada propiedad mientras que las líneas discontinuas muestran el intervalo de confianza correspondiente al 95% (Krzywinski & Altman, 2014). Esta figura fue realizada con el software R (R Core Team, 2013).





**Figura 4.7:** Grupos compactos en común entre las muestras HMCs y SOHNCs. Las galaxias miembro de los HMCs están representadas por cuadrados negros, mientras que los miembros de los SOHNCs son triángulos cian. El círculo interior muestra el círculo mínimo que encierra todos los miembros del grupo ( $\theta_G$ ), mientras que el círculo exterior representa el disco de aislamiento ( $3\theta_G$ ). Cada centro del grupo está representado por un punto. Panel izquierdo: Ejemplo de un CG embebido. El SOHNC (Grupo ID= 600) permanece dentro del HMC (Grupo ID= 53), mientras que no existen dentro del disco de aislamiento para el SOHNC. Panel derecho: Ejemplo de un subgrupo no aislado (SOHNC ID= 1115) que permanece dentro de un HMC (Group ID = 137).

tradas entre nuestros *HMCG* y *SOHNCG*, primero restringimos el análisis a los 49 grupos que son comunes en ambas muestras. Dado que la superposición de grupos no se ha hecho miembro por miembro, los grupos en común todavía pueden diferir en sus propiedades. De hecho, encontramos que las diferencias mencionadas anteriormente se mantienen incluso para las muestras que contienen solo los grupos presentes en ambas muestras. Se puede esperar que los *HMCG* sean más grandes en proyección porque los seleccionamos intencionalmente de esta manera: en los casos en los que hay agrupaciones de galaxias que cumplen todos los criterios de CGs dentro de agrupaciones más grandes que también cumplen con todos los criterios (CGs embebidos), preferimos mantener los grupos más grandes. Para desentrañar este problema, luego realizamos una comparación miembro por miembro y encontramos que solo siete casos (vea el panel izquierdo en la Figura 4.7) están embebidos en CGs de nuestra muestra. Además, observamos que hay 12 *SOHNCG* que son subgrupos dentro de nuestros *HMCG* que no se seleccionan como CG en nuestro algoritmo porque el aislamiento se rompe si las otras galaxias no se incluyen como parte de grupos (vea el panel derecho de la Figura 4.7). Finalmente, hay nueve grupos en los que la mitad de los miembros no se comparten, incluso cuando los centros están muy cerca, porque la galaxia más brillante es diferente y, por lo tanto, el rango de magnitud en el que se considera a los vecinos es diferente. Los 21 grupos restantes son exactamente iguales. En los dos primeros casos (que constituyen  $\sim 40\%$  de la muestra) los *SOHNCG* son básicamente subestructuras más pequeñas dentro de nuestros grupos que conducen a las diferencias en los tamaños que informamos anteriormente. Entonces, concluimos que los tamaños más pequeños de *SOHNCGs* podrían ser una consecuencia de la definición diferente de lo que es un CG para cada autor más que el resultado de los algoritmos en sí. Por otro lado, al analizar los 60 *SOHNCG* que no pudimos recuperar, encontramos que los perdimos porque no cumplieron con nuestro criterio de aislamiento, mientras que Sohn et al. (2016) no aplicó este criterio.

### 4.3. Resumen

En este capítulo, hemos presentado un nuevo algoritmo de identificación de CGs manteniendo el criterio de Hickson (1982) y Hickson et al. (1992).

El algoritmo ha sido aplicado a un catálogo lightcone construido a partir de galaxias del SAM de formación de galaxias (Henriques et al., 2015) corrido sobre la simulación Millennium I. Hemos testeado el nuevo algoritmo comparando la muestra obtenida con este algoritmo *modificado* y la muestra de CGs obtenida usando el algoritmo *clásico*.

Encontramos que el número de CGs identificados con el nuevo algoritmo es el doble que el clásico, mejorando de esta manera la completitud de la muestra. Pero por otro lado, encontramos que el algoritmo no afecta la fracción de grupos que son alineamientos aleatorios a lo largo de la línea de la visual, es decir, que el

*porcentaje de grupos físicamente densos también se mantiene. En el capítulo 6 abordaremos este problema y estudiaremos las posibles opciones para aumentar la pureza de las muestras utilizando información en el espacio real.*

*Además, presentamos un nuevo catálogo de CGs identificado usando la muestra de galaxias con información espectroscópica construida por Tempel et al. (2017) a partir del SDSS DR12. Este catálogo contiene 462 sistemas. Esta muestra de 462 grupos compactos Hickson modificado (HMCs) es el catálogo mas grande hasta el momento construido con información espectroscópica y que cumple con los criterios originales de Hickson: grupos pequeños, compactos, aislados, con sus galaxias con velocidades concordantes.*

*Este trabajo fue publicado en una revista internacional con referato (Díaz-Giménez et al. (2018b)).*

# Capítulo 5

## Análisis de la naturaleza y frecuencia de CGs en diferentes SAMs

*Como se ha mencionado con anterioridad, existen numerosos estudios durante los últimos casi 40 años avocados al análisis de la naturaleza y la frecuencia de CGs. En lo que respecta a los estudios realizados con SAMs, estos tópicos también han sido abordados. Sin embargo, en casi todos estos trabajos el común denominador es la realización del estudio sobre un SAM en particular, por lo que siempre queda el interrogante si estos resultados pueden ser dependientes de la gran cantidad de variables diferentes que uno puede enumerar entre los diferentes SAMs.*

*Por lo tanto, en este capítulo mostraremos un estudio complejo realizado utilizando diferentes SAMs de manera de entender la influencia, entre otras cosas, de los parámetros cosmológicos como también de la resolución de las simulaciones en la identificación de CGs. Para esto, hemos utilizado 9 catálogos lightcones distintos usando diferentes simulaciones (solo materia oscura), resoluciones, SAMs, bandas fotométricas, profundidades y algoritmos de CGs. En particular, analizamos 4 SAMs corridos en la simulación Millennium I, 2 SAMs corridos en la Millennium II y un SAM corrido en la simulación MultiDark.*

### 5.1. Los catálogos lightcones

*Hemos construido catálogos sintéticos con evolución (lightcones) en todo el cielo (i.e.  $4\pi$  sr), siguiendo un procedimiento similar al realizado por Zandivarez et al. (2014), el cual se describe en detalle en la sección 2.3. La mayoría están limitados en una magnitud aparente AB en el sistema de referencia del observador*

Muestra	Catálogos Lightcones			
	banda fotométrica	magnitud límite	# de galaxias	# de CGs
DLB	$r$	17.77	3 757 143	3387
G11	$r$	17.77	3 149 024	3175
GII	$r$	17.77	3 214 602	2558
G13	$r$	17.77	2 982 462	1682
HrI	$r$	17.77	3 087 401	1291
HrII	$r$	17.77	2 519 119	800
HkI	$K_s$	13.57	632 224	251
HrIc	$r$	16.54	611 008	276
SAG	$r$	17.77	2 941 613	723
cDLB	$r$	17.77	3 757 143	1812
cHrI	$r$	17.77	3 087 401	684
SDSS	$r$	17.77	557 517	462
c2MASS	$K_s$	13.57	408 618	85

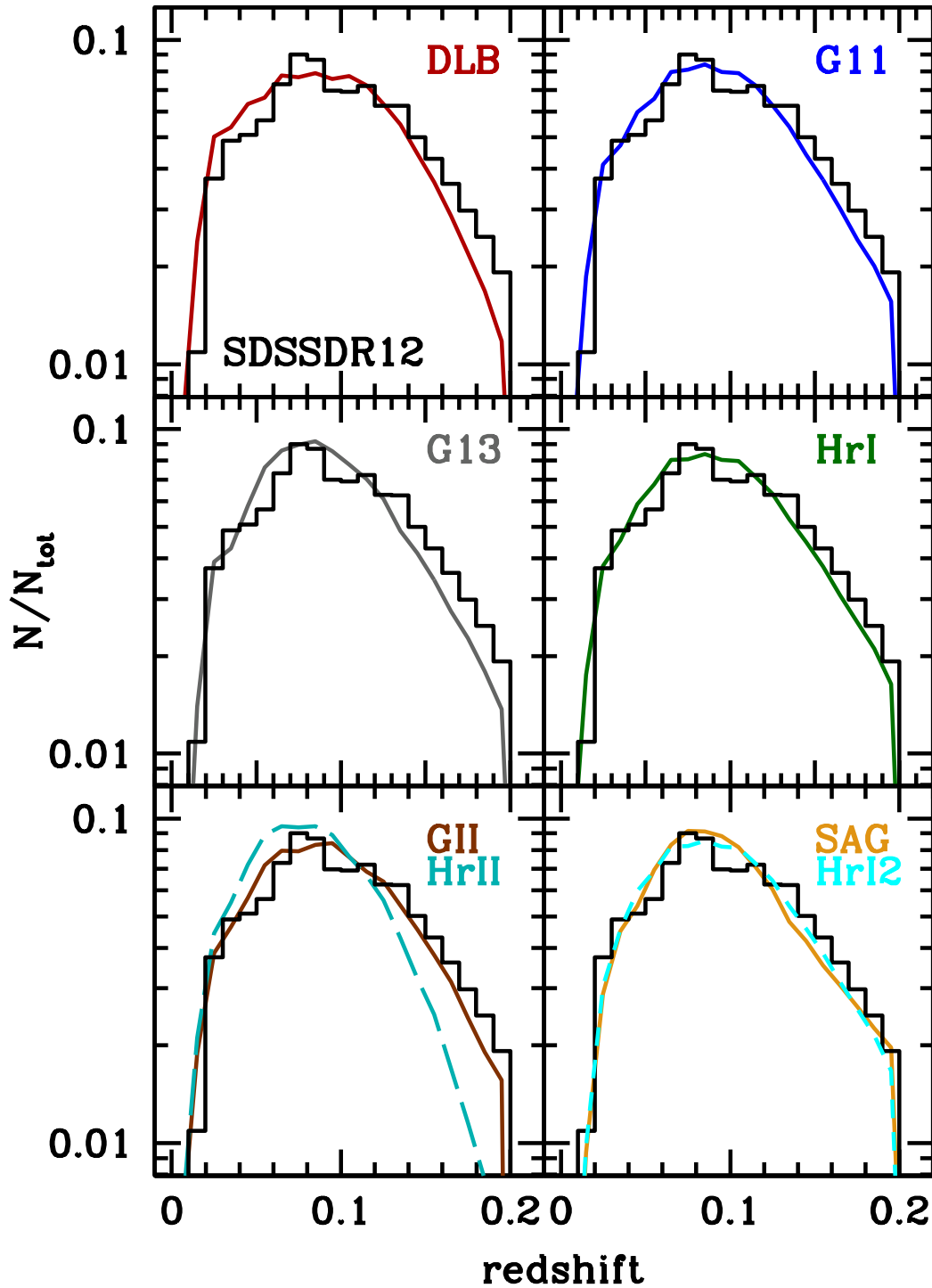
**Tabla 5.1:** CGs identificados en catálogos lightcones. HkI: es el mismo que HrI pero construido usando la banda fotométrica  $K_s$ . HrIc es construido similarmente HrI pero con un magnitud aparente límite menor. El número de CGs fue obtenido a partir del algoritmo de identificación modificado definido en el capítulo 4. cDLB: identificación con algoritmo clásico de CGs en el lightcone DLB; cHrI: identificación con algoritmo clásico de CGs en el lightcone HrI; SDSS: muestra observacional obtenida en el capítulo 4 dentro de un ángulo sólido de  $0,66 \pi$  sr; c2MASS: muestra observacional construida por Díaz-Giménez et al. (2012) dentro de un ángulo sólido de  $2,3 \pi$  sr.

de  $r \leq 17,77$ . Pero, también hemos construido catálogos lightcones seleccionando galaxias en la banda  $K_s$ <sup>1</sup> con una magnitud aparente límite de  $K_s = 13,57$ . Otro de los catálogos lightcones utilizados es limitado en  $r \leq 16,54$ , el cual es equivalente al límite en la banda  $r$  de el Two Micron All Sky Survey (2MASS)  $K_s$ . El número total de galaxias en cada catálogo esta mostrado en la tabla 5.1. La figura 5.1 muestra una comparación de la distribución de redshift normalizada de las galaxias sintéticas con magnitud límite de  $r < 17,77$  y redshift<sup>2</sup>  $z \leq 0,2$  sumado a la distribución observacional de redshift de las galaxias en el SDSS DR12 construido por Tempel et al. (2017).

Los 7 catálogos lightcones construidos a partir de diferentes SAMs recuperan muy bien la distribución de redshifts observacional, salvo en el rango  $0,14 < z < 0,2$  donde subestiman el número de galaxias.

<sup>1</sup>las magnitudes han sido convertidas desde el sistema AB al sistema Vega para imitar las magnitudes 2MASS

<sup>2</sup>el límite en redshift fue impuesto solo para construir esta figura para realizar una comparación adecuada con los datos observacionales



**Figura 5.1:** Distribuciones de redshifts normalizadas de galaxias en los diferentes catálogos lightcones limitados a una magnitud aparente en el sistema de referencia del observador  $r_{\text{lím}} = 17,77$  y redshift de  $z_{\text{lím}} = 0,2$  para los SAMs (*curvas*) y para los datos del SDSS DR12 (Tempel et al., 2017) usando la misma magnitud y límites en redshift (*histogramas negros*).

## 5.2. Muestras de CGs sintéticos

Identificamos CGs en los catálogos lightcones usando el algoritmo modificado, siguiendo el criterio básico definido por Hickson (Hickson, 1982; Hickson et al., 1992), el cual se encuentra detallado en la sección 3.1.2.

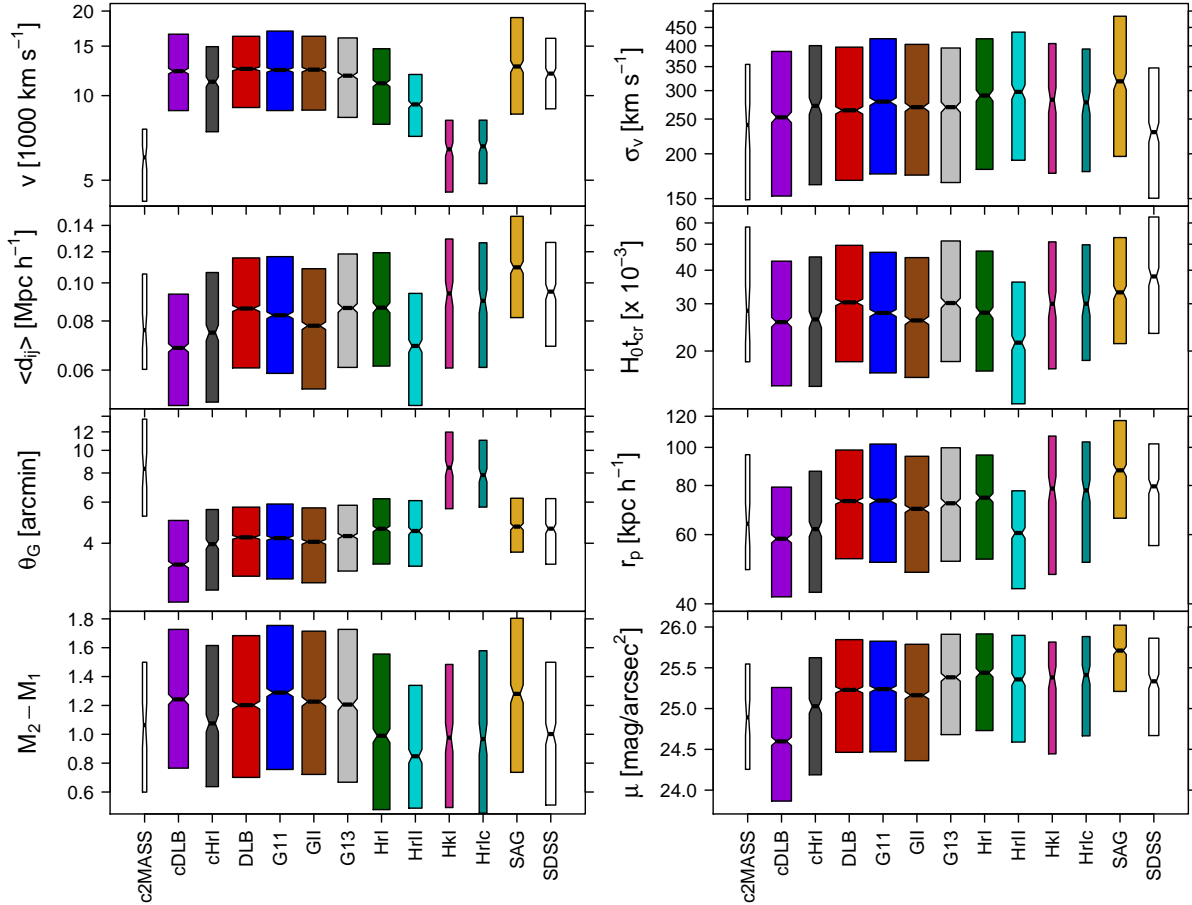
Los valores de  $\mu_{\text{lim}}$  (brillo superficial límite) depende de la banda fotométrica en la cual se realiza la selección. Para catálogos en la banda r-SDSS, adoptamos  $\mu_{\text{lim}} = 26,33$ , mientras que el valor equivalente para la muestra identificada en la banda  $K_s$  es  $\mu_{\text{lim}} = 23,6$  (ver capítulo 3). Los límites en magnitud se muestran en la tabla 5.1.

Dado que las galaxias en los lightcones simulados son partículas puntuales, y las galaxias son objetos extendidos, de la misma manera que lo hicimos en los capítulos anteriores, hemos corregido la muestra por este efecto siguiendo Díaz-Giménez & Mamon (2010) ya que esto puede modificar el número de galaxias detectables y cambiar la población de CGs. El número de CGs identificados en cada lightcone esta mostrado en la tabla 5.1.

Para cada grupo identificado en los diferentes catálogos hemos calculado las siguientes propiedades: el redshift; la dispersión de velocidades en al línea de la visual ( $\sigma_v$ ); el tiempo de cruce adimensional ( $H_0 t_{\text{cr}}$ ); la mediana de la separación proyectada entre galaxias ( $\langle d_{ij} \rangle$ ); el diámetro angular del círculo mínimo que encierra todos los miembros ( $\theta_G$ ); el radio proyectado del círculo mínimo ( $r_p$ ); el brillo superficial del grupo en las banda r ( $\mu$ ); la diferencia en magnitud absoluta entre la galaxia más brillante y la segunda galaxia más brillante en el sistema de referencia en reposo ( $M_2 - M_1$ ). Las fórmulas utilizadas para el cálculo de estas propiedades se detallan en el apéndice B.

En la figura 5.2 mostramos las distribuciones de las diferentes propiedades de CGs en todos los catálogos lightcones en forma de diagramas de caja (McGill et al., 1978; Krzywinski & Altman, 2014). También hemos incluido propiedades de CGs extraídos en la banda  $K_s$  a partir del catálogo 2MASS (Díaz-Giménez et al., 2012) usando el algoritmo clásico (llamado como c2MASS), y la muestra de CGs identificada en el capítulo anterior (Díaz-Giménez et al. (2018b)) usando el algoritmo modificado en la banda r en el SDSS DR12 (llamado como SDSS).

Existen dos propiedades mostradas en esta figura que dependen explícitamente de la distancia ( $v$ ,  $\theta_G$ ), por lo tanto, las muestras extraídas de HkI, HrIc y c2MASS muestran diferencias en esas propiedades por construcción. Para todas las demás propiedades que no dependen de la distancia, observamos buen acuerdo entre todas las muestras provenientes de diferentes SAMs y observaciones. Sin embargo, realizando un análisis estadístico mas detallado hemos encontrado algunas diferencias. Hemos usado el test Kolmogorov-Smirnov de dos muestras y el 95% de los intervalos de confianza para las medianas lo que nos permitió realizar un comparación más cuantitativa entre las muestras (ver tabla 5.2).



**Figura 5.2:** Distribuciones de propiedades de CGs identificados en diferentes catálogos lightcones. *Panel Izquierdo, desde el superior al inferior:* velocidad heliocéntrica, mediana de la separación proyectada intergaláctica, diámetro angular, gap en magnitud entre las galaxias mas brillantes. *Panel derecho, desde el superior al inferior:* dispersión de velocidades a lo largo de la línea de la visual estimada usando el estimador gapper (Beers et al., 1990), tiempo de cruce, radio proyectado, brillo superficial medio (para muestras identificadas en la banda  $K_s$ , el brillo superficial ha sido aumentado por 2,73 magnitudes para compararlo con la banda  $r$  de otros SAMs). Para realizar una comparación con propiedades de CGs observacionales sumamos las muestras: SDSS (identificadas por Díaz-Giménez et al. 2018b con el nuevo algoritmo) y la 2MASS (identificado por Díaz-Giménez et al. (2012) con el algoritmo clásico).



	$d_{ij}$		$M_2 - M_1$		$\sigma_v$		$H_0 t_{cr}$		$r_p$		$\mu$	
	KS- $p$	M-CI	KS- $p$	M-CI	KS- $p$	M-CI	KS- $p$	M-CI	KS- $p$	M-CI	KS- $p$	M-CI
c2MASS-cDLB	0.02	✓	0.04	✓	0.61	✓	0.07	✓	0.16	✓	0.07	✓
c2MASS-cHrI	0.19	✓	0.51	✓	0.30	✓	0.14	✓	0.35	✓	0.66	✓
SDSS-DLB	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	0.10	✓
SDSS-G11	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	0.09	✓
SDSS-GII	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×
SDSS-G13	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	0.77	✓
SDSS-HrI	0.89	✓	0.89	✓	< 0.01	×	< 0.01	×	0.03	✓	0.08	✓
SDSS-HrII	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	0.95	✓
SDSS-HkI	0.07	✓	0.70	✓	< 0.01	×	< 0.01	×	0.03	✓	0.51	✓
SDSS-HrIc	0.12	✓	0.90	✓	< 0.01	×	< 0.01	×	0.15	✓	0.51	✓
SDSS-SAG	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×	< 0.01	×

**Tabla 5.2:** Comparación de propiedades físicas de CGs extraídos a partir de SAMs y observaciones. La comparación fue realizada usando el *valor - p* del test Kolmogorov-Smirnov, y el intervalo de confianza del 95 % para las medianas. La primera columna indica las muestras que son comparadas, mientras que las columnas restantes muestran las propiedades que se muestran en la figura 5.2. Para cada propiedad, la columna KS- $p$  hace referencia al valor- $p$  del test Kolmogorov-Smirnov obtenido de la comparación de las muestras, mientras que la columna M-CI indica cuando los intervalos de confianza para las medianas de una dada propiedad se superponen (✓) o no (×).

A partir de este análisis, se puede notar que el algoritmo clásico, ambos SAMs (cDLB y cHrI) reproducen de buena manera las propiedades observacionales de CGs. Sin embargo, cuando usamos el algoritmo modificado, las muestras HrI, HkI y HrIc reproducen algunas propiedades ( $d_{ij}$ ,  $M_2 - M_1$ ,  $r_p$  y  $\mu$ ), pero no reproduce muy bien otras propiedades ( $\sigma_v$  y  $H_0 t_{cr}$ ). También se pueden observar algunos comportamientos diferentes cuando comparamos dos SAMs entre sí. Las muestras identificadas con el algoritmo clásico muestran el brillo superficial, radio proyectado y tiempo de cruce más bajos como consecuencia de la incompletitud de la indentificación. Los CGs extraídos a partir de los SAMs de Henriques et al. (2015) (cHrI, HrI, HrII, HkI y HrIc) presentan las medianas más bajas de las diferencias entre las magnitudes de las dos galaxias más brillantes (i.e., evita los CGs dominados por una sola galaxia), mientras que los CGs SAG muestran los gaps en magnitud más grandes. También, los HrII muestran las diferencias en magnitudes, radios proyectados y tiempo de cruce más bajos de todas las muestras, mientras que los CGs SAG se comportan de manera opuesta en esas propiedades.

### 5.3. Frecuencia y naturaleza de CGs

Para continuar con el análisis de los CGs, hemos estudiado la frecuencia y la naturaleza tridimensional de estos grupos y su dependencia con los SAM y parámetros cosmológicos. Para caracterizar la frecuencia de CGs, hemos medido la densidad espacial de CGs en cada lightcone simulado hasta la distancia mediana de la muestra menos profunda. Hemos definido

$$\eta_{60} = \frac{3N_{\text{CG}}(r < r_{60})}{\Omega r^3},$$

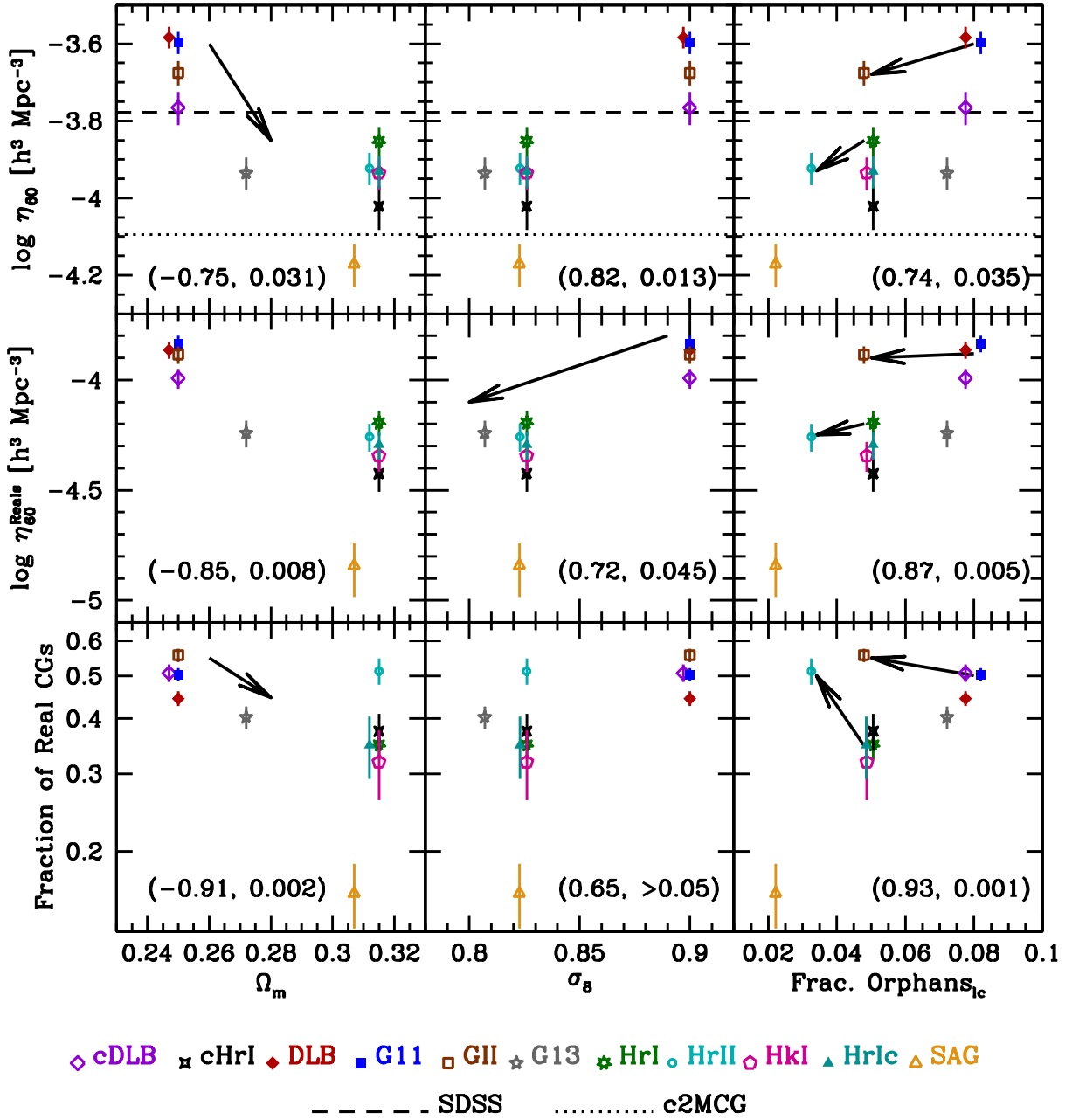
con un ángulo sólido  $\Omega = 4\pi$  y una distancia  $r_{60} = 60 h^{-1}\text{Mpc}$ , que es cercana a la mediana de las distancias comóviles de los centros de los grupos en los catálogos lightcones HkI y HrIc los cuales tienen el límite más restrictivo en flujo (ver panel superior izquierdo de la figura. 5.2).

Los paneles superiores de la figura 5.3 muestran que la densidad espacial de CGs en cada catálogo lightcone como función de los parámetros cosmológicos: densidad adimensional de materia,  $\Omega_m$  (izquierda), amplitud en densidad de fluctuaciones primordial a  $8 h^{-1}\text{Mpc}$ ,  $\sigma_8$  (centro) y fracción de galaxias huérfanas en el catálogo lightcone (derecha). En los paneles superiores hemos incluido la densidad espacial de dos muestras observacionales ya descritas: c2MASS (línea a puntos) y SDSS (línea discontinua). Dadas las limitaciones inherentes de los catálogos observacionales (incompletitud en redshift, colisión de fibras, saturación por brillo), la completitud de las muestras esta comprometida, por lo tanto, los valores de densidad encontrados en observaciones pueden ser tomados como un límite inferior. Los paneles superiores de la figura 5.3 indican que la densidad espacial de las muestras encontradas con el algoritmo clásico (cDLB y cHrI) están por encima de los valores de densidad espacial de la muestra de CGs 2MASS identificados con el mismo algoritmo. Sin embargo, los valores de las densidades espaciales obtenidas a partir de las muestras basadas en Henriques et al. (2015), G13 y SAG usando el nuevo algoritmo esta por debajo del umbral dado por el los CGs SDSS, por lo tanto esos SAMs parecen que pierden CGs comparados con las observaciones.

Por otro lado, para estudiar la naturaleza de los CGs, dividimos las muestras en diferentes categorías: CGs físicamente densos (*Reals*) y alineamientos aleatorios (CAs) siguiendo la clasificación proporcionada por Díaz-Giménez & Mamon (2010) (ver capítulo 4 para más detalles).

Los alineamientos aleatorios (CAs) pueden ocurrir de dos maneras: CAs de galaxias dentro de grupos (difusos) más grandes (CALGs) y CAs en el campo (CAFs), i.e. los filamentos orientados en la línea de la visual, donde las velocidades peculiares anulan por azar las diferencias en las velocidades causadas por el flujo de Hubble para permitir que el CG cumpla con el criterio de concordancia de velocidades. Por lo tanto,

- si las 4 galaxias mas cercanas en un CA CG pertenece solo a un grupo FoF,



**Figura 5.3:** Densidad espacial de CGs (*Paneles superiores*), densidad espacial de CGs *Reals*, (*Paneles medios*), y fracción de CGs *Reals* (*Paneles inferiores*) como función de parámetros cosmológicos: densidad adimensional de materia ( $\Omega_m$ ) (*paneles izquierdos*), amplitud de las fluctuaciones de masa en una esfera de  $8 h^{-1}$  Mpc ( $\sigma_8$ ) (*paneles centrales*), fracción de galaxias huérfanas en los lightcones (*paneles derechos*). Cada SAM es descrito con su correspondiente acrónimo definido en la tabla 5.1. En los *panel superiores*, la línea horizontal discontinua corresponde a la densidad espacial de la muestra observacional obtenida del SDSS DR12 por Díaz-Giménez et al. 2018b, mientras que la línea horizontal a puntos corresponde a la muestra observacional identificada con el algoritmo clásico 2MASSXCG por Díaz-Giménez et al. 2012. Las flechas en los paneles izquierdos y centrales indican la transición en parámetros cosmológicos manteniendo la mayoría de las otras cosas iguales (i.e. desde G11 hasta G13). Las flechas en los paneles derechos indican la transición en resolución (G11 a GII y HrI a HrII). En cada panel los números  $(r_s, p)$  indican coeficiente de Spearman ( $r_s$ ) y su correspondiente valor- $p$ , calculados solo para los lightcones construidos a partir de los SAMs que utilizan la simulación Millennium I.

el grupo es clasificado como CALG;

- si las 4 galaxias mas cercanas en un CA CG pertenecen a mas de un grupos FoF o ninguno, entonces los grupos se clasifican como CAF.

Para realizar esta subdivisión de grupos formados por alineamientos aleatorios (CAs) es necesario contar con una muestra de grupos FoF (difusos) en el cubo de la simulación. Por lo tanto, identificamos grupos de galaxias usando el algoritmo de percolación “amigos de los amigos”, (FoF, por sus siglas en inglés Friends-of-Friends) en el espacio real (Davis et al., 1985).

### Friends-of-Friends

El algoritmo asocia galaxias que comparten vecinos comunes. Para comenzar con la identificación, se toma una galaxia que no haya sido asignada previamente a un grupo, y se buscan a su alrededor galaxias vecinas que satisfagan que la separación entre ellas sea una distancia menor que un dado umbral. Si ninguna galaxia cumple con esta condición de proximidad, entonces dicha galaxia no tiene un grupo asociado; de lo contrario, la galaxia junto con todas sus vecinas son incluidas en un “grupo FoF”. Este procedimiento es aplicado iterativamente sobre las galaxias vecinas que cumplieron con la condición hasta que no se encuentren más miembros y, cada vez que se halla una galaxia que la satisface, es incluida en el grupo. El umbral es llamado longitud de percolación ( $r_0$ ) y está relacionada directamente con el valor de la sobredensidad con la que se desea identificar a los grupos. Este método de identificación de estructuras tiene importantes ventajas: en primer lugar, produce un único catálogo de grupos para cada valor elegido de  $r_0$ ; en segundo lugar, todos los miembros de un grupo definido para un  $r_0$  particular, serán miembros del mismo grupo para un catálogo definido con un valor de  $r_0$  más alto; en tercer lugar los grupos FoF no pueden cruzarse, lo que significa que una galaxia puede asignarse de forma única a un solo grupo FoF (para una longitud de percolación dada). Finalmente, el método no necesita hacer ninguna suposición sobre las formas de los grupos, lo cual permite la identificación de estructuras con formas diversas.

Por lo tanto, el algoritmo asocia galaxias especificando solamente una longitud de percolación o sobredensidad. En general, ésta longitud de asociación  $r_0$  se define como un factor  $b$  de la longitud media inter-objeto del catálogo a utilizar. Dicho factor  $b$  es el que típicamente se usa en la literatura para denotar la variación propuesta para la longitud de percolación  $r_0$ .

Para identificar grupos FoF en el cubo definido por la simulación, siguiendo a Zandivarez et al. (2014), debemos tener en cuenta que la sobredensidad de virialización cambia con el tiempo, y por lo tanto, la longitud de percolación varía con el redshift según:

$$\frac{\delta\rho}{\rho} = b^{-3}(z) = b_0^{-3} \left( 0,24 \frac{\Delta_{vir}(z)}{178} + 0,68 \right) \quad (5.1)$$

donde  $\Delta_{vir}(z)$  es la sobredensidad de materia encerrada en halos virializados que depende de la cosmología y del redshift según:

$$\Delta_{vir}(z) = 18\pi^2 \left[ 1 + 0,399 \left( \frac{1}{\Omega_m(z)} - 1 \right)^{0,941} \right] \quad (5.2)$$

$$\text{con } \left( \frac{1}{\Omega_m(z)} - 1 \right) = \left( \frac{1}{\Omega_0} - 1 \right) (1+z)^{-3}$$

En general, debido a que los halos de galaxias son más concentrados que los halos de materia oscura, para identificar halos que simulen las propiedades de halos de galaxias, se usa un contraste de densidad más alto de lo que se usaría para materia oscura. En consecuencia, en vez de usar el típico valor de  $b_0 = 0,2$  que se suele usar en la literatura para los halos de materia oscura, se usa  $b_0 = 0,14$  (Eke et al., 2004; Berlind et al., 2006; Zandivarez et al., 2014), lo cual se corresponde aproximadamente con una sobredensidad de contorno de 377 veces la media, para la cosmología WMAP1 a  $z = 0$ , 358 para WMAP7 y 328 para Planck, y permite la identificación de una muestra de grupos difusos.

Para encontrar el valor de  $b(z = 0)$  de la ecuación 5.1, se debe reemplazar en dicha ecuación el miembro derecho por  $b_0 = 0,14$  y evaluar los parámetros cosmológicos correspondientes a cada modelo cosmológico utilizado. De esta manera se obtienen grupos FoF que reflejan los grupos de galaxias observados.

### 5.3.1. Dependencia con parámetros cosmológicos

La variación de la densidad numérica de CGs *Reals* y la fracción de CGs *Reals* como función de los parámetros cosmológicos se muestran en el panel medio e inferior de la figura 5.3. La tendencia de la densidad espacial de los CGs *Reals* sigue el mismo comportamiento que la densidad espacial de la muestra total de CGs. Considerando las muestras construidas a partir de la misma simulación (MS), las fracciones de *Reals* decrece cuando aumenta  $\Omega_m$ , mientras no hay tendencias obvias con  $\sigma_8$ .

Los puntos más sobresalientes de la figura 5.3 se evidencian al comparar G11 y G13 (ver flechas), los cuales utilizan similares SAMs en una misma simulación (MS), pero corrida con dos cosmologías diferentes. Asumiendo un comportamiento como ley de potencias hemos encontrado que:

$$\eta_{60} \propto \Omega_m^\alpha, \quad (5.3)$$

$$\eta_{60}^{\text{Reals}} \propto \sigma_8^\beta, \quad (5.4)$$

$$f_{\text{Real}} \propto \Omega_m^\gamma, \quad (5.5)$$

donde

$$\alpha = \begin{cases} -8,7 \pm 1,4 & (\text{G11} \rightarrow \text{G13}) \\ -2,7 \pm 0,8 & (\text{todos MS}) \end{cases}, \quad (5.6)$$

$$\beta = \begin{cases} 8,5 \pm 1,5 & (\text{G11} \rightarrow \text{G13}) \\ 9,8 \pm 1,9 & (\text{todos MS}) \end{cases}, \quad (5.7)$$

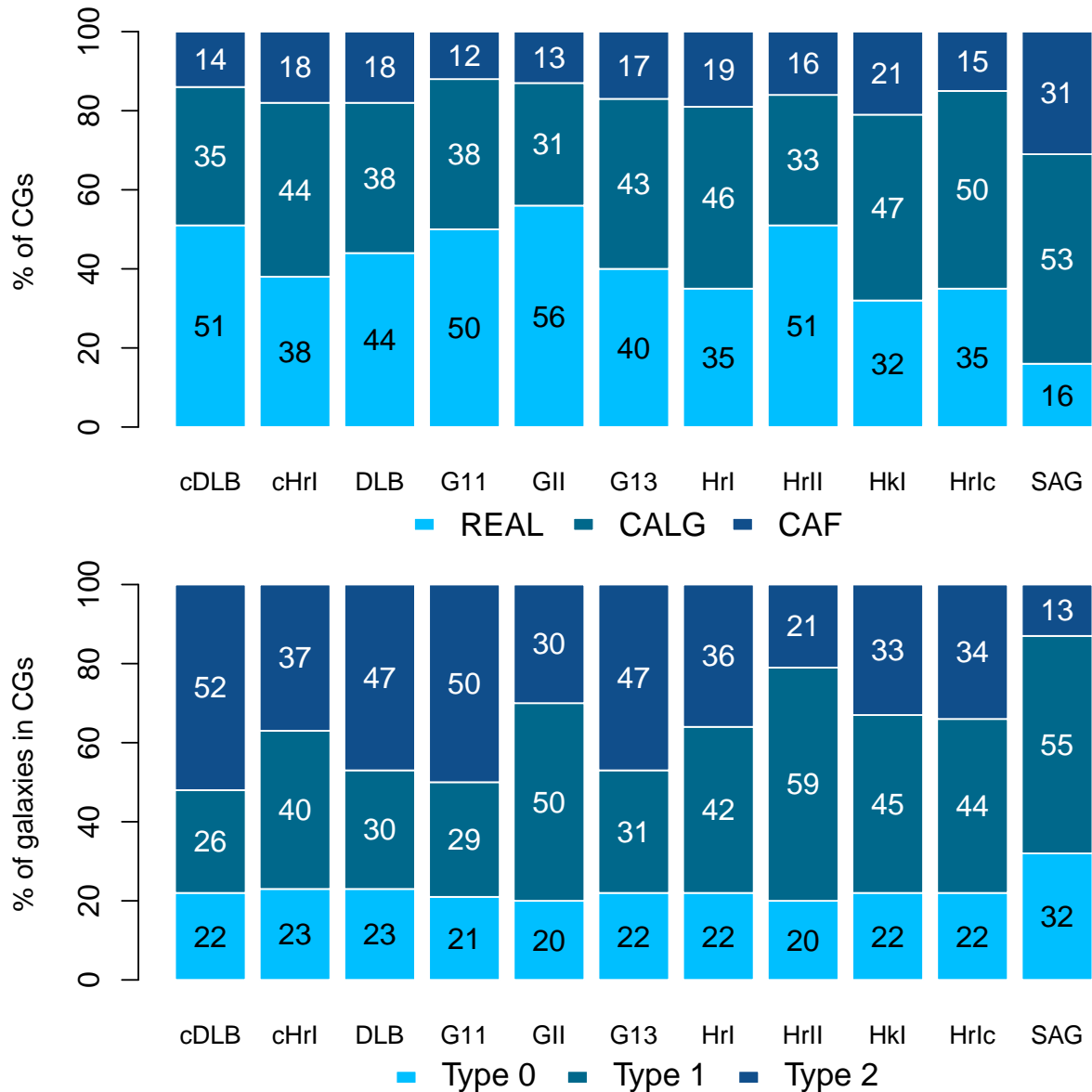
$$\gamma = \begin{cases} -2,7 \pm 0,8 & (\text{G11} \rightarrow \text{G13}) \\ -1,4 \pm 0,2 & (\text{todos MS}) \end{cases}. \quad (5.8)$$

Estas tendencias estadísticamente significativas pueden ser interpretadas de la siguiente manera:

1. un aumento en la contaminación del anillo de aislación de CGs cuando aumenta  $\Omega_m$  (ecuación [5.3] y [5.6]),
2. un aumento en la densidad espacial de grupos físicamente densos a medida que aumenta la normalización del espectro de potencia de materia (ecuación [5.4] y [5.7]),
3. un incremento en la fracción de alineamientos aleatorios con el incremento de  $\Omega_m$  (ecuación [5.5] y [5.8]).

Este análisis asume que las recetas físicas en los SAMs G11 y G13 son idénticas, lo cual no es del todo cierto. De hecho, acorde a la tabla 2 de Guo et al. (2013), relativo a G11, la eficiencia en formación estelar en G13 es la mitad de la que es en G11, los mecanismos de retroalimentación (feedback) son la mitad en G13 para el feedback de SN y AGN, mientras que el feedback térmico de SNE es más o menos el mismo en los dos SAMs. Sin embargo, las tendencias obtenidas con el aumento de  $\Omega_m$ : decrecimiento de la densidad espacial de CGs y decrecimiento de las fracciones de CGs físicamente densos, como así también el aumento de la densidad espacial de los CGs físicamente densos con el aumento de  $\sigma_8$ , se pueden ver para todos los SAMs que se muestran en la figura 5.3, y particularmente entre los SAMs que comparten la misma simulación MS con la misma resolución, tendencia que se encuentra detallada en las ecuaciones (5.6), (5.7) y (5.8) con “(todos MS)”. Hemos analizado la intensidad de esas correlaciones obtenidas a partir de la simulación MS usando el coeficiente de correlación de Spearman y su valor- $p$ , los cuales se muestran en la figura 5.3. Hemos encontrado una fuerte anti-correlación entre la densidad total, la densidad de CGs *Reals* y la fracción de CGs *Reals* con  $\Omega_m$ , y una fuerte correlación entre la densidad total y la densidad de CGs *Reals* con  $\sigma_8$ , pero no hay una correlación estadísticamente significativa de la fracción de CGs *Reals* con  $\sigma_8$ .

Por otro lado, en el panel superior de la figura 5.4 se muestra, para los diferentes catálogos lightcones, el porcentaje de CGs que han sido clasificados como *Reals*, CALGs y CAFs. El porcentaje de CGs clasificados como *Reals* varía desde 16% hasta 56% dependiendo de la muestra que esta siendo analizada, las muestras construídas a partir de la simulación con más alta resolución (GII y



**Figura 5.4:** *Panel Superior:* Porcentajes de las diferentes clases de CGs para los 9 catálogos lightcones: físicamente densos (*Real*), alineamientos aleatorios dentro de grupos difusos (CALG) y en el campo (CAF). *Panel Inferior:* Porcentajes de galaxias en CGs que son clasificadas como galaxias centrales de un grupo FoF (Type 0), galaxias centrales de un subhalo (Type 1), o una galaxia satélite o huérfana (Type 2).

HrII) presentan las fracciones más altas de *Reals*, mientras que las muestras extraídas de SAG, que es el más pobre en resolución espacial (0.4 dex) como también en masa (0.2 dex) presenta el menor porcentaje de *Reals*. Entre las diferentes categorías de alineaciones aleatorias, en todos los catálogos lightcones, el porcentaje de CALGs es mayor que el porcentaje de CAFs. Las alineaciones aleatorias en el campo aumentan al aumentar el parámetro de densidad cosmológica como:

$$f_{\text{CAF}} \propto \Omega_m^\delta, \quad (5.9)$$

con

$$\delta = \begin{cases} 4,1 \pm 1,7 & (\text{G11} \rightarrow \text{G13}) \\ 1,1 \pm 0,3 & (\text{todos MS}) \end{cases}, \quad (5.10)$$

i.e. una mayor dependencia que la fracción de grupos físicamente densos con  $\Omega_m$  (ecuación [5.5]), mostrando cómo el aumento de la densidad del Universo lleva a alineamientos más frecuentes de galaxias en el campo. El ajuste utiliza todas las muestras basadas en la simulación MS y produce una dependencia ligeramente más débil con  $\Omega_m$  de  $f_{\text{CAF}}$  en comparación con los  $f_{\text{Real}}$ .

### 5.3.2. Dependencia con la fracción de galaxias huérfanas

Como ya hemos discutido en el capítulo 2, las diferencias en el tratamiento de las galaxias huérfanas en los SAMs tienen un impacto importante en el clustering de las galaxias a escalas pequeñas, las cuales son las escalas que nos interesan. Por lo tanto, hemos examinado como la fracción de galaxias huérfanas en los catálogos lightcones afecta la ocurrencia de CGs.

El panel derecho de la figura 5.3 muestra la densidad espacial de CGs (panel superior), la densidad espacial de CGs *Reals* (panel medio) y la fracción de CGs *Reals* (panel inferior) como función de la fracción de galaxias huérfanas en los lightcones en los que fueron identificados. El porcentaje de huérfanas en los lightcones varía de  $\sim 2$  hasta  $\sim 9\%$ . El panel superior de la figura 5.3 muestra que mientras más alta sea la fracción de galaxias huérfanas, más alta será la densidad de CGs. Sin embargo, podría ser complicado mezclar todas las muestras ya que la frecuencia de las galaxias huérfanas en los SAMs no solo está determinada por las recetas particulares, sino que también está vinculada a la resolución de la simulación de materia oscura que usaron para seguir los árboles de fusión. Considerando solo las muestras extraídas de los catálogos lightcones basado en la MS, el coeficiente de correlación de Spearman (mostrado en figura 5.3) indica fuertes correlaciones entre las 3 propiedades analizadas y la fracción de huérfanas: mientras más alta sea la fracción de huérfanas, mayor es la densidad espacial de CGs y de *Reals* CGs. Especialmente, mientras más alta es la fracción de huérfanas, más alta es la fracción de *Reals*.

Con respecto al tipo de galaxias que conforman a los CGs, en el panel inferior de la figura 5.4, mostramos los porcentajes de galaxias de diferentes tipos que pueblan los CGs. Las galaxias centrales (Type 0) son casi una cuarta parte de



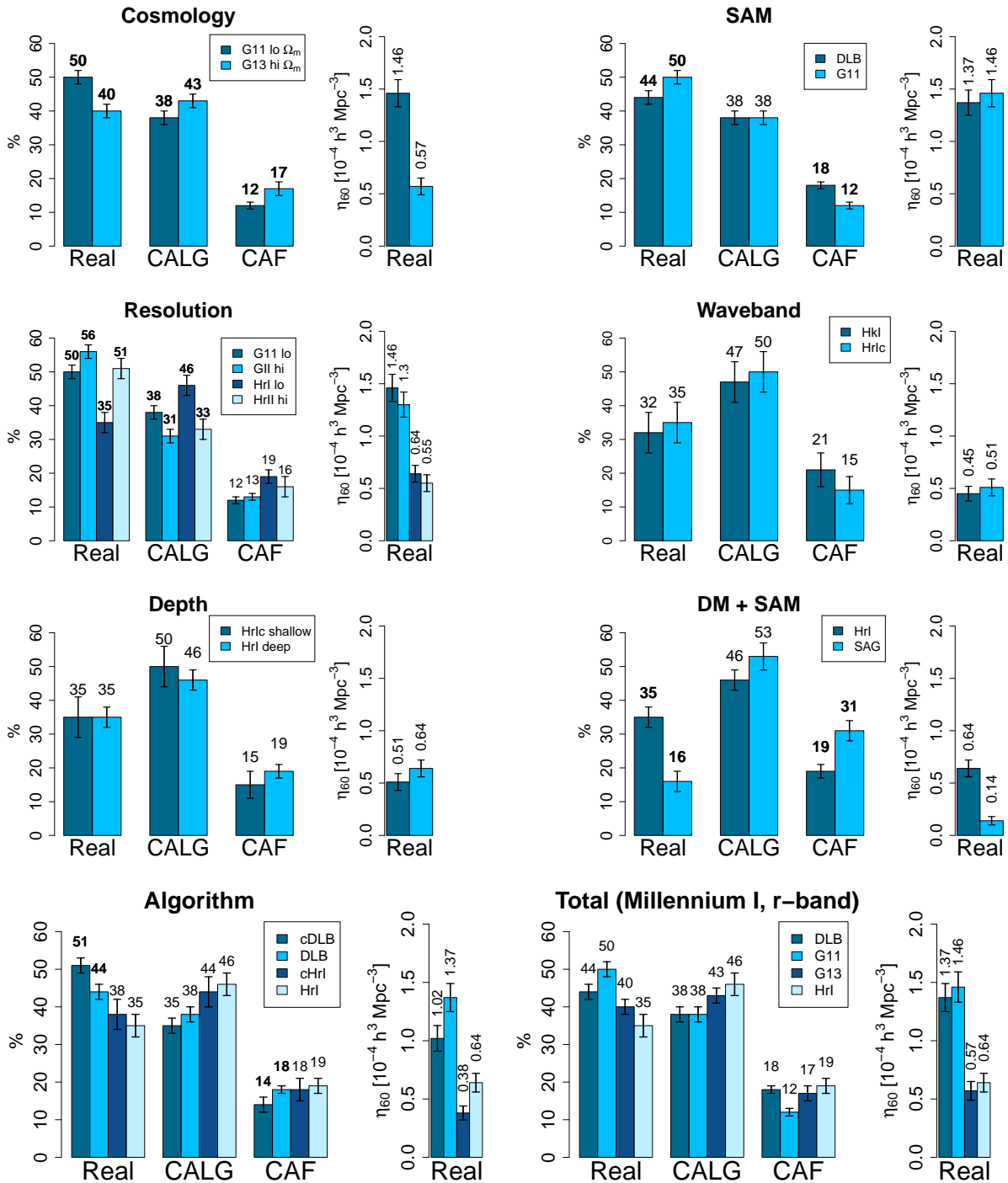
las galaxias en CGs en todos los catálogos. El porcentaje de galaxias huérfanas (Type 2) varía de 13 a 52% dependiendo del SAM y de la resolución de la simulación. Incluso, si solo se consideran galaxias en CGs *Reals*, la fracción de galaxias huérfanas es más alta que en las muestras totales (van desde 19 a 61%), lo cual señala la importancia de las galaxias huérfanas en los sistemas de gran densidad. El aumento de CGs *Reals* con la fracción de galaxias huérfanas conducirá a un exceso del número total de CGs pero con un correlación más pequeña y también conducirá a un exceso en la fracción de CGs *Reals* (como se muestra en los paneles derechos de la figura 5.3).

Una de las posibles causas de las correlaciones con la fracción de huérfanas pueden ser el hecho de que las huérfanas tienen posiciones inciertas desde que ellas pierden sus subhalos y por lo tanto sus posiciones son inferidas usando recetas para simular su evolución orbital incluyendo fricción dinámica. Algunos SAMs permiten que las huérfanas vivan demasiado tiempo, mientras que otros permiten que desaparezcan demasiado rápido. En el primer caso, deberíamos ver un exceso de galaxias en las regiones centrales donde los subhalos deberían verse más afectados por fuerzas de marea y calentamientos tidal y destruirse o reducirse a menos del número mínimo permitido de partículas. Este excedente de galaxias en las regiones internas de los halos permitirá formar más CGs (cuando esté aislado), lo que puede explicar las correlaciones positivas entre la densidad espacial de CGs y la fracción de huérfanos. Pujol et al. (2017) muestra que las galaxias huérfanas en el SAM SAG tienden a estar en las afueras de los grupos, mientras que las huérfanas en HrI no muestran tendencia con el radio.

Es interesante comparar las fracciones de galaxias huérfanas en SAMs similares, pero con diferentes resoluciones. La figura 5.3 muestra que las fracciones de huérfanas en el lightcone disminuyen típicamente en un 40 por ciento al pasar de SAMs de baja resolución (G11 y HrI) a alta (GII y HrII, respectivamente), mientras que la correspondiente fracción de CGs *Reals* aumenta ligeramente (G11 a GII) o en un factor 1.5 (HrI a HrII) (ver flechas en el panel inferior derecho de la figura 5.3). La fracción de CGs *Reals* parece menos dependiente de la fracción huérfanas a mayor resolución. Esto puede deberse al hecho de que los límites de flujo relativo de los CGs evitan que contengan las galaxias de menor masa (por lo tanto, subhaloes de menor masa) en los SAM basados en MS-II, por lo tanto, estos CGs tienen menos huérfanas y la fracción de *Reals* depende menos en la fracción de huérfanas.

### 5.3.3. Dependencias con otros parámetros

Dado que las muestras difieren en una serie de características, en la figura 5.5, mostramos la comparación de los porcentajes de los CGs clasificados como *Reals*, CALGs y CAFs entre pares de muestras que comparten características en común pero difieren en otra a la vez, la cual puede ser cosmología, modelo semianalítico, resolución, banda, profundidad, o algoritmo de identificación.



**Figura 5.5:** Fracción de CGs clasificados como físicamente densos (*Reals*), alineamientos aleatorios en grupos difusos (CALG) o en el campo o en filamentos (CAF). Cada panel muestra una comparación diferente en función de la cosmología, modelos semianalíticos (SAM), resolución de la simulación, banda fotométrica (waveband), o magnitud aparente límite (depth). También mostramos una comparación de todas las muestras obtenidas a partir de la simulación Millennium I en la banda  $r$  con  $r_{\text{lim}} = 17,77$ . Las barras de error muestran el intervalo de confianza del 95% binomial para cada porcentaje computado como  $\pm 1,96\sqrt{f(1-f)/N_{\text{CG}}}$ , donde  $f$  es la fracción. Los valores porcentuales marcados en negrita indican resultados estadísticamente diferentes teniendo en cuenta los intervalos de confianza (para todos los paneles, excepto para la comparación de las muestras Millennium - panel inferior derecho).

Una comparación detallada nos proporciona los siguientes resultados:

- para entender la relevancia de los parámetros cosmológicos en la formación de CGs, comparamos G11 y G13 (= SAM,  $\neq$  cosmología, = resolución en masa, = profundidad, = banda, = algoritmo de identificación): cuanto menor sea la densidad de materia ( $\Omega_m$ ), mayor será la fracción de *Reals* y menor será la fracción de CAF, mientras que la fracción de CALG permanecerá casi sin cambios (ver panel superior izquierdo de la figura 5.5).
- para entender la influencia de los SAMs, usamos G11 y DLB ( $\neq$  SAM, = cosmología, = resolución en masa, = profundidad, = banda, = algoritmo de identificación): el modelo G11 es más eficiente para identificar CGs *Reals* que el DLB. Esta disminución en la fracción de *Reals* en DLB es trasladada a un incremento de la fracción de sistemas CAF (ver panel superior derecho de la figura 5.5).
- para entender la influencia de la resolución espacial y en masa en la ocurrencia de CGs, comparamos G11 y G12, o HrI y HrII (= SAM, = cosmología,  $\neq$  resolución en masa, = profundidad, = banda, = algoritmo de identificación): cuanto mayor sea la resolución de masa de la simulación, mayor será la fracción de CG *Reals* y menor la fracción de CALG (ver panel izquierdo de la segunda fila de la figura 5.5).
- para evaluar la influencia de la banda fotométrica, analizamos HrI y HkI (= SAM, = cosmología, = resolución en masa, = profundidad,  $\neq$  banda, = algoritmo de identificación): no existen diferencias significantes en la fracción de *Reals*, CALG o CAF identificados en las diferentes bandas (ver panel derecho de la segunda fila de la figura 5.5).
- para entender la influencia de la profundidad del catálogo, usamos HrI y HrIc (= SAM, = cosmología, = resolución en masa,  $\neq$  profundidad, = banda, = algoritmo de identificación): no existen diferencias significativas en la fracción de CGs de ningún tipo cuando se utilizan catálogos lightcones restringidos en diferentes límites en flujo (ver panel izquierdo de la tercer fila de la figura 5.5).
- para comparar las diferentes simulaciones de N-cuerpos y diferentes SAM de formación de galaxias con una única cosmología, comparamos HrI con SAG ( $\neq$  SAM,  $\neq$  simulación, = cosmología, = profundidad, = banda, = algoritmo de identificación): SAG produce una fracción más pequeña de CGs *Reals* (menos de la mitad que en HrI), mientras que tiene una fracción más alta de CALG y CAF (ver panel derecho de la tercer fila de la figura 5.5).
- para entender la relevancia al usar diferentes algoritmos de identificación, analizamos cDLB y DLB, o cHrI y HrI (= SAM, = cosmología, = resolución en masa, = profundidad, = banda,  $\neq$  algoritmo de identificación): la fracción de CGs *Reals* cuando usamos el algoritmo clásico o el modificado es dependiente del SAM. Para DLB, la fracción de *Reals* es más alta cuando usa el

Catálogo Algoritmo	DGM10	DG+12		DGZ15	TDGZ16		DGZT18		este trabajo		
	cubo	lightcone		lc H+12	lc H+12		lightcone		lightcone		
	HC	HC		HC	HC		HC	HM	HC	HM	
Banda	$R$	$r$	$r$	$K_s$	$r$	$K_s$	$r$	$r$	$r$	$r$	$K_s$
Mag. lim	17,44	16,3	17,77	13,57	16,54	13,57	17,77	17,77	17,77	16,54	13,57
DLB	59	53	58	–	–	–	–	51	44	–	–
G11	–	–	–	54	52	56	–	–	50	–	–
GII	–	66	69	–	–	–	–	–	56	–	–
HrI	–	–	–	–	–	–	38	35	38	35	32

**Tabla 5.3:** Descripción de las muestras y porcentajes de CGs *Reals* en trabajos previos. Los acrónimos son: DGM10 (Díaz-Giménez & Mamon, 2010), DG+12 (Díaz-Giménez et al., 2012), DGZ15 (Díaz-Giménez & Zandivarez, 2015), TDGZ16 (Taverna et al., 2016), y DGZT18 (Díaz-Giménez et al., 2018b). lightcone H+12 hace referencia al catálogo ( $4\pi$  sr) construido por Henriques et al. (2012) usando el SAM G11; los tamaños de las galaxias fueron estimados utilizando las recetas de Shen et al. (2003) para el caso de DGM10 y DG+12, y de Lange et al. (2015) para DGZ15, TDGZ16 y DGZT18.

*algoritmo clásico en vez del modificado; sin embargo, estas diferencias no son significantes cuando comparamos cHrI y HrI (esto fue demostrado por Díaz-Giménez et al. (2018b)) (ver panel izquierdo inferior de la figura 5.5) .*

### Comparación con trabajos previos

*En la tabla 5.3 hemos comparado nuestros resultados en la fracción de CGs que son *Reals* con las fracciones que calculamos en trabajos anteriores. Todos los resultados son consistentes, dadas las diferencias en catálogos lightcones, bandas fotométricas, magnitudes límites, algoritmos y asignación de tamaños a galaxias simuladas (blending).*

*Díaz-Giménez & Mamon (2010) (DGM10) y Díaz-Giménez et al. (2012) (DG+12) han usado el SAM DLB (y otros SAM no considerados en este trabajo) para predecir fracciones ligeramente mayores de CGs *Reals* que las encontradas en este trabajo. Las pequeñas diferencias son causadas por las bandas fotométricas, los catálogos lightcones y el blending. Similarmente, la fracción de CGs *Reals* encontrada por Díaz-Giménez & Zandivarez (2015) (DGZ15) y Taverna et al. (2016) (TDGZ16, capítulo 3) fueron levemente más altas, debido a que en esos trabajos se utilizó el algoritmo clásico, y por nuestros resultados sabemos que el algoritmo modificado produce un fracción más baja de grupos *Reals*. También la fracción de CGs *Reals* encontradas en el SAM GII fue más alta en Díaz-Giménez et al. (2012) (DG+12), pero de nuevo estamos comparando dos algoritmos diferentes. Finalmente, las fracciones de CGs *Reals* obtenidas en el capítulo anterior (Díaz-Giménez et al. (2018b), DGZT18) son las mismas que en este trabajo.*

## 5.4. Resumen

*En este trabajo hemos construido diferentes catálogos lightcone con diferentes SAMs, diferentes bandas y diferentes magnitudes límites para estudiar la ocurrencia y naturaleza de CGs en diferentes SAMs. Encontramos que la densidad espacial y la naturaleza de CGs en los catálogos depende de los parámetros cosmológicos de las simulaciones de materia oscura (WMAP1, WMAP7 y Planck) donde fueron aplicados los SAMs, pero no así de los límites en flujo o la banda fotométrica que fueron utilizados para identificar los CGs en el espacio de redshift. Por otro lado, encontramos que el porcentaje de grupos físicamente densos aumenta para los SAMs con mayor resolución espacial y en masa, siendo los SAMs con menor porcentaje de galaxias huérfanas en CGs. Sin embargo, para todos los SAM, cuando aumenta el porcentaje de galaxias huérfanas, también aumenta la densidad de CGs Reals, siendo la fracción de huérfanas muy importante para la identificación de estos objetos en simulaciones.*

*El número de CGs clasificados como alineamientos aleatorios a lo largo de la línea de la visual (CALG o CAF) representa igual o más de la mitad de la muestra de CGs, llegando en algunos casos a un  $\sim 85\%$ . En el capítulo siguiente procuraremos cambiar el enfoque a la hora de estudiar la naturaleza de las muestras de CGs identificadas con el criterio de Hickson. Esto nos permitirá entender las posibles causas por las cuales los porcentajes de grupos espurios parecen ser altos para diferentes muestras de CGs.*

# Capítulo 6

## Naturaleza 3D de los CGs

*Como hemos visto en los capítulos anteriores, los criterios de identificación de grupos compactos de galaxias dependen en gran medida de las propiedades observacionales de sus galaxias como también de las propiedades utilizadas para definir dichos criterios. Debido a esto, el objetivo de este capítulo ha sido desarrollar un criterio independiente de propiedades observacionales capaz de identificar grupos físicamente densos y aislados en el espacio real, libres de sesgos observacionales. Contar con una muestra de grupos compactos en el espacio real nos permitirá un estudio exhaustivo sobre la pureza y la completitud de las muestras de grupos compactos identificados en catálogos observacionales con los criterios usuales, estudio que hasta el momento no ha sido abordado en la literatura.*

### 6.1. Criterio 3D para identificar CGs

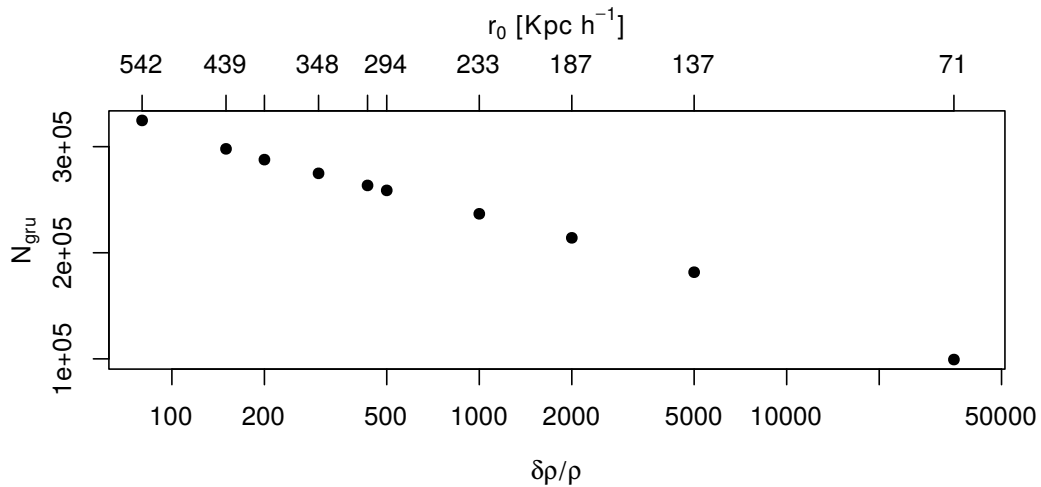
*Idealmente, los grupos compactos son grupos con un bajo número de miembros, muy densos y localmente aislados, en donde se produce una alta tasa de interacciones entre galaxias que han evolucionado relativamente aisladas del entorno en gran escala. Aquí, hemos desarrollado un criterio de identificación de CGs aplicado sobre simulaciones numéricas ejecutadas con distintos parámetros cosmológicos combinados con diferentes modelos semianalíticos de formación de galaxias. Inicialmente, construimos este criterio tridimensional en el espacio real (sin distorsión por velocidades peculiares), usando sólo propiedades físicas absolutas provenientes de las simulaciones para la identificación.*

### 6.1.1. Compacticidad

Dado que los grupos compactos deberían ser grupos altamente densos, hemos definido la compacticidad de los CGs en 3D basándonos en la sobredensidad de contorno de los halos de galaxias sintéticas. Para ello, comenzamos identificando grupos de galaxias usando el algoritmo de percolación FoF en el espacio real (Davis et al., 1985). La descripción del algoritmo se detalló previamente en la sección 5.3.

Hemos adoptado aquí un límite del contraste de sobredensidad alto para asegurar la alta densidad esperada para los grupos identificados. Para poder decidir este límite, hemos estudiado como cambia la identificación de grupos en función de dicha sobredensidad.

Al introducir el valor  $(\delta\rho/\rho)_{lim}$  a la ecuación 5.1, obtenemos el valor  $b(z=0)$  que es parámetro necesario para calcular la longitud de percolación  $r_0$  para luego identificar grupos densos. Como se mencionó anteriormente, la longitud de percolación es una fracción de la densidad media de galaxias en la simulación ( $\bar{\rho}$ ) y se obtiene de la siguiente manera:  $r_0 = \bar{\rho} \times b(z=0)$ . En la figura 6.2 se muestran los números de grupos identificados para diferentes valores de sobredensidad y sus correspondientes longitudes de percolación.



**Figura 6.1:** Número de grupos con 4 o más miembros identificados en el SAM G11, en función de la sobredensidad utilizada para seleccionar estos grupos. En el eje superior, se muestra la equivalencia entre la sobredensidad utilizada y la longitud de percolación.

Hemos seleccionado el valor de sobredensidad 5000 ya que es equivalente a un valor de separación entre galaxias comparable con la separación media entre galaxias en los grupos compactos observacionales identificados en el SDSS DR12 en el capítulo 4 (en la figura 4.2 se muestra la separación media proyectada entre galaxias, que podemos traducir a separación 3D con un factor

multiplicativo  $\pi/2$  para hacer la comparación). Esta sobredensidad supera en al menos un factor diez el contraste obtenido para grupos difusos en el capítulo anterior ( $\frac{\delta\rho}{\rho} \geq 433$ , para G11).

Una vez que los grupos fueron identificados con  $\delta\rho/\rho \geq 5000$ , hemos tomado sólo aquellos con 4 o más miembros,  $N \geq 4$ . A estos grupos los denominamos “grupos densos”.

### 6.1.2. Aislamiento

Una vez que tenemos seleccionados grupos sobredensos, nos interesan aquellos que se encuentren relativamente aislados en el universo, en donde sus galaxias tengan poca interacción con el entorno. El aislamiento de los grupos densos fue definido en tres diferentes etapas utilizando la información espacial 3D, es decir, las posiciones comóviles de las galaxias sintéticas en el cubo de la simulación,  $\vec{r} = (x, y, z)$  y sus propiedades absolutas, como la masa estelar. En general, el procedimiento establecido para definir grupos aislados puede describirse con los siguientes pasos: primero, hemos eliminado de la muestra de grupos densos a aquellos que consideramos que son subestructura de un grupo difuso, luego, eliminamos grupos que son sistemas donde en realidad hay sólo una galaxia grande junto a sus pequeñas galaxias satélites como el sistema Vía Láctea y las Nubes de Magallanes, y por último, eliminamos grupos que no pueden considerarse aislados en base a la densidad local.

#### a) Subestructura

En esta etapa hemos seleccionado sólo los grupos densos que no son subestructura de un grupo con sobredensidad más baja, es decir, que no están embebidos en un grupo de menor densidad (difuso), ya fuera que son el núcleo de un grupo difuso o alguna otra estructura densa en el interior de los difusos. Para ello, primero hemos identificado grupos difusos en 3D de la manera que mencionamos en la sección 5.3, y obtuvimos una muestra de grupos difusos FoF en los cubos de cada simulación utilizada. Una vez que fueron identificados, eliminamos de la muestra de difusos a los grupos que se encuentran en la muestra de grupos densos y que son exactamente el mismo grupo. Esto es posible ya que la sobredensidad utilizada en las identificaciones es un límite inferior, y pueden existir algunos grupos difusos cuya sobredensidad sea mayor a 5000 y que sus miembros sean los mismos que algún denso.

Todo grupo denso que está embebido en un grupo difuso más numeroso, es decir, que comparte sus miembros con algún difuso, puede ser considerado subestructura de ese difuso, pero hay algunos casos particulares que necesitan ser tenidos en cuenta. Para seleccionar los grupos densos considerados subestructura hemos definimos un parámetro “p” el cual nos



permite conocer el porcentaje de miembros coincidentes con respecto al total de miembros del grupo difuso:

$$p = \frac{\# \text{ miembros coincidentes}}{\# \text{ miembros grupo difuso}} \quad (6.1)$$

donde el número de miembros coincidentes son los miembros que pertenecen a ambos grupos.

Este parámetro nos da información de cuán más numeroso es el difuso en el cual se encuentra el grupo denso. Si  $p \leq 0,5$  significa que el grupo difuso tiene más de dos veces el número de miembros coincidentes. Por otro lado, si  $p > 0,5$ , esto quiere decir que el grupo difuso tiene a lo sumo dos veces el número de miembros coincidentes.

Por lo tanto, para el primer caso ( $p \leq 0,5$ ) el grupo denso está embebido en un difuso el doble de grande, y uno puede tender a pensar que estos tipos de grupos son los que podemos considerar como subestructura. Pero este parámetro sólo no es suficiente ya que, depende de la población de cada grupo. Si elegimos seleccionar como subestructura a grupos densos con  $p$  menor a 0.5, eliminamos a grupos pobres con pocas coincidencias pero donde el difuso no es mucho más numeroso y no sería del todo correcto pensar al grupo denso como una subestructura dentro de otro sistema (por ejemplo, un grupo denso de 4 miembros y un grupo difuso con 9 miembros, donde coinciden todos los miembros). Por otro lado, si  $p > 0,5$  el grupo difuso comparte una gran cantidad de sus miembros con el grupo denso. Por lo tanto, uno puede pensar que son grupos muy parecidos: un grupo denso embebido en un grupo difuso el cual tiene unos pocos miembros más. El grupo denso se vería como un grupo en el espacio real acompañado por unas galaxias más a su alrededor. Pero nuevamente, qué grupo considerar subestructura depende del número de miembros de los grupos involucrados. Según nuestro criterio, el grupo podría ser considerado aislado localmente ó no.

Por estas razones, hemos tomado un criterio relajado para eliminar a grupos densos por ser subestructura de difusos, permitiendo la supervivencia de casos poco claros, pero sabiendo que luego aplicaremos otros criterios de aislamiento local, que se encargarán de incluir o eliminar esos casos problemáticos utilizando otras propiedades más robustas.

Definimos que los grupos densos son subestructura de grupos difusos (con alto grado de certeza) cuando cumplen con las siguientes condiciones:

- Si  $p \leq 0,5$ , son subestructura si  $Nm_{\text{denso}} - Nm_{\text{difuso}} > 10$
- Si  $p > 0,5$ , son subestructura si  $Nm_{\text{denso}} > 20$

donde,  $Nm_{\text{denso}}$  es el número de miembros del grupo denso y  $Nm_{\text{difuso}}$  es el número de miembros del grupo difuso que lo contiene.

Los grupos que no cumplen con esas condiciones son grupos densos pequeños donde su difuso asociado tiene unos pocos miembros más. Estos

grupos serán analizados en los siguientes pasos para seguir estudiando su aislamiento.

### b) Satélites

Así como verificamos si un grupo denso es subestructura de un grupo difuso, queremos averiguar si estos grupos densos no están formados por una galaxia brillante junto a varias galaxias satélites considerablemente más débiles. El ejemplo clásico de este tipo de grupos es el sistema conformado por nuestra galaxia Vía Láctea (VL) y las Nubes de Magallanes, las cuales son galaxias satélites a nuestra galaxia. La masa de estos objetos es al menos 10 veces menor respecto a la masa de la VL.

Entonces, para asegurarnos que el grupo denso no sea de este tipo de sistemas, seleccionamos los grupos que cumplen con las siguientes condiciones:

- $\frac{M_1^*}{M_2^*} \leq 10$
- $M_2^* \geq 1 \times 10^9 M_{\odot} h^{-1}$

donde  $M_1^*$  es la masa estelar de la galaxia más masiva en el grupo y  $M_2^*$  es la masa estelar de la segunda galaxia más masiva en el grupo.

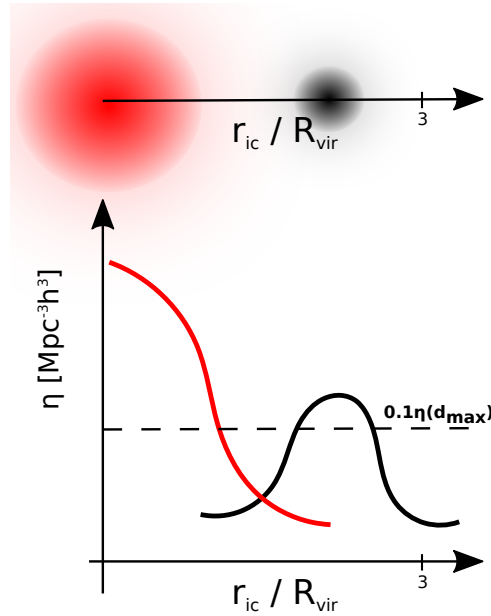
La primera condición evita que las dos galaxias con mayor masa estelar del sistema sean muy diferentes comparando sus masas estelares, es decir, no permite que sistemas como la VL y las Nubes de Magallanes sean considerados densos (y posiblemente compactos si pasan el último criterio de aislamiento); y la segunda condición la hemos considerado para evitar casos donde los sistemas cumplen con el primer ítem pero todos los miembros del grupo son galaxias muy pequeñas y poco masivas, los cuales tampoco nos interesan ya que observacionalmente no representan a los grupos compactos conocidos.

### c) Perfil de densidad

Finalmente, nos interesa que los grupos densos que han sobrevivido a los criterios anteriores puedan considerarse que están aislados localmente. Para ello utilizamos el perfil de densidad numérica de sus miembros y de las galaxias a su alrededor. El criterio adoptado para el aislamiento local califica que un grupo es aislado dentro del rango de distancias desde el centro del sistema entre la distancia a la que se encuentra el miembro más alejado,  $d_{\max}$ , y tres veces el radio virial del grupo,  $3R_{\text{vir}}$ , si el perfil de densidad de las galaxias no miembros es menor que un cierto umbral de densidad. Es decir,

$$\eta_{nm}(r) \leq 0,1 \times \eta_m(d_{\max}) \quad \text{con } r \in [d_{\max}, 3R_{\text{vir}}]$$

donde  $\eta_{nm}(r)$  es la densidad numérica de las galaxias no miembros y  $\eta_m$  es la densidad numérica de las galaxias miembros.



**Figura 6.2:** Ejemplo de cómo el perfil de densidad numérica fue utilizado para analizar aislamiento. En este ejemplo, el grupo 3D representado por el perfil rojo no se encuentra aislado de acuerdo con el criterio adoptado.

En la figura 6.2 se muestra de manera ilustrativa cómo funciona el criterio. La esfera roja hace referencia a un grupo denso y la esfera negra a un conjunto de galaxias que se encuentran a menos de 3 radios viriales del centro del grupo. En el gráfico se puede observar que el perfil de la esfera negra se encuentra por encima de la línea a trazos horizontal, que es el umbral de densidad permitido. Por lo tanto, para alguna distancia  $d_{\max} < r < 3R_{\text{vir}}$ , el perfil negro supera el umbral de aislamiento ( $0,1 \times \eta_{d_{\max}}$ ), lo que indica que el grupo rojo no satisface el criterio, y por lo tanto no está aislado localmente.

Este último criterio permite seleccionar grupos que se encuentran en una zona de muy baja densidad, aunque permite que existan algunas pocas galaxias (sin importar su masa ni su magnitud) dentro de la cáscara de aislamiento (desde  $d_{\max}$  hasta  $3R_{\text{vir}}$ ), siempre y cuando su densidad no supere el umbral de aislamiento.

### Criterio 3-D de identificación de grupos compactos

De esta manera, construimos un criterio 3-D de identificación de grupos compactos manteniendo las características usuales de los grupos compactos observacionales (población, compacticidad y aislamiento):

**Población:** Número de miembros  $N \geq 4$

**Compacticidad:** Grupos identificados con una sobredensidad de contraste  $> 5000$ , cuya separación típica es  $\sim 140$  kpc.

**Aislamiento:** Grupos densos que no son subestructura de un grupo difuso, no son sistemas formados por una galaxia y sus satélites, y están aislados localmente.

Los grupos que cumplen con estos criterios serán nombrados como “grupos compactos 3D”.

## 6.2. CGs en las simulaciones en espacio real

Utilizando el criterio desarrollado en la sección anterior, identificamos grupos compactos en las simulaciones Millennium I y II combinadas con varios SAMs. En particular, en este capítulo trabajamos con cuatro muestras de galaxias:  $G11^b$ ,  $G13^b$ ,  $HrI^b$  y  $GII^b$ . Hemos decidido utilizar sólo estos cuatro SAMs teniendo en cuenta los resultados obtenidos en el capítulo 5. Guo et al. (2011) introduce varias modificaciones con respecto a el SAM DLB realizando cambios en el tratamiento de la acreción del gas, retroalimentación, por supernovas, tamaños de galaxias y pérdida de gas y estrellas. La implementación de estas modificaciones brinda un modelo que reproduce muy bien las abundancias observadas y el clustering en la estructura en gran escala de galaxias a baja redshift como función de la masa estelar, luminosidad y color, y reproduce la distribución de color y el clustering a pequeña escala de galaxias del SDSS. Dado entonces que el SAM de G11 es una actualización mejorada del SAM de DLB, y en base a los análisis realizados hasta aquí, el SAM G11 parecería ser más eficiente en producir sistemas físicamente densos (*Reals*, ver figura 5.5). El SAM SAG no recupera de buena manera el clustering en pequeñas escalas (ver figura 2.7), lo que impacta directamente en la baja detección de CGs, de los cuales solo el 16% de los CGs identificados son *Reals*. Por último el SAM HrII no recupera de buena manera la función de luminosidad para magnitudes brillantes (ver figura 2.4) ni la función de masa estelar (ver figura 2.5) para masas altas, lo que también perjudica la búsqueda de grupos compactos.

Empezando con la identificación de grupos densos, las longitudes de percolación (en unidades de  $\text{kpc } h^{-1}$ ) utilizadas para identificar grupos densos con contraste de densidad mayor que 5000 fueron: 137, 134, 141 y 130 para G11, G13, HrI y GII, respectivamente; mientras que las longitudes (y sobredensidades) para identificar grupos difusos fueron: 308(433), 306 (423), 320(407) y 293 (433) para G11, G13, HrI y GII, respectivamente. Vale notar que las pequeñas diferencias en el contraste de densidad de grupos difusos se debe a las variaciones de los parámetros cosmológicos utilizados en cada simulación. En la tabla 6.1

describimos el número de grupos identificados en cada uno de los cubos de los SAMs en cada una de las etapas de definición: grupos densos, que no son subestructura ni tampoco una galaxia con sus satélites y que están aislados localmente.

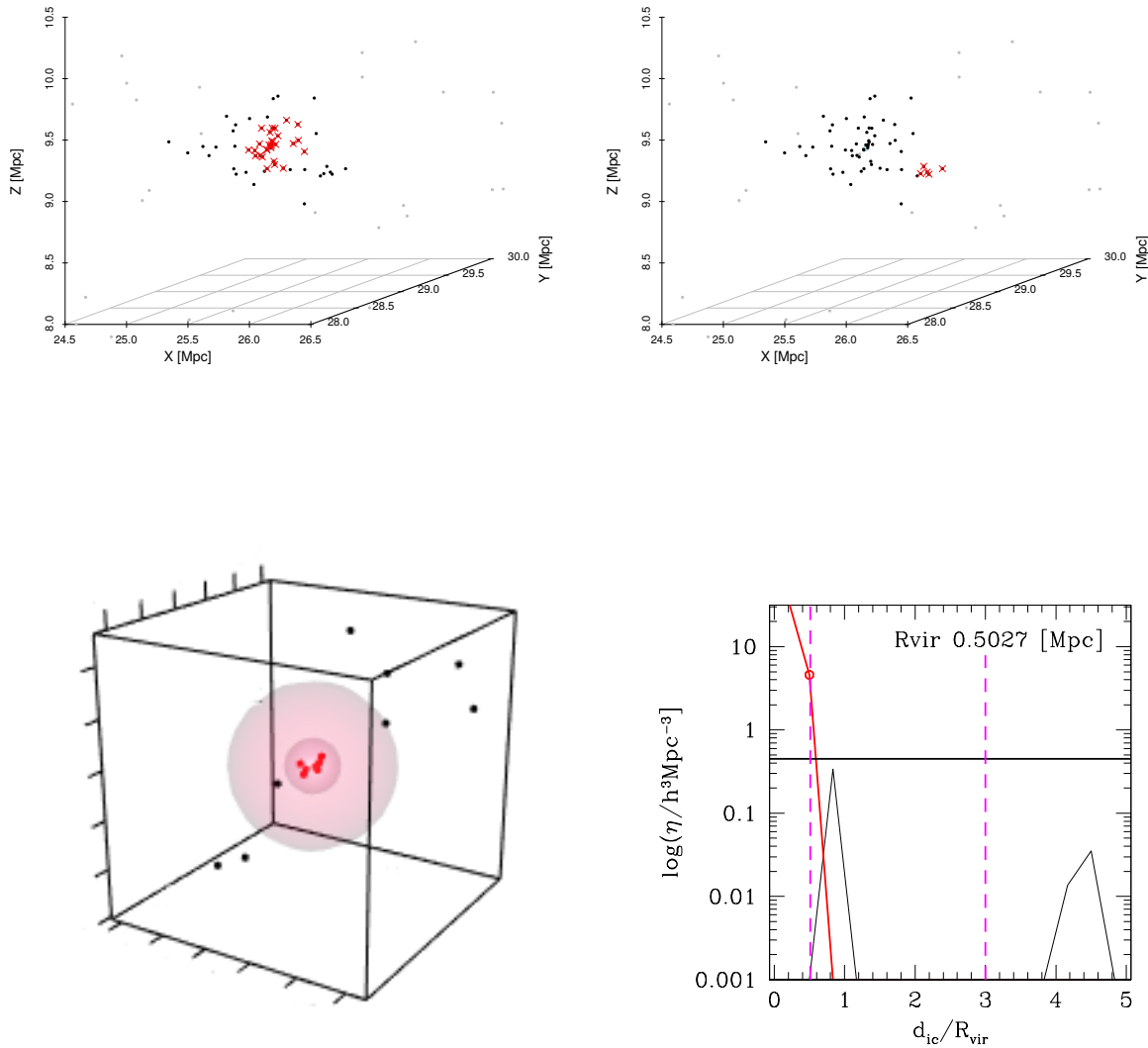
	G11 <sup>b</sup>	G13 <sup>b</sup>	HrI <sup>b</sup>	GII <sup>b</sup>
1 Densos	178599	190848	92331	1737
2 Difusos	263462	291556	208664	2436
3 $1 \cap 2$	32129 (12%)	35297 (12%)	14743 (7%)	328 (13%)
4 Subestr.	22286 (12%)	23122 (12%)	11357 (12%)	226 (13%)
5 No-subestr.	156313 (87%)	167726 (88%)	80974 (88%)	1511 (87%)
6 S/satélites	118859 (66%)	123757 (65%)	65235 (71%)	1025 (60%)
7 Aislados	47418 (27%)	49414 (26%)	24613 (27%)	421 (24%)

**Tabla 6.1:** Grupos identificados en cada SAM. Las filas 1 y 2 representan los números de grupos densos y difusos respectivamente; la fila 3 muestra el número de grupos difusos que son sobredensos, la fila 4 es el número de grupos densos que son subestructura dentro de grupos difusos; y en las últimas filas se indican el número de grupos densos que han si conservados luego de pasar cada una de los filtros del criterio de aislamiento: subestructura (fila 5), satélites (fila 6) y perfil de densidad (fila 7). Los porcentajes de las cuatro últimas filas son calculados sobre el número de grupos densos (fila 1).

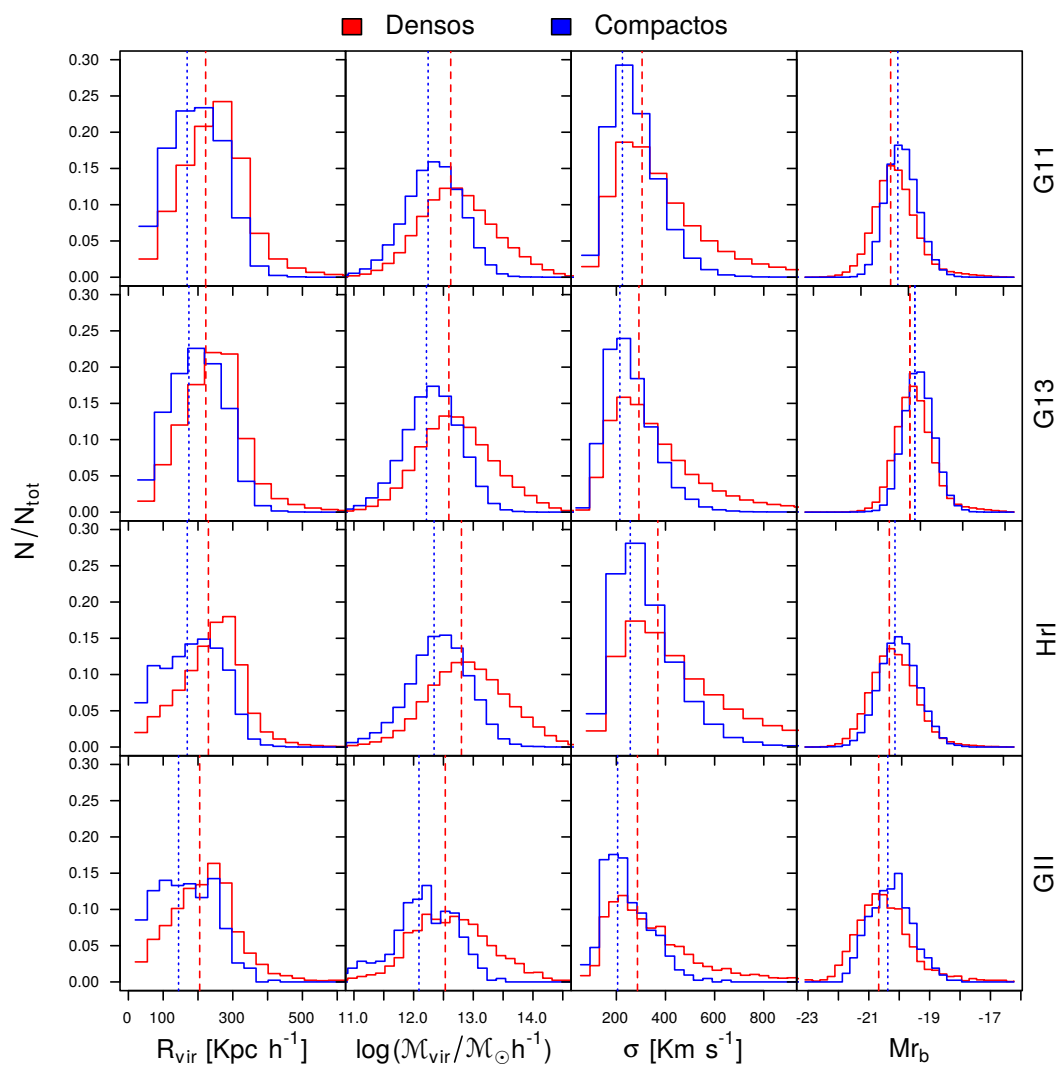
Encontramos, en casi todos los SAMs, que aproximadamente el 12% de los grupos difusos fueron identificados también como grupos densos (ver fila 3 de la tabla 6.1). El 88% restante tiene contraste de densidad más bajo que el de los grupos densos. En el SAM HrI varían levemente estos porcentajes.

Además encontramos que un muy bajo porcentaje (12%) de los grupos densos son subestructura dentro de grupos difusos (ver fila 4 de la tabla 6.1). En los paneles superiores de la figura 6.3 se observan dos ejemplos de grupos densos (puntos rojos) que fueron eliminados por ser considerados subestructura de un grupo difuso (puntos negros). En el panel superior izquierdo, el grupo denso es el núcleo de un cúmulo difuso, mientras que en el panel superior derecho, el grupo denso se encuentra alejado del centro hacia la derecha en la imagen, en lo que podría ser un grupo en caída dentro del grupo difuso.

De entre los grupos densos que no son subestructura (fila 5), eliminamos los grupos densos que son sistemas del tipo una galaxia central con sus satélites, quedándonos con aproximadamente un 65% de la muestra de grupos densos inicial (fila 6). Y finalmente, cuando filtramos con el criterio de aislamiento local con los perfiles de densidad numérica, nos quedamos con nuestra muestra final de grupos compactos 3D (fila 7). En el panel inferior izquierdo de la figura 6.3 podemos observar un ejemplo de un grupo compacto identificado a partir del criterio 3D y su correspondiente perfil de densidad en el panel inferior derecho. Podemos notar que la densidad numérica de las galaxias no miembros (perfil negro) no se encuentra por encima de la línea horizontal negra, la cual representa el umbral de aislamiento, en el intervalo entre las dos líneas verticales rosadas.



**Figura 6.3:** *Paneles superiores:* Ejemplos de grupos densos que son subestructuras de un grupo difuso. La figura izquierda muestra un grupo denso en el centro de un grupo difuso grande, y la figura derecha, un grupo denso que no es centro del grupo difuso pero es parte de uno. *Paneles inferiores:* La figura izquierda muestra un grupo compacto (aislado 3D) identificado con el criterio 3D. La figura derecha muestra el perfil de densidad numérica de galaxias en donde las distancias galaxia-centro están normalizadas al radio virial del grupo. El color rojo representa el perfil de densidad de las galaxias miembro del grupo denso, y el perfil negro hace referencia al perfil de las galaxias no miembros. Las líneas rosadas verticales y la línea negra horizontal corresponden a los límites utilizados para el criterio de aislamiento.



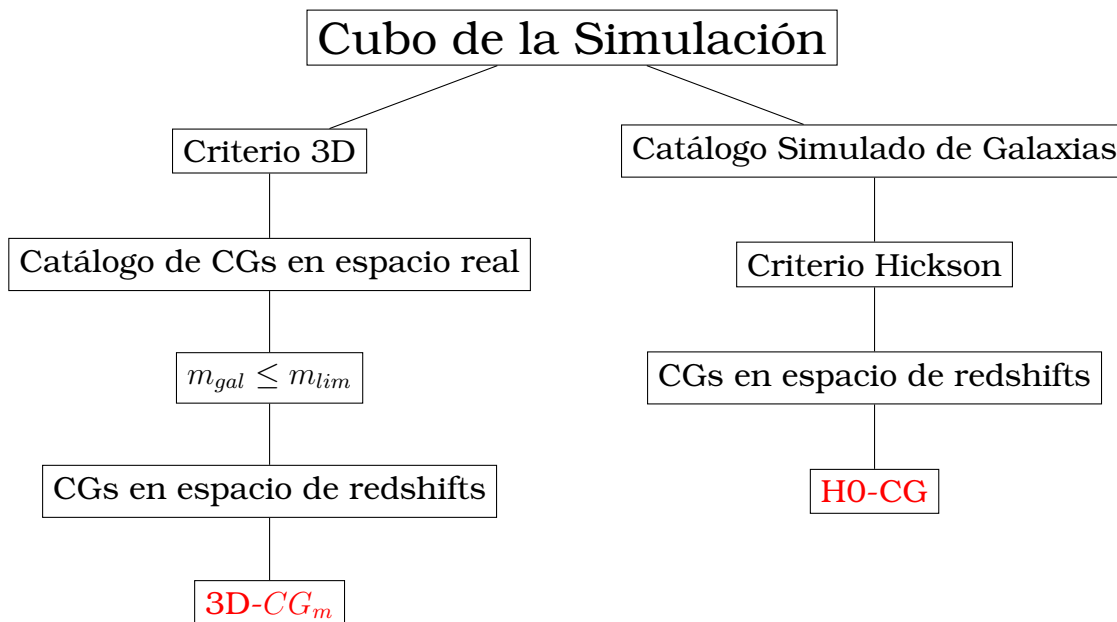
**Figura 6.4:** Distribución de propiedades de los grupos identificados. El color rojo representa los grupos densos y el color azul los grupos compactos (aislados). Las líneas verticales corresponden a la mediana de cada distribución.

La muestra final de grupos 3D que cumplen con los criterios desarrollados previamente los denominamos "grupos compactos", y es entre un 24% y un 27% de la muestra de grupos densos, dependiendo del SAM utilizado.

En la figura 6.4 mostramos las distribuciones de las distintas propiedades de los grupos densos y de los grupos compactos identificados con el criterio 3D para los diferentes SAMs. Las propiedades fueron calculadas utilizando las ecuaciones B.2, B.3 y B.6 detalladas en apéndice B. Los grupos compactos tienen tamaños más pequeños, masa más chica, dispersión de velocidades más baja y su galaxia más brillante tiende a ser más débil que la de los grupos densos en general (las medianas son estadísticamente diferentes utilizando el Intervalo de Confianza del 95% para distinguirlas).

### 6.3. ¿Puede el algoritmo de Hickson recuperar los CGs densos en 3D?

El trabajo realizado en la sección anterior nos permitió definir una muestra de grupos compactos identificados en el espacio real 3D. Esta muestra es de gran importancia porque constituye la muestra ideal de sistemas densos que quisiéramos ser capaces de identificar en un catálogo observacional. Por lo tanto, en estas próximas secciones analizaremos la capacidad del conocido identificador de grupos compactos observacionales (criterio de Hickson), para recuperar la muestra ideal de grupos compactos definida en la sección anterior.



**Figura 6.5:** Diagrama de construcción de dos muestras de CGs a partir de el cubo de una Simulación utilizando diferentes criterios de identificación.



### 6.3.1. CGs 3D en el espacio de redshifts

La comparación entre los CGs ideales y los observacionales la llevaremos a cabo en un catálogo sintético que imita las limitaciones observacionales del catálogo SDSS pero sin máscara angular. La muestra de comparación considerada ideal es construida a partir del catálogo 3D de CGs definido en la sección 6.2. A partir de dicha muestra identificada en el cubo de la simulación hemos construido una muestra de CGs en un super cubo, es decir, adosando al cubo original replicas del mismo cubo que nos permiten observar estos grupos hasta el redshift máximo del catálogo SDSS ( $\sim 0,3$ ), en un área de  $\pi/2$  sr. Luego, para generar nuestra muestra limitada en flujo, seleccionamos las galaxias miembros de los 3D-CGs que satisfacen que su magnitud aparente en la banda  $r$  sea menor a  $m_{lim}$ , donde  $m_{lim} = 17,77$  es la magnitud aparente límite del Sloan. Para más detalles de la construcción del supercubo leer la sección 2.3.

Finalmente, en una primera etapa, hemos seleccionado sólo los CGs que tienen 4 o más miembros (3D-CG<sub>m</sub>) luego del corte en flujo. Este procedimiento está esquematizado en el lado izquierdo del diagrama de la figura 6.5. De esta manera, hemos construido la muestra de CGs 3D en el espacio de redshifts. Ésta es la muestra que consideraremos ideal para su comparación con una identificación observacional de CGs.

3D-CG	G11	G13	HrI	GII
Cubo	47418	49414	24613	421
Supercubo	379344	395312	196904	421000
$n_{mi} \geq 1$	13311	12310	7162	14607
$n_{mi} = 2$ (pares)	2328	2211	1439	2565
$n_{mi} = 3$ (tripletes)	524	497	387	526
$n_{mi} \geq 4$	<b>280</b>	<b>233</b>	<b>156</b>	<b>243</b>

**Tabla 6.2:** Número de grupos compactos 3D en el supercubo luego de cortar la muestra por magnitud aparente límite.  $n_{mi}$  es el número de miembros que ha conservado el grupo luego del corte en flujo.

En la tabla 6.2 se muestran el número de 3D-CG que sobrevivieron luego de imponer el corte en magnitud aparente límite a sus miembros. Hemos incluido en esta tabla también el número de grupos que tiene menos de 4 miembros luego del límite en flujo, que serán utilizados más adelante. Los números obtenidos con los cuatro SAMs son similares, salvo para el SAM HrI donde los valores obtenidos son un poco menores. Puede verse que del total de grupos que sobrevivieron al límite en flujo con al menos una galaxia, aproximadamente el 2% poseen al menos cuatro miembros con  $r < 17,77$ , mientras que cerca de un 4% son tripletes.

En la figura 6.6 mostramos la distribución de dos propiedades observacionales: la mediana de la separación proyectada entre galaxias y el brillo superficial. En esta figura mostramos la comparación de estas propiedades en los 3D-CG<sub>m</sub> en

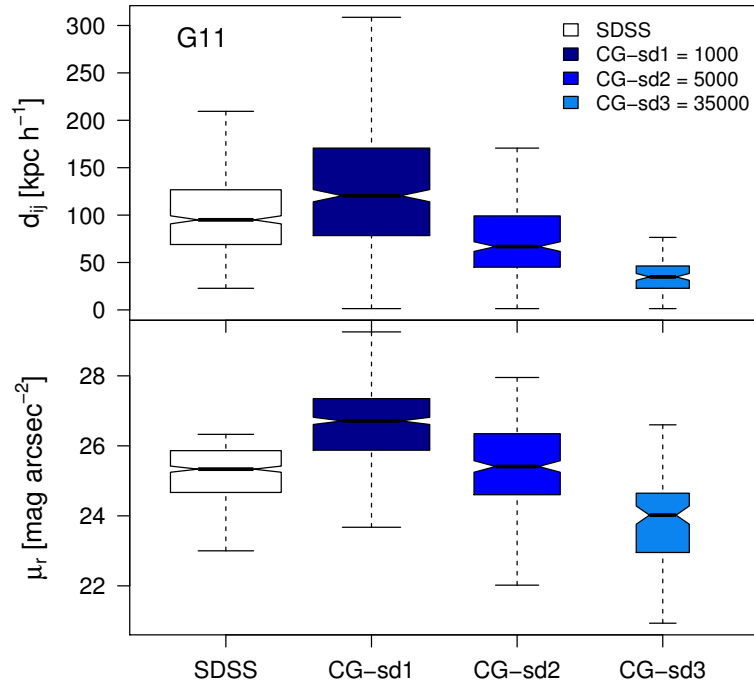
el espacio de redshifts con los resultados obtenidos para una muestra observacional de grupos compactos. La muestra observacional son los CGs identificados en el SDSS previamente mostrados en el capítulo 4, identificados utilizando el algoritmo de Hickson modificado. El objetivo de esta comparación es verificar que la sobredensidad elegida para identificar los 3D-CG, genera muestras de CGs en el espacio de redshifts con propiedades similares a los grupos observacionales. Además, hemos agregado a la comparación dos muestras con dos sobredensidades diferentes (1000 - CG-sd1 - y 35000 - CG-sd3) para realizar la identificación 3D y poder comparar con lo obtenido para nuestra muestra 3D-CG en el espacio de redshifts (llamada CG-sd2 en este gráfico). Los diagramas de caja de la figura 6.6 muestran las comparaciones entre las 3 muestras de grupos identificados con diferentes sobredensidades: CG-sd1, CG-sd2, CG-sd3 y la muestra de CGs identificada en el SDSS. Se puede ver que la muestra CG-sd2 es la que mejor recupera las distribuciones de  $d_{ij}$  y  $\mu$  observadas para la muestra de CGs en el SDSS. Identificar grupos con sobredensidad muy alta (CG-sd3) genera muestras con muy bajo brillo superficial (alta compacticidad), es decir, con separaciones proyectadas intergalaxia demasiado pequeñas. Por otro lado, si identificamos grupos con sobredensidad más baja (CG-sd1), las separaciones proyectadas son más elevadas. La elección de 5000 como contraste de densidad para la identificación de grupos densos se ve justificada empíricamente a través de este gráfico.

### 6.3.2. Identificación de CGs tipo Hickson en el espacio de redshifts

Con el objetivo de comparar nuestros 3D-CG<sub>m</sub> con grupos compactos identificados a partir del criterio de Hickson, hemos construido un catálogo sintético de galaxias sin evolución, limitado por flujo de manera que alcanza el redshift máximo del SDSS, en un ángulo sólido de  $\pi/2$  sr, para cada SAM con el método desarrollado en el capítulo 2.3. En la figura 6.7 se muestran las distribuciones de redshifts de las galaxias en cada catálogo sintético de cada SAM, comparadas con la distribución de las galaxias en el SDSS. Puede verse que todos los catálogos reproducen bastante bien la distribución de redshifts hasta  $z \sim 0,13$ , y en todos los SAMs se ve un defecto de galaxias para  $z$  más altos.

A partir de estos catálogos sintéticos de galaxias, hemos identificado CGs utilizando el algoritmo de Hickson modificado, desarrollado en el capítulo 4. Básicamente, se utiliza el criterio original de Hickson en el espacio de redshifts pero aplicado de manera automática e incorporando el filtro en velocidad al mismo tiempo que los otros criterios, lo que mejora la completitud de las muestras de grupos identificados. Los números de grupos compactos tipo Hickson (H-0) obtenidos para cada SAM son los siguientes: 478, 289, 188 y 521 CGs en los SAMs G11, G13, HrI y GII, respectivamente.

Nos interesa saber si las propiedades de los grupos identificados de manera

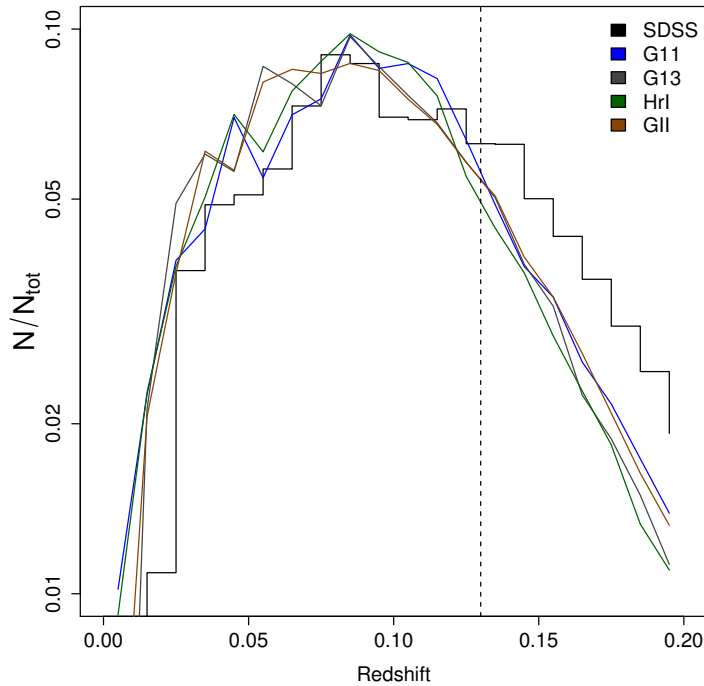


**Figura 6.6:** Distribuciones de la mediana de la separación proyectada entre galaxias ( $d_{ij}$  y el brillo superficial de los CG-3D $_m$  para el SAM G11 obtenidos a partir de un contraste de densidad inicial de identificación mayor a 1000 (CG-sd1), a 5000 (CG-sd2) y a 35000 (CG-sd3). Estas muestras son comparadas con las distribuciones de los HMCg del catálogo observacional SDSS que fueron identificados en el capítulo 4.

observacional son similares a las propiedades de los grupos ideales identificados en 3D. Pero, debido a que el criterio de Hickson considera como miembros únicamente a las galaxias en un rango de 3 magnitudes a partir de su galaxia brillante, para hacer una comparación más justa, hemos incorporado una nueva muestra de grupos 3D en el espacio de redshifts en los que se consideran sólo sus miembros en un rango de 3 magnitudes desde la galaxia brillante. Todas las propiedades de los grupos fueron calculadas usando las ecuaciones detalladas en el apéndice B.

En la figura 6.8 se muestran los diagramas de caja de las distribuciones de las propiedades de los grupos en todos los SAMs analizados. Se comparan las propiedades de los grupos 3D trasladados al espacio de redshifts y más de cuatro miembros después del límite en flujo (3D-CG $_m$ ), los grupos 3D trasladados al espacio de redshifts con límite en flujo y considerando sólo sus miembros en un rango de 3 magnitudes (m3), y los grupos compactos tipo Hickson identificados directamente en el espacio de redshifts (H-0).

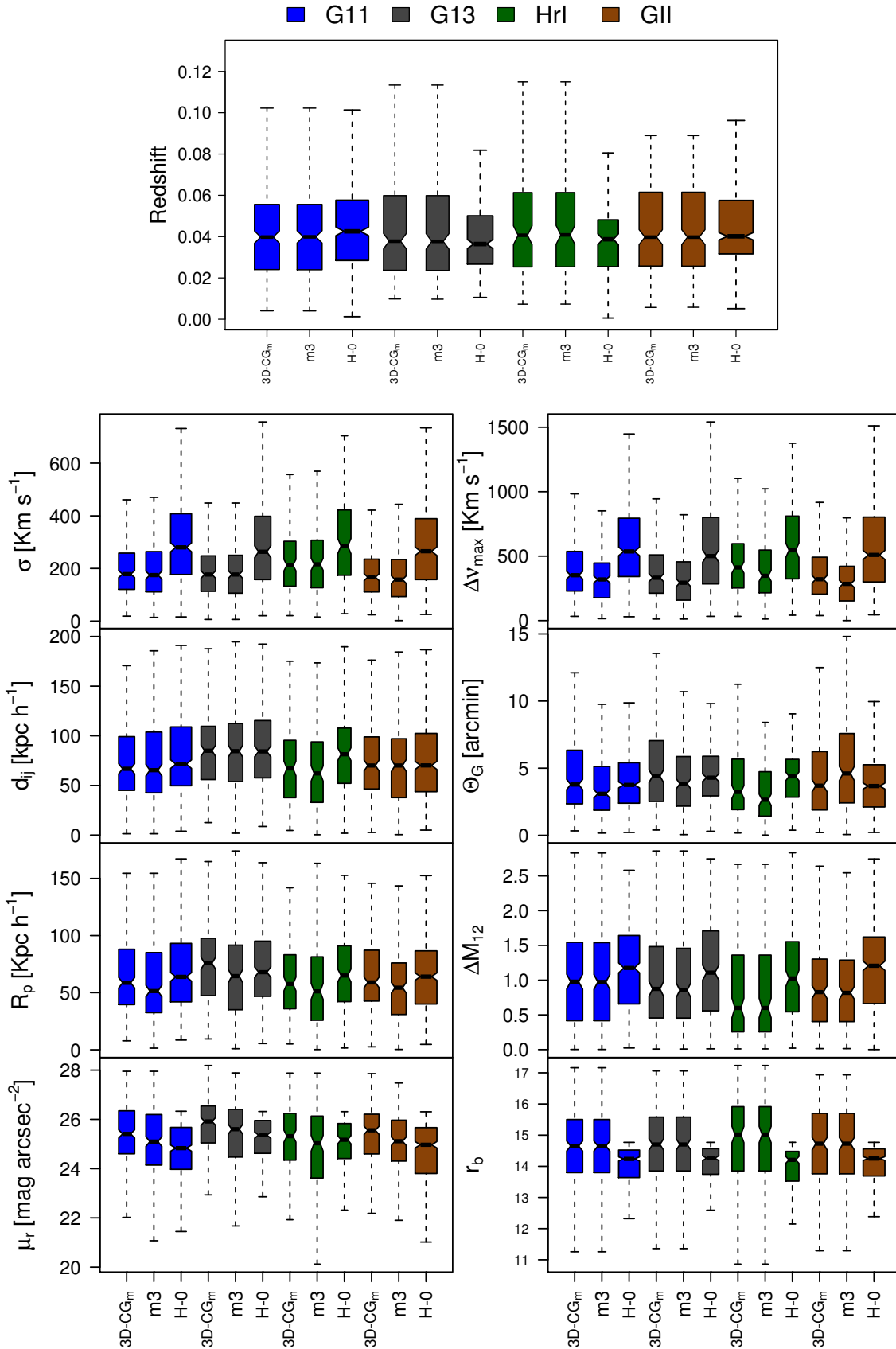
En primer lugar, se puede ver que no hay cambios significativos entre las mues-



**Figura 6.7:** Distribución de redshifts de las galaxias en los catálogos sintéticos sin evolución para cada SAM comparadas con la distribución de galaxias en el SDSS.

tras  $3D-CG_m$  y  $m3$ . Las propiedades que se ven levemente afectadas son las que tienen que ver con el número de miembros como el diámetro angular del círculo mínimo ( $\Theta_G$ ), el brillo superficial ( $\mu_r$ ), el radio proyectado ( $R_p$ ), y la diferencia máxima en velocidad de las galaxias miembro ( $\Delta v_{max}$ ).

Por otro lado, en la figura 6.8 se puede ver que en general, para todas los SAMs, los H-0 son más grandes en tamaños ( $d_{ij}$ ,  $\Theta_G$  y  $R_p$ ), tienen dispersión de velocidades más alta ( $\sigma$ ), brillo superficial ( $\mu_r$ ) menor (mayor compactidad), diferencia en velocidades entre sus miembros ( $\Delta v_{max}$ ) considerablemente más alta, diferencia en magnitudes ( $\Delta M_{12}$ ) más alta y magnitud aparente de la galaxia brillante ( $r_i$ ) más brillante que la de los grupos 3D. Estas dos últimas propiedades están directamente relacionadas con el criterio utilizado. Sin embargo, las distribuciones de redshifts de estos grupos son similares, lo cual elimina la posibilidad de que algunas de estas diferencias sean debido a que los grupos se encuentren a diferentes redshift. Debido a estas diferencias, los grupos compactos 3D y los H-0 parecen ser poblaciones de grupos diferentes.



**Figura 6.8:** Diagramas de caja de propiedades observacionales en el espacio de redshifts de grupos compactos. Para cada SAM se comparan las propiedades de los grupos 3D-CG<sub>m</sub>, los grupos m3 y los grupos H-0.

### 6.3.3. Pureza y Completitud de CGs

Existen dos propiedades que permiten caracterizar una muestra de grupos: la completitud y la pureza. La completitud nos habla de la bondad del identificador, es decir, si podemos identificar el total de los grupos ideales existentes. Mientras que la pureza nos dice si los grupos que se han identificado, son los ideales que se querían identificar. Si la completitud es un porcentaje alto, significa que mi muestra identificada contiene a la mayoría de los grupos que se querían identificar, pero al mismo tiempo la muestra puede estar contaminada con otros tipos grupos espurios dando un bajo porcentaje de pureza (muestra completa e impura). Por otro lado, si el porcentaje de pureza es alto, significa que muy pocos de los grupos identificados son espurios, pero pueden existir muchos otros grupos ideales que no han sido identificados, dando una baja completitud (muestra pura e incompleta). Por lo tanto, se define a una muestra “confiable” si tiene un alto porcentaje de pureza y de completitud al mismo tiempo.

Si asumimos que nuestra muestra limitada en flujo de grupos compactos 3D trasladados al espacio de redshifts,  $3D-CG_m$ , es la muestra ideal que queremos recuperar, podemos analizar la pureza y la completitud de la muestra de CGs H-0. Para eso, definimos la pureza  $P$  y la completitud  $C$  de la siguiente manera. Sean  $N_{rec}$  el número de grupos que pertenecen tanto a la muestra ideal como a la muestra tipo Hickson,  $N_{3D}$  es el número total de grupos ideales,  $3D-CG_m$  y  $N_H$  es el número de CGs H-0, entonces,

- **Pureza:** es el porcentaje de grupos H-0 que son grupos ideales  $3D-CG_m$

$$P = 100 \times \frac{N_{rec}}{N_H}$$

- **Completitud:** es el porcentaje de grupos ideales  $3d-CG_m$  que fueron recuperados por el identificador tipo Hickson

$$C = 100 \times \frac{N_{rec}}{N_{3D}}$$

Para calcular el número de grupos recuperados realizamos una búsqueda de coincidencias espacial: en proyección y en la línea de la visual. Hemos seleccionado los grupos H-0 y  $3D-CG_m$  cuyos centros se encuentran angularmente cerca y además la diferencia de los redshifts entre sus centros es relativamente baja. Tomamos como coincidentes aquellos grupos que satisfacen:

- Distancia angular entre los centros del círculo mínimo es menor a la suma de los radios angulares del círculo mínimo:  $\gamma_{cc} \leq \theta_G(3D-CG) + \theta_G(CG H-0)$
- Distancia a lo largo de la línea de la visual entre los centros menor a  $1800 \text{ km s}^{-1}$ :  $c|z(3D-CG) - z(CG H-0)| \leq 1800 \text{ km s}^{-1}$

		G11	G13	HrI	GII	
		3D-CG <sub>m</sub>	280	233	156	243
H-0	$\Delta v = 1000$ c/n3+rb $\mu=26.33$	# CG H-0	478	289	188	521
		$N_{\text{rec}}$	67	40	31	60
		P	14 ± 3	14 ± 4	16 ± 5	12 ± 3
		C	24 ± 5	17 ± 5	20 ± 6	25 ± 5
H-1	$\Delta v = 500$ c/n3+rb $\mu=26.33$	# CG H-1	429	254	154	469
		$N_{\text{rec}}$	64	39	28	58
		P	15 ± 3	15 ± 5	18 ± 6	12 ± 3
		C	23 ± 5	17 ± 5	18 ± 6	24 ± 5
H-2	$\Delta v = 500$ c/n3+rb $\mu=27.0$	# CG H-2	558	362	242	565
		$N_{\text{rec}}$	78	50	30	62
		P	14 ± 3	14 ± 4	12 ± 4	11 ± 3
		C	28 ± 5	21 ± 5	19 ± 6	26 ± 6
H-3	$\Delta v = 500$ s/n3-rb $\mu=26.33$	# CG H-3	1669	1045	775	2009
		$N_{\text{rec}}$	176	134	87	169
		P	11 ± 2	13 ± 2	11 ± 2	8 ± 1
		C	63 ± 6	58 ± 6	56 ± 8	70 ± 6
H-4	$\Delta v = 500$ s/n3-rb $\mu=27.0$	# CG H-4	2398	1630	1223	2705
		$N_{\text{rec}}$	197	170	97	189
		P	8 ± 1	10 ± 2	8 ± 2	7 ± 1
		C	70 ± 5	73 ± 6	62 ± 8	78 ± 5
H-5	Submst H-4 c/ $z_{cg} \leq 0,1$ $Mr_b \leq -22$	# CG H-5	1527	1295	937	1919
		$N_{\text{rec}}$	192	158	92	189
		P	13 ± 2	12 ± 2	10 ± 2	10 ± 1
		C	69 ± 5	68 ± 6	59 ± 8	78 ± 5

**Tabla 6.3:** Pureza y Completitud para distintas muestras tipo Hickson identificadas con criterios levemente diferentes (los parámetros principales de cada criterio están en la 2da columna). Los intervalos de confianza binomial del 95% para la mediana fueron calculados por:  $IC_C = 1,96\sqrt{\frac{C(100-C)}{N_{3D}}}$ ,  $IC_P = 1,96\sqrt{\frac{P(100-P)}{N_H}}$

La elección, algo permisiva, de 1800 km/s a lo largo de la línea de la visual fue inspirada en el hecho de que la identificación tipo Hickson permite que la separación máxima entre galaxias sea de hasta 2000 km/s. En esta etapa, no se tomaron en cuenta los miembros de cada grupo ya que debido a las diferencias en la identificación, es probable que estos sean diferentes. Por lo tanto, los grupos recuperados puede que no coincidan exactamente en población, pero, de hecho, la diferencia en el número de miembros de los recuperados en todas las muestras es muy baja (entre 1 y 5 en su mayoría).

Teniendo el número de grupos recuperados en cada uno de los SAMs, hemos calculado los porcentajes de pureza y completitud de la muestra H-0. Los mismos se encuentran detallados en las primeras filas de la tabla 6.3.

Tanto la pureza como la completitud de la muestra H-0 son muy bajas para todos los SAMs, cercana al 14% para la pureza y alrededor del 20% para la completitud. Esto significa que la muestra H-0 está extremadamente contaminada por grupos espurios y que además existen muchos grupos compactos ideales que no se lograron identificar. Por ejemplo, del total de grupos H-0: 478 identificados en el SAM G11, sólo 67 coinciden con grupos 3D, lo que significa que el resto de los grupos de la muestra H-0 (411) son grupos que contaminan la muestra (baja pureza). Los grupos 3D que no fueron identificados por la muestra Hickson son un total de 213 de los 280 que tiene la muestra 3D-CG<sub>m</sub>, lo que genera un bajo porcentaje de la completitud de la muestra.

### **Mejorando el criterio tipo Hickson**

Para intentar mejorar la completitud y/o pureza de la muestra observacional, hemos realizado modificaciones al criterio H-0 pero manteniendo la idea original sólo modificando los umbrales de los parámetros de identificación. Es decir, hemos variado el umbral máximo de velocidades concordantes  $\Delta v$ , el brillo superficial límite,  $\mu_r$ , y los cortes en magnitudes que tienen que ver con la membresía,  $n_3$ , y el límite de magnitud aparente de la galaxia brillante del grupo,  $r_b$ . Las elecciones de los nuevos valores de los parámetros fueron realizadas analizando las distribuciones y valores mínimos y máximos de las propiedades de los grupos 3D-CG<sub>m</sub> y H-0 de la figura 6.8.

En la tabla 6.4 se muestran todas las combinaciones de parámetros que utilizamos para intentar mejorar las muestras.

Para construir la muestra H-1, el único parámetro que modificamos respecto del criterio Hickson modificado original que crea la muestra H-0, fue la máxima separación a lo largo de la línea de la visual entre las galaxias y su centro, disminuyendo a la mitad respecto del valor original, y adoptando un valor límite  $\Delta v = 500$ . Esta modificación no introduce grandes variaciones en cuanto a pureza o completitud.



Muestra	Miembros	Aislación	Compacticidad	Filtro velocidad	$n_{3+r_b}$
H-0	$4 \leq N \leq 10$	$\theta_N > 3\theta_G$	$\mu_r \leq 26,33$	$\leq 1000 \text{ km/s}$	si
H-1	$4 \leq N \leq 10$	$\theta_N > 3\theta_G$	$\mu_r \leq 26,33$	$\leq 500 \text{ km/s}$	si
H-2	$4 \leq N \leq 10$	$\theta_N > 3\theta_G$	$\mu_r \leq 27,0$	$\leq 500 \text{ km/s}$	si
H-3	$N \geq 4$	$\theta_N > 3\theta_G$	$\mu_r \leq 26,33$	$\leq 500 \text{ km/s}$	libre
H-4	$N \geq 4$	$\theta_N > 3\theta_G$	$\mu_r \leq 27,0$	$\leq 500 \text{ km/s}$	libre

**Tabla 6.4:** Parámetros utilizados para definir las diferentes muestras identificadas con los criterios tipo Hickson.

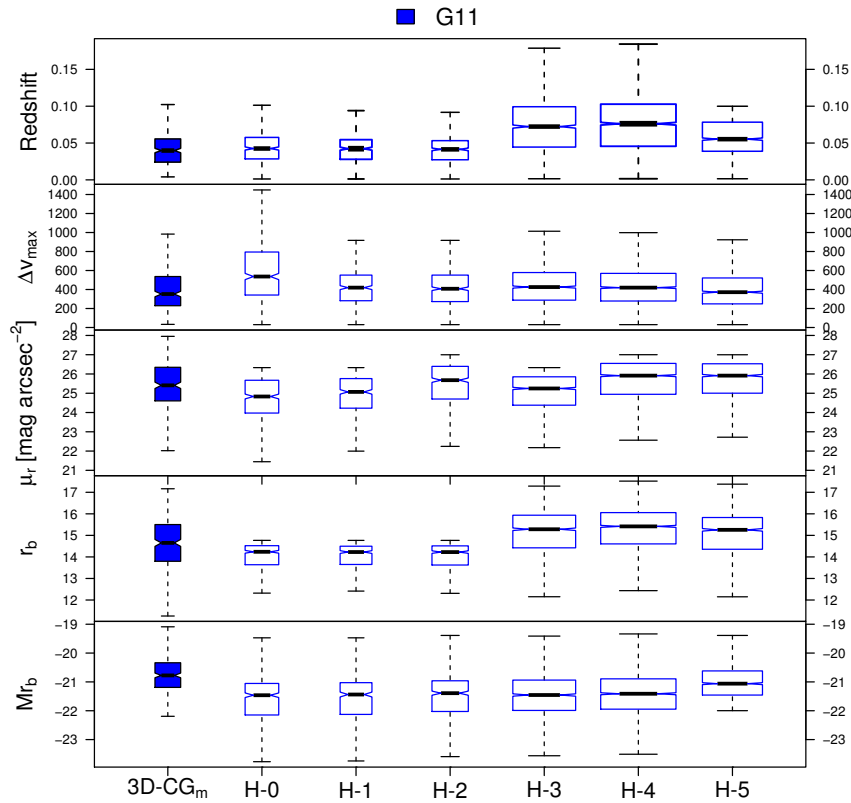
En segundo lugar, debido a que la mediana del brillo superficial para la muestra 3D-CG<sub>m</sub> es más alta que para H-0, hemos aumentado el valor de  $\mu_{lim}$  de 26.33 a 27 [mag arcsec<sup>-2</sup>], y hemos mantenido el cambio en velocidad de la muestra anterior. Así construimos la muestra H-2. La completitud aumentó para todos los SAMs, pero la pureza se mantuvo.

Por último, dado que el algoritmo de Hickson identifica sus miembros, calcula la compacticidad, y controla el aislamiento, todo considerando sólo las galaxias en un rango de 3 magnitudes a partir de la magnitud de la galaxia más brillante, hemos optado por dejar libre el rango de magnitudes y permitirle al algoritmo encontrar grupos con más miembros débiles, y aislados de galaxias que tengan cualquier magnitud aparente. Esto lo hicimos para dos casos particulares de brillo superficial:  $\mu_{lim} = 26,33$  (muestra H-3) y  $\mu_{lim} = 27$  (muestra H-4). Estos últimos dos criterios lograron aumentar de forma significativa la completitud: cerca de un 60% para la muestra H-3 y cerca de un 70% para la muestra H-4, dependiendo del SAM; Sin embargo, estas modificaciones actúan muy pobremente en relación a la pureza, que decae hasta valores cercanos al 10%, para la mayoría de los SAMs.

Entonces, las pequeñas modificaciones introducidas en el criterio de identificación, no contribuyeron a mejorar la pureza de las muestras. Si tomamos un brillo superficial un poco más alto (H-2), la completitud de la muestra aumenta levemente sin cambios en la pureza. Cuando tomamos una muestra con brillo superficial alto y sin restricciones en magnitudes (H-4), la completitud aumenta considerablemente, esto quiere decir que mi nueva muestra H-4 recupera la gran parte de los 3D-CG<sub>m</sub> (197 de 283 para G11), pero esa muestra está extremadamente contaminada (existen 2201 grupos que son espurios).

Podemos comparar ahora las propiedades de los grupos en cada una de estas muestras modificadas con respecto a la muestra de CG-3D<sub>m</sub>. En la figura 6.9 se muestran las distribuciones de algunas de las propiedades relacionadas con los parámetros que utiliza el criterio de Hickson y los cuales hemos ido variando para obtener las diferentes muestras. Ésta figura fue realizada sólo para el SAM G11, pero los demás tienen comportamiento similar.

En primer lugar, se puede ver que en las muestras H-3 y H-4, las distribuciones de redshifts no son similares a la muestra 3D-CG<sub>m</sub>, sino que están siendo identificados hasta redshifts más altos. El haber liberado del límite en flujo de



**Figura 6.9:** Diagramas de caja de diferentes propiedades de CGs para diferentes muestras de CGs para el SAM G11.

la galaxia brillante produjo que sea posible identificar CGs hasta mayores distancias, pero esto parece ser contraproducente en términos de querer reproducir la muestra ideal. En segundo lugar, vemos que la máxima separación entre galaxias a lo largo de la línea de la visual ( $\Delta v_{\max}$ ) para la muestra H-0 es muy alta comparada con la muestra 3D-CG<sub>m</sub>, pero todas las otras muestras solucionan esa diferencia utilizando un umbral más bajo para la identificación de galaxias concordantes ( $\Delta v$ ). En tercer lugar, puede verse que el brillo superficial abarca un amplio rango de valores para la muestra 3D, pero la muestra H-0 sólo permitía  $\mu \leq 26,33$ , por lo tanto, aumentamos este valor a 27 para permitir la identificación de esos grupos que se estaban restringiendo en el criterio original. De esta manera, las muestras H-2 y H-4, recuperan de mejor manera la distribución 3D de brillo superficial. En cuarto lugar, podemos notar que la restricción en la magnitud aparente de la galaxia brillante del grupo ( $r_b$ ) es muy determinante al identificar grupos tipo Hickson. El criterio de identificación de grupos 3D permite incluir en su muestra grupos con galaxias más débiles que  $m_{lim} - 3 = 14,77$ . En las muestras H-3 y H-4, las cuales se identificaron liberando esta restricción del criterio, las galaxias brillantes son más débiles que la brillante de los 3D-CG. Y por último, hemos estudiado la distribución de la magnitud absoluta de la galaxia más brillante del grupo ( $Mr_b$ ), cantidad que en los catálogos sintéticos no depende del observador como sí lo hacen todas las

anteriores. Ninguna muestra tipo Hickson ha logrado reproducir la distribución de  $M_{r_b}$ , es decir, que los grupos 3D tienen galaxias con magnitud absoluta de la galaxia brillante más débil que la galaxia brillante identificada por los criterios tipo Hickson.

Teniendo en cuentas estas diferencias en las distribuciones entre los grupos 3D y los tipo Hickson, y además los bajos porcentajes de pureza de las muestras tipo Hickson, hemos construido una nueva muestra de CGs (H-5) a partir de la muestra H-4 para intentar mejorar la pureza y así tener una muestra más confiable de CGs. Hemos utilizado como punto de partida los grupos H-4 —ya que es la muestra con mayor completitud— y, en base al análisis de las propiedades que mayores diferencias arrojaron en la figura 6.9, hemos decidido restringir la muestra H-4 en redshift y en magnitud absoluta de la galaxia brillante, quedándonos con aquellos grupos que además cumplen con las siguientes restricciones:  $z_{cm} \leq 0,1$  y  $M_{r_{bri}} \geq -22$ . En el último diagrama de caja de la figura 6.9 se muestran las distribuciones de las propiedades para la muestra H-5, y se puede ver que las distribuciones de la muestra H-5 intentan acercarse mejor a las distribuciones de las propiedades de los 3D-CG<sub>m</sub>. La construcción de esta nueva muestra permitió aumentar la pureza manteniendo un alto porcentaje en completitud. En las últimas filas de la tabla 6.3 se detallan los valores obtenidos para la completitud y la pureza de la muestra H-5. Sin embargo, el porcentaje de pureza sólo ha aumentado un levemente respecto de las muestras H-4, llegando a valores similares a los de las muestra H-0. Es importante destacar que, a pesar de que la completitud aumentó a porcentajes deseables, la pureza no supera el 13%. Es más, con ningún criterio tipo Hickson en cualquier SAM se pudo obtener una pureza superior al 18%.

Podemos concluir que los criterios tipo Hickson no logran identificar una muestra de grupos compactos con alta completitud y pureza al mismo tiempo.

#### 6.3.4. Análisis de grupos espurios

En esta sección, nos preguntamos qué tipo de grupos son aquellos identificados utilizando algún criterio tipo Hickson, pero que no tienen una contraparte en los grupos compactos ideales 3D.

Hasta aquí, hemos considerado como ideales a los grupos 3D-CG<sub>m</sub> que fueron identificados en 3D en un supercubo y que luego quedaron con 4 o más miembros al cortar el catálogo sintético por magnitud aparente. Aquellos compactos 3D que tuvieron menos de 4 miembros más brillantes que la magnitud aparente límite, han sido descartados de la muestra de comparación. Esta restricción observacional nos lleva a descartar grupos compactos 3D que observacionalmente aparecen como tripletes, dobles o una sola galaxia, pero el hecho de que no se puedan observar sus galaxias compañeras no hace que ese grupo deje de existir.

Ahora, ¿podría pasar que el identificador tipo Hickson encuentre esos “miembros visibles” de los grupos 3D y que al mismo tiempo le agregue otras galaxias intrusas para lograr llegar a los 4 miembros requeridos por el criterio observacional? Para tener en cuenta estos casos, seremos más permisivos en cuanto a considerar la muestra ideal 3D extendiendo la muestra de grupos 3D en el espacio de redshifts ( $3D-CG_m$ ) y permitiendo considerar grupos a los que les sobrevive al menos un miembro luego de cortar por magnitud aparente ( $3D-CG_{+1}$ ). En la tabla 6.2 se puede ver la cantidad de grupos que cumplen con esta condición.

También modificamos la forma en la que buscamos grupos recuperados (grupos Hickson que viven en la muestra extendida de grupos compactos ideales en el espacio de redshifts  $3D-CG_{+1}$ ): primero buscamos coincidencias viendo si la galaxia brillante del grupo Hickson es parte de un grupo  $3D-CG_{+1}$ . Si esto ocurre, seleccionamos al grupo Hickson como recuperado. En caso contrario, si la galaxia brillante no pertenece a un grupo  $3D-CG_{+1}$ , seleccionamos al grupo Hickson como recuperado si cumple con las condiciones de cercanía en proyección angular y redshift tal como lo hicimos anteriormente con los grupos  $3D-CG_m$  con 4 o más miembros<sup>1</sup>. Los grupos restantes, es decir los grupos Hickson que no se encuentran en la muestra de  $3D-CG_{+1}$  (ni por cercanía ni por su galaxia brillante), son los grupos a los que llamaremos espurios.

En esta sección nos concentraremos sólo en la muestra H-0, por cuestiones históricas, y en la muestra H-5 que es la que mejor lograba reproducir las propiedades de los grupos  $CG-3D_m$  con una alta completitud.

En la tabla 6.5 mostramos los valores del número de nuevos grupos recuperados para las muestras H-0 y H-5. Comparando con la tabla 6.3, se puede ver que para las muestras H-0, el número de recuperados no aumenta considerablemente a pesar de que la muestra de control  $3D-CG_{+1}$  es treinta veces más grande. Esto significa que los grupos H-0 están asociados a muy pocos tripletes, pares y grupos de un miembro de los grupos ideales. La pureza aumenta levemente, ya que los recuperados que se agregaron son pocos. Por otro lado, al agrandar considerablemente la muestra de comparación 3D, la completitud disminuye drásticamente. Por ejemplo, para el SAM G11, se tenían 67 grupos recuperados en la muestra H-0, dando una pureza de 14% y pasó a 81 recuperados, con una pureza de 17%, mientras que la completitud pasó de 24% con la muestra  $3D-CG_m$  con 280 grupos, a un valor de 0,6% con la muestra  $3D-CG_{+1}$  que tiene 13311 compactos. Sin embargo, para las muestras H-5 el número de recuperados aumentó considerablemente, llevando la pureza a valores cercanos al 30% para casi todos los SAMs, un resultado mucho más optimista que el anterior. Por supuesto hay que recordar que en estos casos, estamos considerando como “puro” a grupos H-5 que tienen al menos 4 miembros pero que sabemos que en la gran mayoría de ellos menos de 4 miembros son parte del grupo ideal,

<sup>1</sup>Ya que en el criterio de cercanía angular intervienen los radios angulares de los círculos mínimos de los grupos involucrados, en el caso de que el grupo  $3D-CG_{+1}$  sólo tenga 1 miembro, se le asigna radio angular=0

y podemos decir entonces con certeza que muchos contienen miembros intrusos.

Muestras		G11	G13	HrI	GII
#	3D-CG <sub>+1</sub> ( $n_{mi} \geq 1$ )	13311	12310	7162	14607
	# CG H-0	478	289	188	521
H-0	$N_{rec}$	81	52	36	85
	P	$17 \pm 3$	$18 \pm 4$	$19 \pm 6$	$16 \pm 3$
	# CG-esp	397 (83 %)	237 (82 %)	152 (81 %)	436 (84 %)
	# CG H-5	1527	1295	937	1919
H-5	$N_{rec}$	494	381	285	479
	P	$32 \pm 2$	$29 \pm 2$	$30 \pm 3$	$25 \pm 2$
	# CG-esp	1033 (68 %)	914 (71 %)	652 (70 %)	1440(75 %)

**Tabla 6.5:** Número de grupos Hickson que están asociados a un CG-3D<sub>+1</sub> ( $N_{rec}$ ), pureza de la muestra tipo Hickson (P) y número de grupos Hickson espurios (CG-esp). El porcentaje de espurios es calculado con respecto a las muestras tipo Hickson.

Nos centraremos ahora en el análisis de los grupos espurios: una pregunta razonable que surge es si al menos estos grupos espurios que no están asociados con grupos compactos 3D-CG<sub>+1</sub> (=densos+aislados), lo están con algún grupo sobredenso.

En la sección 6.1.1 hemos identificado grupos densos en los cubos de las simulaciones en el espacio real. Dado que los grupos espurios son grupos Hickson identificados en el espacio de redshifts, utilizamos la información de las posiciones 3D que nos provee el catálogo sintético obtenida a partir de simulaciones. Con esta información puedo conocer las posiciones comóviles  $\vec{r} = (x, y, z)$  de los miembros de los grupos espurios. Es decir, que hemos trasladado la muestra de grupos espurios al espacio real.

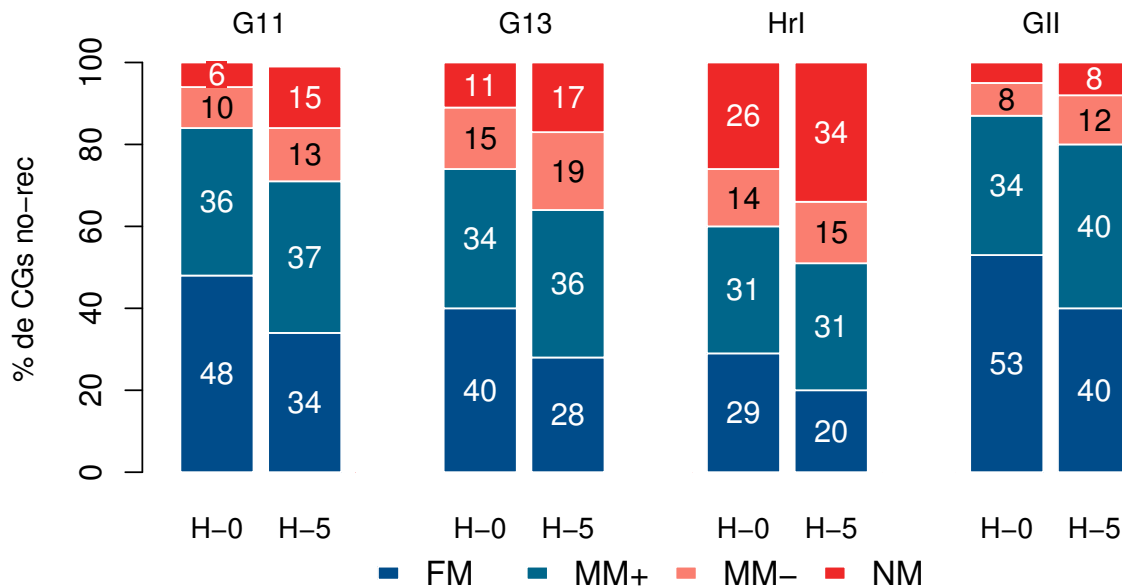
Esta muestra de grupos Hickson espurios en 3D, nos ha permitido buscar coincidencias miembro-a-miembro para estudiar si los miembros del grupo espurio coinciden con los miembros de algún grupo sobredenso. Si ninguno coincide, significa que estos grupos espurios no tienen una contraparte en un grupo denso. Si coinciden algunos miembros entonces significa que el grupo espurio sí tiene asociado un grupo denso. Dividiremos a los grupos espurios en cuatro categorías dependiendo del número de miembros coincidentes con los miembros de los grupos densos:

**FM** (Full Match): todos los miembros del grupo espurio coinciden con los miembros de algún grupo denso

**MM+** (Medium Match +): más del 50% de los miembros del grupo espurio coinciden con los miembros de algún grupo denso

**MM-** (Medium Match -): menos de la mitad de los miembros del grupo espurio coinciden con los miembros de un grupo denso

**NM** (*No Match*): no coincide ninguno de los miembros del grupo espurio con ningún miembro de los grupos densos, i.e., no están asociados a un grupo denso.



**Figura 6.10:** Porcentajes de grupos espurios en cada una de las categorías para las muestras H-0 y H-5 y para todos los SAMs.

En la figura 6.10 se muestran los porcentajes de grupos espurios que pertenecen a cada categoría, tanto para la muestra tipo Hickson H-0 como para la H-5, en los cuatro SAMs.

Las categorías FM y MM+ contienen grupos Hickson espurios, donde todos sus miembros o más del 50% son miembros de un grupo denso 3D. La suma de estos porcentajes para la muestra H-0 supera el 60%, alcanzando para GII el 87%. Para la muestra H-5, esta supera el 50%. Este resultado es alentador respecto de los algoritmos observacionales: los criterios tipo Hickson no logran recuperar los grupos compactos 3D (densos y aislados), pero la eficiencia para identificar grupos densos es muy alta. Un muy bajo porcentaje de grupos tipo Hickson no son densos en 3D (NM).

Para profundizar aún más este análisis, nos preguntamos por qué los grupos densos asociados a los grupos Hickson de las categorías FM, MM+ y MM- no lograron llegar a ser grupos compactos 3D. En alguna etapa de la identificación de grupos compactos 3D, estos grupos no satisficieron alguno de los criterios y no llegaron a ser clasificados como grupos compactos. Existen diferentes etapas donde estos grupos pueden no pasar el filtro para ser clasificados compactos 3D:

- es subestructura de un grupo difuso

- es un grupo formado por una galaxia brillante y sus pequeñas satélites
- no está aislado localmente de acuerdo con el perfil de densidad

En la tabla 6.6 se muestra en qué etapa de la identificación 3D se han perdido estos grupos, ya que no llegaron a ser grupos aislados (compactos), para todos los SAMs. Para cada SAM, se muestran el número de grupos Hickson espurios que se relaciona con un grupo denso: FM, MM+, MM- ( los porcentajes son respecto del total de espurios, tal como se muestran en la figura 6.10). Luego, dentro de cada categoría, se ha calculado el porcentaje de grupos que no ha pasado el criterio de subestructura, luego por el de satélites, y lo mismo para el criterio de perfil de densidad.

Puede verse que, en general, el criterio de satélites no elimina muchos grupos, es decir, los grupos densos asociados a los espurios no suelen ser del tipo una galaxia y sus satélites pequeñas, lo que tal vez era esperable ya que el criterio observacional fue pensado originalmente para evitar esos casos, y entonces parece estar funcionando efectivamente. Lo más notable es que la mayoría de los grupos densos asociados a espurios que no llegaron a ser clasificados como compactos 3D fracasaron en estar aislados localmente de acuerdo con el perfil 3D de densidad numérica. Es decir, esos grupos espurios si bien son grupos densos no cumplen la aislación 3D. Podríamos decir que tal vez el criterio observacional para definir aislación estaría funcionando pobremente.

### 6.3.5. *Reals* vs CG-3D

Hasta aquí, hemos estudiado y caracterizado a las muestras identificadas con un criterio tipo Hickson y las hemos comparado con la muestra ideal de CGs 3D (densos y aislados) trasladada al espacio de redshifts. Hemos calculado los porcentajes de grupos recuperados por las muestras observacionales tipo Hickson, y por lo tanto pudimos tener una estima del porcentaje de grupos observacionales que son compactos 3D.

Ahora, en los capítulos previos, para estudiar el porcentaje de los grupos que eran considerados físicamente densos, utilizamos la definición de grupos *Reals* (ver sección 3.1.2 en el capítulo 3) definida por Díaz-Giménez & Mamon (2010). Esta definición no puede aplicarse a muestras observacionales sino sólo en catálogos sintéticos ya que necesita información 3D para la selección. Para que un CGs sea considerado *Real*, debe cumplir: que la distancia comóvil 3D máxima entre las 4 galaxias más cercanas ( $s_4$ ) sea menor a  $100 h^{-1} \text{kpc}$  ó que  $s_4 < 200 h^{-1} \text{kpc}$  mientras que la elongación de grupo ( $s_{\parallel}/s_{\perp}$ ) sea menor a 2. Observar que este método es construido en base a una muestra de CGs identificados al estilo observacional (en el espacio de redshifts) pero con restricciones impuestas sobre sus posiciones en el espacio real. A partir de este método, en todos los trabajos previos hemos podido clasificar cada catálogo por su porcentaje de grupos compactos *Reals*.

Muestra Categoría	H-0		
	FM	MM+	MM-
<b>G11</b>	# espurios = 397		
densos	191 ( 48%)	142 ( 36%)	39 ( 10%)
subestructura	30 ( 16%)	30 ( 21%)	11 ( 28%)
satélites	1 ( 1%)	7 ( 5%)	6 ( 15%)
perfil numérico	160 ( 84%)	105 ( 74%)	22 ( 56%)
<b>G13</b>	# espurios = 237		
densos	95 ( 40%)	80 ( 34%)	36 ( 15%)
subestructura	16 ( 17%)	22 ( 28%)	10 ( 28%)
satélites	0 ( 0%)	2 ( 3%)	0 ( 0%)
perfil numérico	79 ( 83%)	56 ( 70%)	26 ( 72%)
<b>Hr1</b>	# espurios = 152		
densos	44 ( 29%)	47 ( 31%)	22 ( 14%)
subestructura	8 ( 18%)	10 ( 21%)	7 ( 32%)
satélites	0 ( 0%)	3 ( 6%)	7 ( 32%)
perfil numérico	36 ( 82%)	34 ( 72%)	8 ( 36%)
<b>GII</b>	# espurios = 436		
densos	234 ( 54%)	148 ( 34%)	34 ( 8%)
subestructura	66 ( 28%)	45 ( 30%)	20 ( 59%)
satélites	0 ( 0%)	6 ( 4%)	3 ( 9%)
perfil numérico	168 ( 72%)	97 ( 66%)	11 ( 32%)
Muestra	H-5		
<b>G11</b>	# espurios = 1033		
densos	356 ( 34%)	387 ( 37%)	135 ( 13%)
subestructura	81 ( 23%)	83 ( 21%)	33 ( 24%)
satélites	18 ( 5%)	18 ( 5%)	17 ( 13%)
perfil numérico	257 ( 72%)	286 ( 74%)	85 ( 63%)
<b>G13</b>	# espurios = 914		
densos	259 ( 28%)	327 ( 36%)	174 ( 19%)
subestructura	52 ( 20%)	69 ( 21%)	55 ( 32%)
satélites	15 ( 6%)	15 ( 5%)	17 ( 10%)
perfil numérico	192 ( 74%)	243 ( 74%)	102 ( 59%)
<b>Hr1</b>	# espurios = 652		
densos	130 ( 20%)	205 ( 31%)	96 ( 15%)
subestructura	21 ( 16%)	45 ( 22%)	25 ( 26%)
satélites	3 ( 2%)	11 ( 5%)	12 ( 13%)
perfil numerico	106 ( 82%)	149 ( 73%)	59 ( 61%)
<b>GII</b>	# espurios = 1440		
densos	570 ( 40%)	577 ( 40%)	172 ( 12%)
subestructura	140 ( 25%)	162 ( 28%)	46 ( 27%)
satélites	22 ( 4%)	30 ( 5%)	32 ( 19%)
perfil numérico	408 ( 72%)	385 ( 67%)	94 ( 55%)

**Tabla 6.6:** Número de grupos espurios (con un grupo denso asociado) que no satisficieron el criterio de subestructura (primera fila), satélites (segunda fila) y perfil de densidad (tercera fila) para cada una de las categorías, en las muestra H-0 y H-5, y para los 4 SAMs.



El objetivo de esta sección es estudiar los grupos clasificados como *Reals* en las muestras 3D-CG<sub>m</sub>, H-0 y H-5, y de esta manera, poner a prueba qué tan bueno es el método de Díaz-Giménez & Mamon para efectivamente seleccionar grupos compactos 3D.

En la tabla 6.7 se muestran los números de grupos identificados para cada una de las muestras junto al número y porcentaje de grupos clasificados como *Reals*. Como primer resultado, puede observarse que los grupos de la muestra 3D-CG<sub>m</sub> (grupos por definición físicamente compactos 3D) son clasificados en su totalidad como *Reals* para los 4 SAMs. Esto demuestra que el método ideado por Díaz-Giménez & Mamon (2010) funciona a la perfección en una muestra de grupos que son efectivamente físicamente compactos.

Por otro lado, el porcentaje de *Reals* varía desde un 36% a un 60% para el caso de la muestra H-0 en los diferentes SAMs, y entre un 27% a un 52% para el caso de la muestra H-5. Estos porcentajes son similares a los obtenidos para los grupos compactos identificados con el criterio Hickson modificado en diferentes lighthcones en el capítulo 5 (Ver figura 5.4).

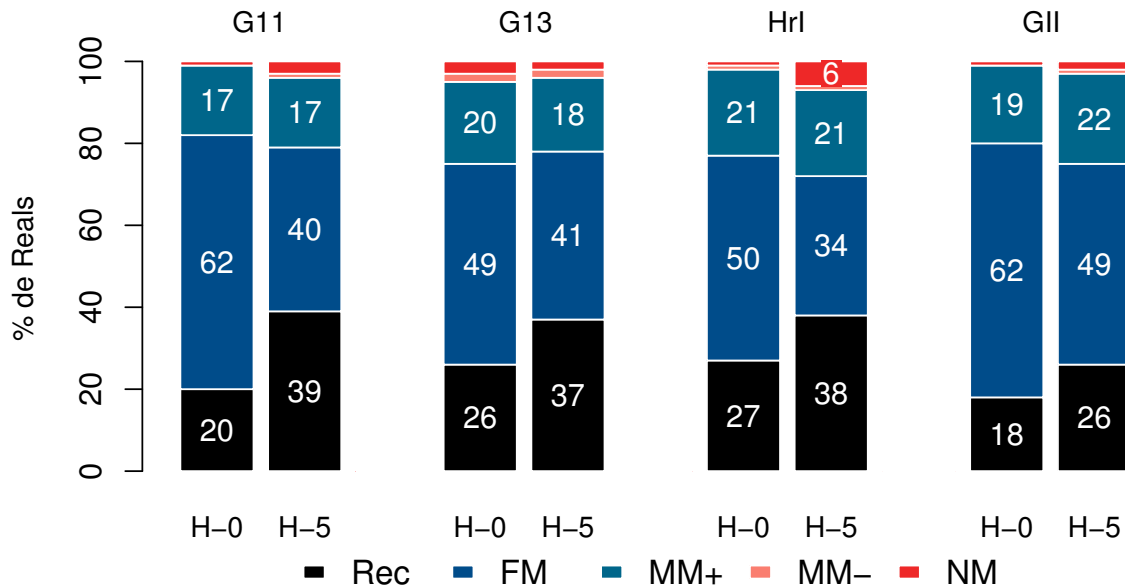
Comparado con la pureza determinada previamente para las muestras H-0 y H-5 (ya sea usando los 3D-CG<sub>m</sub> como los 3D-CG<sub>+1</sub>), vemos que el porcentaje de *Reals* es superior a lo que da *P* para cada muestra. Pensar a los *Reals* como grupos verdaderamente compactos en 3D es una extrapolación que no parece ser cierta.

	G11	G13	HrI	GII
3D-CG <sub>m</sub>	280	233	156	243
<i>Reals</i>	280 ( 100 %)	233 ( 100 %)	156 (100 %)	243 ( 100 %)
H-0	478	289	188	521
<i>Reals</i>	245 ( 51 %)	126 ( 44 %)	68 ( 36 %)	312 ( 60 %)
H-5	1527	1295	937	1919
<i>Reals</i>	636 ( 52 %)	440 ( 34 %)	256 ( 27 %)	934 ( 49 %)

**Tabla 6.7:** Números de CGs tipo Hickson y 3D trasladados al espacio de redshifts, y número (y porcentaje) de grupos clasificados como *Reals* para cada uno de los SAMs.

Para entender la naturaleza de los grupos que fueron seleccionados como *Reals* en las muestras tipo Hickson, hemos clasificado a los grupos *Reals* de acuerdo con las categorías definidas en la sección previa: primero vemos cuáles son recuperados (están asociados a un compacto 3D), y luego entre los espurios los clasificamos en las cuatro categorías de acuerdo con su asociación a algún grupo sobredenso. En la figura 6.11 se muestra la distribución de porcentajes de *Reals* en cada categoría, para las muestras H-0 y H-5, en cada uno de los SAMs. Podemos observar que la gran mayoría de los grupos *Reals* (más del 90% en todos los SAMs) fueron clasificados como Recuperados o como FM, MM+, es decir que tienen un grupo físicamente denso asociado con el cual comparten todos o la mayoría de sus miembros. Aunque menos del 40% de esos grupos *Reals*

es verdaderamente un compacto 3D (*Rec*). En general, menos del 10% (para algunos SAMs es del 1%) fueron clasificados como *MM-* ó *NM*. Esto nos dice que el método que clasifica los grupos como *Reals* es muy eficiente para definir una submuestra que verdaderamente está relacionada con un grupo físicamente denso.



**Figura 6.11:** Porcentajes de CGs clasificados como *Reals* para las muestras tipo Hickson en función de las categorías *Rec*, *FM*, *MM+*, *MM-* y *NM*.

Para complementar, podemos pensar las muestras al revés: ¿cuántos grupos Recuperados y espurios (*FM*, *MM+*, *MM-*, *NM*) en las muestras *H-0* y *H-5*, son clasificados como *Reals*? En la figura 6.12 se muestran, para cada SAM, el porcentaje de grupos *Reals* respecto de total de CGs *H-0* o *H-5* en cada una de esas categorías. La categoría *Rec* y *FM* son las que tienen más grupos clasificados como *Reals*. Los grupos *Reals* representan un  $\sim 30\%$  de la categoría *MM+*, mientras que en las categorías "*MM-z*" "*NM*" no llegan al 10%. En particular, el porcentaje de *Reals* varía entre un 34% y un 64% en la categoría recuperados, pero nunca llega al 100%. Por lo tanto, existen CGs clasificados como recuperados, es decir, grupos físicamente densos y aislados que no fueron clasificados como *Reals*! Esto es debido a que aún cuando todos los miembros de un CG-3D estén dentro del CG tipo Hickson, este último puede tener miembros intrusos que hacen que no se cumplan las condiciones establecidas para ser clasificado como *Real*.

En conclusión, el método propuesto por Díaz-Giménez & Mamon (2010) para clasificar a los CGs como sistemas físicamente densos (*Reals*) es muy bueno para definir una submuestra de sistemas que están de una manera u otra asociados con un sistema físicamente denso en 3D, aunque no necesariamente sean compactos (densos+aislados) en 3D. Sin embargo, eso no significa que aquellos

CGs que no cumplen con las condiciones para ser considerados *Reals* (los CALG y CAF) no estén asociados a un grupo físicamente denso en 3D, ya que hemos visto que casi el 90% de las muestras tipo Hickson están asociadas a sistemas sobreensos. En todo caso, el método es muy restrictivo para asegurarse de que todo lo que clasifica como *Real* estaría asociado con un sistema denso.

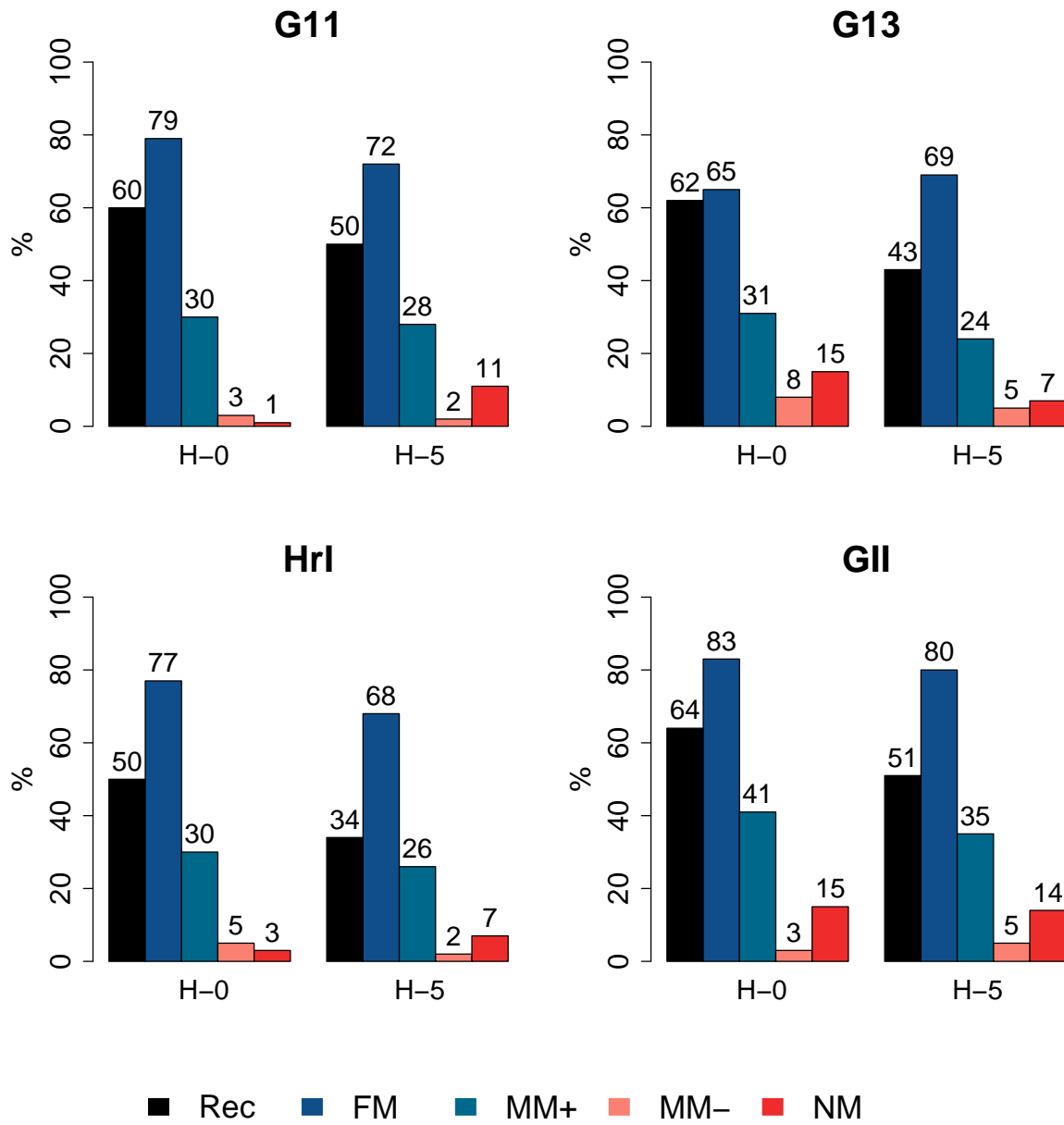
## 6.4. Resumen

Hemos diseñado un criterio de identificación de CGs en el espacio real usando sólo información 3D, definiendo como grupo compacto en el espacio real a aquellos que son sobredensos y se encuentran aislados en 3D. Esta definición nos ha permitido estudiar la naturaleza de los CG que resultan de identificaciones en el espacio de redshifts.

Encontramos que los CGs identificados en el espacio de redshifts por un criterio observacional tipo Hickson no son en su mayoría CGs en el espacio real, sólo entre un 18% y un 27% para los grupos identificados con el criterio original, H-0, y entre un 25% y un 39% para los grupos identificados con un criterio que intenta mejorar al criterio original (los H-5). Es decir que la pureza de las muestras observacionales es muy baja.

Analizando los restantes grupos, llamados espurios, encontramos que la gran mayoría son grupos que tienen un grupo denso asociado, es decir, que los criterios tipo Hickson identifican grupos físicamente densos, pero no logran ser grupos compactos 3D porque, o bien son subestructura de un grupo difuso, o bien no se encuentran aislados localmente en 3D.

Por otra parte, aproximadamente entre un 35% y un 50% de estos grupos tipo Hickson fueron clasificados como *Reals* (físicamente densos), dependiendo del SAM, como ya lo habíamos visto en capítulos anteriores. Pudimos comprobar que ese criterio es eficiente para seleccionar grupos físicamente densos en 3D, pero es extremadamente restrictivo, ya que encontramos que entre el 70% y el 90% (dependiendo del SAM) de los grupos tipo Hickson están asociados a grupos sobredensos en el espacio real, por lo que el criterio que los clasifica como *Reals* subestima el número de grupos que son sobredensos en 3D.



**Figura 6.12:** Porcentajes de grupos *Reals* relativo al total de CGs que hay en cada categoría.

# Capítulo 7

## Conclusiones

*Para finalizar, haremos un breve resumen de los resultados obtenidos en cada uno de los temas estudiados en esta tesis y enumeraremos nuestras principales conclusiones.*

### **7.1. Comparación de CGs identificados en diferentes bandas fotométricas**

*Hemos analizado las propiedades de los CGs identificados utilizando un mismo algoritmo de identificación pero seleccionando galaxias en diferentes bandas fotométricas. Utilizamos galaxias sintéticas usando el catálogo lightcone construido por Henriques et al. (2012), el cual usa el SAM de formación de galaxias de Guo et al. (2011) en la simulación Millennium Run I de materia oscura (Springel et al., 2005b).*

*Tres catálogos en diferentes bandas fotométricas,  $K$ ,  $r$  y  $u$ , fueron construidos seleccionando las galaxias más brillantes que una dada magnitud aparente límite. Hemos adoptado la magnitud aparente límite en la banda  $K$  para imitar la muestra observacional del relevamiento 2MASS ( $K_{\text{lim}} = 13,57$ ). Analizando las distribuciones magnitud-magnitud de las galaxias en el catálogo lightcone, hemos determinado los límites de las otras dos bandas para que sean equivalentes al límite en la banda  $K$  ( $r_{\text{lim}} = 16,4$  y  $u_{\text{lim}} = 18,73$ ).*

*En cada uno de los tres catálogos sintéticos se identificaron CGs aplicando un algoritmo automático tipo Hickson clásico. Primero, el algoritmo identifica CGs en proyección en el cielo, y luego utiliza el redshift de las galaxias miembro para eliminar intrusos. El criterio considera la membresía, la compacticidad, el aislamiento, y un límite en flujo de la galaxia brillante. Los valores límites para cada muestra fueron cambiados acorde a la banda fotométrica para obtener resultados similares.*

En un ángulo sólido de  $4\pi$ , hemos encontrado 447  $K$ -CGs, 406  $r$ -CGs y 276  $u$ -CGs. La comparación de las propiedades de los grupos compactos identificados en las 3 bandas ha revelado que los  $K$ -CGs y los  $u$ -CGs presentan radios proyectados más chicos que los  $r$ -CGs; sin embargo, analizando las separaciones comóviles 3D interpartículas entre los cuatro miembros más cercanos, los  $u$ -CGs tienen la separación más grande mientras que los  $K$ -CGs tienen la más pequeña. No hemos encontrado diferencias en el tiempo de cruce entre las tres muestras de CGs. La mayoría de estos resultados están en conflicto con previas comparaciones reportadas en la literatura (Díaz-Giménez et al., 2012; Hernández-Fernández & Mendes de Oliveira, 2015). La principal razón es debido a la dificultad de comparar muestras de grupos identificados con diferentes algoritmos lo que puede cambiar las propiedades resultantes, particularmente los tamaños que son parte del criterio de identificación. También hemos encontrado que las galaxias más brillantes en los  $u$ -CGs tienden a ser más brillantes que las galaxias más brillantes en los  $r$ - o  $K$ -CGs. La diferencia en magnitud entre las dos galaxias más brillantes en un grupo es más grande para los  $K$ -CGs que para los  $u$ -CGs. Los  $r$ -CGs presentan los valores más altos de brillos superficiales en cualquiera de las tres bandas, lo que significa que presentan la más baja compacticidad.

Hemos chequeado si estos resultados eran dependientes de las diferentes elecciones de los valores límites dependientes de la banda en el criterio de identificación. Definimos muestras "restringidas" de grupos compactos en las 3 bandas restringiendo las muestras usando también los valores límites de las otras dos bandas. Asimismo, se confirmó que los  $u$ -CGs son más pequeños en proyección, los  $K$ -CGs tienen la diferencia en magnitud más grandes y más pequeña separación 3D interpartícula, mientras que los  $r$ -CGs tienen el brillo superficial más alto y los tamaños proyectados más grandes. Varios de estos resultados han sido mejorados en las muestras restringidas comparando con los resultados obtenidos por las muestras totales.

A pesar de que normalizamos el criterio, las muestras de CGs aún son diferentes. No todos los CGs identificados en una banda están presentes también en las otras bandas. Para estudiar más específicamente qué diferencias son intrínsecamente debidas a la banda fotométrica en la que se observaron las galaxias, de las muestras de CGs restringidos hemos seleccionado aquellos grupos que son comunes en las tres identificaciones y aquellos que sólo existen en uno de los catálogos simulados (CGs puros). Los grupos comunes son el 51% de los  $u$ -CGs, sólo 1/3 de los  $K$ -CGs, y un 41% de los  $r$ -CGs. Los grupos que sólo fueron identificados en la banda  $K$  representan el 39% de los  $K$ -CGs restringidos, mientras que el porcentaje de  $r$ -puros es 17% de los  $r$ -CGs restringidos, y el 23% de los  $u$ -CGs restringidos fueron identificados sólo en la banda  $u$ . La comparación entre estas muestras indica que los CGs  $r$  puros son más grandes en proyección y ellos también tienen los valores más altos de brillo superficial (baja compacticidad). Los  $u$ -CGs puros poseen las más brillantes galaxias brillante del grupo, y las pequeñas diferencias entre las dos galaxias más brillantes del grupo. Los  $K$ -CGs puros tienen la más alta compacticidad y las más pequeñas

masas viriales y relación masa luminosidad. Esta muestra presenta las diferencias de magnitud más grande entre las dos galaxias más brillantes del grupo en comparación con todas las muestras de CGs utilizadas en este trabajo. Este resultado es una clara indicación de que esta característica es inherente a los grupos identificados sólo en esta banda fotométrica en particular. Este resultado también está relacionado con los valores muy bajos obtenidos para las estadísticas de Tremaine-Richstone,  $T1$  y  $T2$ , que se piensan comúnmente como indicación de fusiones de galaxias. Finalmente, los CGs que son comunes en las tres bandas presentan las separaciones 3D interpartículas más pequeñas entre las cuatro galaxias más cercanas, lo que los hace entidades físicas muy compactas en el espacio 3D, por lo tanto, esta muestra tiene el mayor porcentaje de CG *Reals* (usando la definición de *Reals* de Díaz-Giménez & Mamon 2010).

Nuestros resultados indican que la comparación de CGs identificados en diferentes bandas fotométricas tiene que hacerse cuidadosamente para no introducir sesgos relacionados a las diferentes funciones de selección, pero también hemos demostrado que hay características intrínsecas que difieren de banda a banda. Algunas de esas diferencias se puede difuminar en la mayor parte de los datos dada la existencia de grupos compactos comunes a todas las bandas fotométricas. A pesar de que estos resultados se obtuvieron a partir de un único SAM de formación de galaxias, este modelo reproduce muy bien la función de luminosidad local observable en las bandas fotométricas  $K$ ,  $r$  y  $u$ , y también el clustering de galaxias en función de las masas estelares y los colores (Guo et al., 2011; Henriques et al., 2012). Por lo tanto, es esperable que las diferencias entre los CGs sintéticos identificados en diferentes bandas sean representativas de las diferencias entre los CGs observacionales. Sin embargo, las predicciones presentadas en este trabajo deben confirmarse a partir de estudios realizados en catálogos observacionales con información multibanda, y/o cuando se realicen otros modelos de formación de galaxias con información en múltiples bandas fotométricas.

## 7.2. Optimización del algoritmo de identificación de CGs basado en el criterio de Hickson

La forma usual de identificar grupos compactos intenta reproducir los pasos seguidos por Hickson: se comienza buscando CGs en proyección en el plano del cielo que satisfacen los criterios de población, aislamiento y compactidad, y sólo cuando la muestra en proyección está disponible se examina si todos sus miembros tienen velocidades concordantes. Otros autores han seguido estos mismos pasos en la literatura (de Oliveira & Hickson, 1991; McConnachie et al., 2009; Díaz-Giménez & Mamon, 2010; Díaz-Giménez et al., 2012; Sohn et al., 2015). La principal desventaja de este algoritmo, que hemos llamado *clásico*, es que al contar las galaxias en un rango de magnitudes en un región proyectada, puede incluir muchas galaxias que se encuentra por delante o por detrás del grupo

en la dirección de la línea de la visual. Estas galaxias pueden ser intrusas que están afectando la población como así también el aislamiento de los potenciales grupos y eventualmente, algunos de estos grupos puede ser descartados de la muestra proyectada. Esta manera de identificar CGs puede comprometer la completitud resultante de la muestra.

Teniendo en cuenta esta posibilidad de perder grupos por la manera en que se lleva a cabo el criterio, hemos construido un nuevo algoritmo *modificado*. En vez de aplicar el filtro en velocidades después de la identificación en proyección, hemos impuesto dicho filtro en el mismo momento en que las galaxias van siendo seleccionadas como miembros o vecinas. El nuevo algoritmo fue puesto a prueba usando un catálogo sintético con evolución (*lightcone*) construido utilizando el SAM de Henriques et al. (2015). Dada la metodología usada para poner a prueba el algoritmo, la elección de un SAM diferente no afecta los resultados de estos tests.

La muestra resultante de CGs obtenida con el nuevo algoritmo fue comparada con la muestra obtenida usando el algoritmo *clásico*. Como resultado, hemos obtenido una muestra que es dos veces más grande que la versión clásica. En comparación, el algoritmo *modificado* incluye CGs con brillos superficiales cercanos al límite del criterio, tamaños proyectados más grandes, la galaxia brillante del grupo es más débil, y con las dos galaxias más brillantes más similares.

Por lo tanto, nuestro algoritmo *modificado* nos ha ayudado a mejorar la completitud de la muestras de CGs. Un interesante resultado es el obtenido al analizar la naturaleza de las muestras. Usando la información 3D en el espacio real de la simulación y siguiendo la definición de grupos físicamente densos y de alineamientos aleatorios de Díaz-Giménez & Mamon (2010), hemos investigado la ocurrencia de grupos densos (*Reals*) en nuestra muestra de grupos utilizando el algoritmo clásico y el modificado. Encontramos que el algoritmo no afecta la fracción de grupos *Reals*. Este resultado nos dice al mismo tiempo que existe la misma fracción de grupos físicamente densos en ambas muestras, aún cuando aumentamos el número de CGs hasta el doble.

Como corolario, presentamos un nuevo catálogo de CGs identificado en la muestra observacional de galaxias del catálogo Sloan Digital Sky Survey (SDSS) con nuevos datos espectroscópicos compilada por Tempel et al. (2017). Este nuevo catálogo contiene 462 sistemas de los cuales, 406 ( $\sim 88\%$ ) puede confirmarse que están efectivamente aislados según el criterio de Hickson original. Los restantes 56 CGs no han podido ser clasificados como aislados debido a posibles contaminaciones de galaxias en los alrededores. La incerteza surge a raíz de la incompletitud espectroscópica del catálogo observacional SDSS (ya sea por la saturación de galaxias brillantes o la colisión de fibras).

Además, realizamos una comparación detallada de esta nueva muestra observacional con la muestra de CGs identificados por Sohn et al. (2016), buscando coincidencias y diferencias. La comparación entre muestras de CGs identificados por diferentes autores, algoritmos, y criterios nunca es directa y debe tener-



se mucho cuidado cuando se sacan conclusiones sobre CGs que son identificados en una muestra pero que no aparecen en la de otro autor.

Siempre es importante contar con una muestra de CGs grande identificada de manera única y homogénea, con criterios de selección bien conocidos para obtener conclusiones estadísticamente confiables. Con este propósito en mente, tanto la muestra de Sohn et al. (2016) como la muestra obtenida en este trabajo cumplen el objetivo, y arrojan un poco de luz sobre las propiedades de estos pequeños sistemas peculiares, aún cuando estas muestras difieren en sus definiciones sobre qué entienden los autores qué son los CGs. Particularmente, los criterios que tienen en cuenta el límite en flujo, la aislación y los CGs embebidos son claves para determinar las propiedades de una muestra de CGs.

### 7.3. Naturaleza y frecuencia de CGs

Por otro lado, con el objetivo de testear el efecto de los parámetros cosmológicos y las recetas semianalíticas para construir galaxias sintéticas, hemos analizado diferentes muestras de CGs extraídas de catálogos lightcones construidos a partir de los cubos de tres simulaciones cosmológicas de materia oscura diferentes, combinadas con cinco modelos semianalíticos de formación de galaxias (SAMs), y usando diferentes funciones selección observacionales. Hemos analizado un total de once muestras de CGs y realizamos un estudio comparativo de la frecuencia y naturaleza de los CGs entre las diferentes muestras. Encontramos que, en general, todos los SAMs recuperan de buena manera las propiedades observacionales de los CGs.

Respecto de la frecuencia de los CGs vemos que la densidad espacial de CGs varía de SAM a SAM, casi un factor 4 entre el SAM con mayor (De Lucia et al., 2010) y menor frecuencia (Cora et al., 2018). En relación a los parámetros cosmológicos, la densidad espacial de CGs decrece fuertemente cuando aumenta la densidad del universo, dado que a mayores densidades se tiene más contaminación del criterio de aislamiento de los CGs. También se observó que la densidad espacial está fuertemente correlacionada con la fracción de galaxias huérfanas en los SAMs, dado que estas tienen un gran impacto en el clustering a pequeñas escalas. Cuando se compararon los resultados obtenidos con los datos observacionales encontramos que la densidad espacial de CGs predicha en los SAMs de Guo et al. (2011), Henriques et al. (2015) y Cora et al. (2018) es menor que la de las observaciones. Finalmente, vimos que la frecuencia y naturaleza de CGs no dependen de la banda fotométrica cuando se toman muestras totales en la banda  $K$  y  $r$ , ni tampoco del límite en flujo de la muestra.

Con respecto a la naturaleza de los CGs, vimos que la fracción de CGs físicamente densos (*Reals*) varía desde un 16% a un 56% dependiendo del SAM. También notamos que la fracción de *Reals* correlaciona con la fracción de galaxias huérfanas en los catálogos lightcones, por lo que las diferencias en los

*tratamientos de galaxias huérfanas en los modelos tienen un impacto directo en estudio de los CGs.*

*Por otro lado, la fracción de CGs calificados como alineaciones aleatorias aumenta a medida que aumenta el parámetro de densidad del universo (produciendo más galaxias a lo largo de la línea de la visual). Por lo tanto, los trabajos iniciales realizados con catálogos provenientes de la simulación Millennium I con cosmología WMAP1 (bajo parámetro de densidad) tienden a subestimar la importancia de las alineaciones aleatorias. También puede verse que la fracción de alineamientos aleatorios disminuye con el aumento de la resolución de la simulación, debido a que los grupos físicamente densos sufrirán menos la sobre fusión entre los subhalos de galaxias en simulaciones con mejor resolución.*

*Nuestro estudio muestra que los grupos físicamente densos pueden representar menos de la mitad de los CGs observados (como propusieron originalmente Mamon 1986, 1987; Walke & Mamon 1989), aunque esta idea podría no mantenerse cuando se usen simulaciones con mejores resoluciones.*

*Por lo tanto, los resultados obtenidos a partir de SAMs particulares sobre la naturaleza y frecuencia de CGs deben tomarse con cautela, ya que los resultados dependen de los parámetros cosmológicos, la resolución de la simulación utilizada sobre la cual se corrieron los modelos, así como también de las recetas de los modelos de formación de galaxias, particularmente aquellas relacionadas con el tratamiento de las galaxias huérfanas.*

## **7.4. CGs en el espacio 3D y su recuperación usando el criterio de Hickson**

*Por último, en la parte final de esta tesis propusimos la realización de un estudio sobre la naturaleza de los CGs pero partiendo de un lugar diferente al utilizado en el trabajo de Díaz-Giménez & Mamon (2010).*

*A lo largo de esta tesis hemos usado el trabajo de Díaz-Giménez & Mamon (2010) para calificar a los CGs obtenidos a partir de catálogos sintéticos en grupos físicamente densos o alineamientos aleatorios. En dicho trabajo, los autores usaron la información del espacio real para de alguna manera caracterizar la naturaleza de los CGs previamente identificados en un catálogo sintético limitado en flujo utilizando un criterio automático tipo Hickson.*

*Ahora, nuestro trabajo propuso partir de un lugar diferente, definiendo qué debería considerarse como un CG en el espacio real 3D y a partir de allí, analizar qué tan bueno es el criterio de Hickson para recuperar esa muestra de CGs, que podemos considerar *ideal*. Por lo tanto, como primera medida diseñamos un criterio de identificación de CGs en el espacio real que se aplicó a cuatro muestras de SAMs diferentes. Para definir qué deberíamos considerar como un*

grupo compacto 3D, se tomaron de guía las características globales que utilizó Hickson (1982): población, compacticidad y aislamiento, pero planteando todo en el espacio real. El proceso adoptado puede resumirse de la siguiente manera: partimos de la identificación de grupos sobredensos utilizando el algoritmo de percolación FoF (Davis et al., 1985) con un contraste de densidad considerablemente más alto que el que usualmente se utiliza, de tal manera que la separación entre galaxias fuese similar a la separación media de las galaxias en CGs observacionales. Una vez que se identificaron estos grupos físicamente densos, seleccionamos aquellos que no fueran subestructuras de grupos difusos. También eliminamos de la muestras de grupos densos a aquellos que son grupos formados por una galaxia masiva y sus galaxias satélites. Por último, aplicamos a los grupos restantes un criterio basado en el perfil de densidad numérica de las galaxias del grupo y las galaxias de su entorno, y nos quedamos con los grupos que están aislados localmente. Aplicando este criterio, obtuvimos la muestra de grupos compactos 3D en el espacio real, los cuales son grupos pequeños, densos y aislados. A partir de los CGs 3D, los “trasladamos” al espacio de redshifts y descartamos galaxias miembros más débiles que un límite en flujo, quedándonos con aquellos grupos que tienen 4 o más miembros más brillantes que el límite impuesto ( $3D-CG_m$ ). Esta muestra puede ser directamente comparable con los grupos identificados en un catálogo sintético limitado en flujo con algún criterio observacional. A la muestra de los  $3D-CG_m$ , las hemos considerado como la muestra ideal de CGs que uno quiere recuperar al identificar CGs en un catálogo observacional de galaxias.

Para llevar a cabo una comparación en el espacio de redshifts entre los grupos 3D y los grupos tipo Hickson, hemos identificado CGs utilizando el criterio modificado de CGs tipo Hickson, presentado en otro capítulo de esta tesis. Encontramos que el número de grupos recuperados, es decir, el número de CGs de las muestras tipo Hickson que también se encuentra en la muestra  $3D-CG_m$  es muy bajo ( $\sim 15\%$ ). Por lo tanto, la baja pureza del catálogo de CGs definido por el criterio de Hickson nos sugirió que dicho criterio tal vez podría ser mejorado variando sus parámetros iniciales. Analizando las propiedades de los  $3D-CG_m$  y comparándolas con las de los CGs tipo Hickson, realizamos modificaciones al criterio original de Hickson para ver si al variar los umbrales de las propiedades involucradas se podía maximizar la pureza de la muestra. Los resultados obtenidos mostraron que, aún cuando la completitud pudo ser mejorada, la pureza de las muestras no supera el 18% para ningunas de las muestras. Este resultado nos dice que el identificador tipo Hickson incluye muchos grupos espurios en su identificación, incluso cambiando sus parámetros, lo que nos lleva a pensar, que este criterio no es suficiente para seleccionar grupos compactos 3D.

Si bien modificando algunos parámetros del algoritmo de identificación se logra aumentar la completitud de la muestra, aún existe un alto porcentaje de CGs 3D que no fueron identificados con el criterio tipo Hickson. Esos CGs tienen que haber sido eliminados por las diferentes restricciones del criterio observacional, ya sea en población, compacticidad, aislamiento, límite en flujo o concordancia en velocidades. Las razones por las que no fueron encontrados no se examinaron

en detalle en este trabajo, serán evaluadas en un trabajo futuro.

Por otro lado, existen grupos identificados por criterios tipo Hickson que en realidad no existen en la muestra de CGs 3D en el espacio de redshift. Estos grupos espurios son los que contaminan la muestra observacional de CGs. Estudiamos la naturaleza de estos grupos espurios en el espacio real y los clasificamos en diferentes categorías dependiendo de si estaban asociados a un grupo denso en el espacio real, total o parcialmente, o no tenían ninguna conexión con un grupo denso 3D. Encontramos que la mayoría de los espurios está asociado a un grupo denso (total o casi totalmente), y un muy bajo porcentaje de los CGs identificados por Hickson están asociados sólo limitadamente con un denso 3D o no están asociados con ninguno. A partir de este resultado, que se encontró para todos los SAMs, concluimos que el identificador tipo Hickson si bien selecciona grupos espurios, más del 60% de ellos son pequeños y densos aunque no están aislados en el espacio real tridimensional.

Por lo tanto, y volviendo a la pregunta original de cuántos sistemas físicamente densos uno esperaría encontrar en un catálogo observacional de CGs identificados con el criterio de Hickson, la respuesta es que entre un 70% y un 90% (dependiendo del SAM) de los CGs tipo Hickson pueden considerarse asociados a un grupo físicamente denso en 3D, aún cuando muchos de estos sistemas no se encuentran aislados en el espacio real. Esto podría explicar la alta tasa de interacciones que se observan en las imágenes de las galaxias miembros de CGs observacionales.

Si observamos el resultado que obtendríamos aplicando el criterio desarrollado por Díaz-Giménez & Mamon (2010), encontraríamos que entre un 30% y un 50% (dependiendo del SAM) de los CGs tipo Hickson pueden considerarse físicamente densos. Esto claramente refleja que el criterio de Díaz-Giménez & Mamon (2010) se aseguró de clasificar sistemas físicamente densos (demostrado también en nuestro estudio) a costa de ser restrictivo. Por lo tanto, en el caso del resto de los sistemas no clasificados como *Reals*, que hemos llamado alineamientos aleatorios, en realidad un gran porcentaje de ellos también están asociados a un sistema físicamente denso aunque muy probablemente ese sistema no esté aislado.

## 7.5. Discusión Final

A lo largo de esta tesis hemos tratado de aportar nueva información a la problemática de la identificación de CGs en catálogos observacionales. Las herramientas por excelencia utilizadas en este trabajo han sido los catálogos sintéticos construidos a partir de simulaciones numéricas cosmológicas en combinación con modelos semianalíticos de formación de galaxias.

Nuestro trabajo ha brindado importantes avances tanto para el entendimiento

*de los diferentes catálogos observacionales, como para la construcción de nuevas muestras de CGs que representen una mejora en términos de completitud comparada con sus precursoras, y para la comprensión de los resultados que intentan describir la naturaleza de estos sistemas peculiares cuando se utilizan catálogos sintéticos. Todo esto centrado en mejorar el entendimiento sobre el tipo de sistemas que identifica el popular criterio de Hickson.*

*Después de todo el trabajo realizado, llegamos a la conclusión de que es necesario desarrollar un mejor identificador de CGs en catálogos observacionales que logre maximizar los porcentajes de pureza y completitud respecto de las muestras ideales definidas en el espacio real tridimensional que satisfacen los requisitos de sobredensidad y aislamiento espacial. Una muestra observacional resultante de ese identificador aportará importante información respecto a los procesos de evolución de galaxias en entornos extremos y aislados de la estructura en gran escala.*

*Continuará ...*

# Apéndice A

## Descarga de datos para los diferentes SAMs

Para este trabajo hemos utilizado las bases de datos de las simulaciones descritas en el capítulo 2. Por un lado hemos utilizado la base de datos de la Simulación Millennium<sup>1</sup>, trabajando con series de múltiples pedidos, para evitar la restricción de 7 minutos del servidor. De esta forma, siguiendo las sugerencias del manual de uso<sup>2</sup>, utilizamos los scripts de múltiples pedidos que se detallan a continuación:

- Los pedidos para la descarga de las galaxias del modelo DLB fueron:

```
1 SELECT c.galaxyID, c.type, c.x, c.y, c.z, c.velX, c.velY, c.velZ, c
    .stellarMass, c.bulgeMass, g.u_sdss, g.g_sdss, g.r_sdss, g.
    i_sdss, g.z_sdss, g.K_2mass, g.J_2mass
2 FROM MPAGalaxies..DeLucia2006a c, MPAGalaxies..
    DeLucia2006a_SDSS2MASS g
3 WHERE g.galaxyID=c.galaxyID AND c.galaxyID between :START
    *8*1e12 and (:START+:BIN)*8*1e12-1 AND c.snapnum=$sp AND
    g.r_sdss<-16 AND c.stellarMass>0.01
```

En este SAM, :START está en un rango 0-63 y :BIN=1 devuelve 64 pedidos por separado; y \$sp se mueve desde 63 ( $z = 0$ ) hasta 48 ( $z \sim 0,5$ ).

- Los pedidos para la descarga de las galaxias del modelo G11 fueron:

```
1 SELECT galaxyID, type, x, y, z, velX, velY, velZ, stellarMass,
    bulgeMass, uDust, gDust, rDust, iDust, zDust
2 FROM Guo2010a..MR
3 WHERE galaxyID between :START*1e12 and (:START+:BIN)*1e12-1
    AND snapnum=$sp and rDust<-16 AND stellarMass>0.01
```

---

<sup>1</sup><https://wwwmpa.mpa-garching.mpg.de/millennium/>

<sup>2</sup><http://gavo.mpa-garching.mpg.de/MyMillennium/Help?page=faqmillennium#IDs>

Este SAM :START se encuentra en el rango 0-511 y :BIN=1; y \$sp se mueve de 63 ( $z = 0$ ) hasta 48 ( $z \sim 0,5$ ).

■ Los pedidos para la descarga de las galaxias del modelo G11MII fueron:

```
1 SELECT galaxyID, type, x, y, z, velX, velY, velZ, stellarMass,
    bulgeMass, uDust, gDust, rDust, iDust, zDust
2 FROM Guo2010a..MRII
3 WHERE galaxyID between :START*16*1e15 and (:START+:BIN)
    *16*1e15-1 AND snapnum=$sp AND rDust<-11 AND stellarMass
    >0.00025
```

En este SAM :START se encuentra en el rango 0-31 y :BIN=1; \$sp se mueve de 67 ( $z = 0$ ) hasta 52 ( $z \sim 0,5$ ).

■ Los pedidos para la descarga de las galaxias del modelo G13 fueron:

```
1 SELECT galaxyID, type, x, y, z, velX, velY, velZ, stellarMass,
    bulgeMass, uDust, gDust, rDust, iDust, zDust
2 FROM Guo2013a..MR7
3 WHERE galaxyID between :START*1e12 and (:START+:BIN)*1e12-1
    AND snapnum=$sp AND rDust<-16 AND stellarMass>0.01
```

Este SAM :START se encuentra en el rango 0-511 y :BIN=1; \$sp se mueve de 61 ( $z = 0$ ) hasta 46 ( $z \sim 0,5$ ).

■ Los pedidos para la descarga de las galaxias del modelo HrI fueron:

```
1 SELECT galaxyID, type, x, y, z, velX, velY, velZ, stellarMass,
    bulgeMass, SDSSu_Dust, SDSSg_Dust, SDSSr_Dust, SDSSi_Dust,
    SDSSz_Dust, Ks_Dust, J_Dust
2 FROM Henriques2015a..MRscPlanck1
3 WHERE galaxyID between :START*16*1e12 and (:START+BIN)
    *16*1e12-1 AND snapnum=$sp AND SDSSr_Dust<-16 AND
    stellarMass>0.01
```

Este SAM :START esta en el rango 0-31 y :BIN=1; y \$sp se mueve desde 58 ( $z \sim 0$ ) hasta 45 ( $z \sim 0,51$ ).

■ Los pedidos para la descarga de las galaxias del modelo HII fueron:

```
1 SELECT galaxyID, type, x, y, z, velX, velY, velZ, stellarMass,
    bulgeMass, SDSSu_Dust, SDSSg_Dust, SDSSr_Dust, SDSSi_Dust,
    SDSSz_Dust, Ks_Dust, J_Dust
2 FROM Henriques2015a..MRIiscPlanck1
3 % WHERE galaxyID between :START*16*1e15 and (:START+BIN)
    *16*1e15-1 AND snapnum=$sp AND SDSSr_Dust<-11 AND
    stellarMass>0.00025
```

donde :START está en el rango 0-31 y :BIN=1; y \$sp se mueve desde 62 ( $z = 0$ ) hasta 49 ( $z \sim 0,51$ ).

Por último, ya que la base de datos de la simulación Multidark no tiene un tiempo de restricción para la descarga, fue posible obtener los datos a través de un formulario de consulta en el navegador web<sup>3</sup>. La obtención de dichos datos se realizó a través del siguiente pedido:

```
1 SELECT GalaxyStaticID, GalaxyType, x, y, z, vx, vy, vz, MstarSpheroid,  
    MstarDisk, MagStarSDSSu, MagStarSDSSg, MagStarSDSSr,  
    MagStarSDSSi, MagStarSDSSz  
2 FROM `MDPL2`.`SAG`  
3 WHERE snapnum=$sp AND MagStarSDSSr <= -18.2 AND (MstarSpheroid  
    +MstarDisk) >= 630957344
```

donde \$sp se mueve desde 125 ( $z = 0$ ) hasta 106 ( $z \sim 0,51$ ).

---

<sup>3</sup><https://www.cosmosim.org/>



# Apéndice B

## Cálculo de propiedades de CGs

### Estimador Biweight

El estimador Biweight ( $C_{BI}$ ) es un estimador robusto de localización de muestra el cual no es sensible a las distorsiones de las distribuciones (como una distribución con cola extensa) y también es resistente a los valores atípicos. Este método proporciona una mejor manera de seleccionar el valor que mejor describe la muestra:

$$C_{BI} = M + \frac{\sum_{|u_i| < 1} (x_i - M)(1 - u_i^2)^2}{\sum_{|u_i| < 1} (1 - u_i^2)^2} \quad (\text{B.1})$$

donde  $x_i$  son los elementos de la muestra,  $M$  es la mediana de la muestra, y  $u_i = (x_i - M)/cMAD$ , y  $MAD = \text{mediana}(|x_i - M|)$  es la desviación mediana absoluta y  $c = 6,0$  (Beers et al., 1990). El valor  $C_{BI}$  se puede pensar como el valor de la mediana más un factor de corrección.

En nuestro trabajo, hemos utilizado el estimador Biweight para el cálculo de todas las propiedades donde se utiliza la mediana:

- Redshift del grupo: mediana de los redshifts de las galaxias miembro;
- $\langle d_{ij} \rangle$  : mediana de la separación proyectada entre partículas.
- $\langle r_{ij} \rangle$  : mediana de la separación 3D entre partículas.

## Propiedades de CGs

**Masa virial** : calculada usando la fórmula de Limber & Mathews (1960):

$$\mathcal{M}_{\text{vir}} = \frac{\sigma^2 R_{\text{vir}}}{G} \quad (\text{B.2})$$

donde  $G$  es la constante de gravitación universal,  $\sigma$  es la dispersión de velocidades de las galaxias miembro, y  $R_{\text{vir}}$  es el radio virial del sistema.

## Radio virial en el espacio real

$$R_{\text{vir}}^{3D} = 2 R_h = 2 \frac{1}{C_{BI}(1/r_{ij})} \quad (\text{B.3})$$

donde  $R_h = 1/C_{bi}$  es el radio armónico calculado utilizando el estimador Biweight. Para este caso en particular  $C_{BI}$  representa la mediana de la distribución  $1/r_{ij}$

**Nota:** El radio virial es comúnmente obtenido usando la fórmula  $R_{\text{vir}}^{3D} = \frac{2N(N-1)}{\sum_{ij} \frac{1}{r_{ij}}}$  donde  $r_{ij}$  es la distancia tridimensional entre los miembros. Debido a que este radio se puede interpretar como la  $1/\text{media}(\frac{1}{r_{ij}})$ , hemos realizado este cálculo utilizando el estimador Biweight.

## Radio virial en el espacio de redshifts

$$R_{\text{vir}}^{3D} = \frac{\pi}{2} R_{\text{vir}}^p \quad (\text{B.4})$$

donde  $R_{\text{vir}}^p$  es el radio virial proyectado y se calcula utilizando la ecuación anterior pero con la separación proyectada entre los miembros  $d_{ij}$ , en vez de  $r_{ij}$ .

$$R_{\text{vir}}^p = 2 \frac{1}{C_{BI}(1/d_{ij})} \quad (\text{B.5})$$

donde  $C_{BI}$  representa la mediana de la distribución  $1/d_{ij}$ .

## Dispersión de velocidades en el espacio real

$$\sigma = \sqrt{\frac{\sum_i [(v_{x_i} - v_{x_{cm}})^2 + (v_{y_i} - v_{y_{cm}})^2 + (v_{z_i} - v_{z_{cm}})^2]}{N - 1}} \quad (\text{B.6})$$

donde,  $v_i = (v_{x_i}, v_{y_i}, v_{z_i})$  es el vector velocidad de las galaxias miembro y  $v_{cm} = (v_{x_{cm}}, v_{y_{cm}}, v_{z_{cm}})$  es el vector velocidad del centro de masa del sistema.

## Dispersión de velocidades en el espacio de redshifts

$$\sigma = \sqrt{3} \sigma_v \quad (\text{B.7})$$

donde  $\sigma_v$  es dispersión de velocidades radial calculada con los métodos de Beers et al. (1990).

Dada un muestra  $x_i, x_{i+1}, \dots, x_n$ , de velocidades radiales de las galaxias de un grupo de  $n$  miembros,

- Si  $n < 15$  (gapper)

$$\sigma_v = \frac{\sqrt{\pi}}{n(n-1)} \sum_{i=1}^{n-1} w_i g_i \quad (\text{B.8})$$

donde  $g_i = x_{i+1} - x_i$ , con  $i = 1, \dots, n-1$ , y  $w_i = i(n-i)$ .

- Si  $n \geq 15$  (biweight)

$$\sigma_v = C_{BI}(x) \quad (\text{B.9})$$

### Tiempo de cruce adimensional

$$H_0 t_{\text{cr}} = \frac{\pi 100 \text{ h} \langle d_{ij} \rangle}{2\sqrt{3} \sigma_v} \quad (\text{B.10})$$

### Brillo superficial del grupo

$$\mu = -2,5 \log(L/\text{area}) \quad (\text{B.11})$$

donde  $L = \sum_i 10^{-0,4m_i}$  es luminosidad total del grupo,  $m_i$  es la magnitud aparente de las galaxias miembros y el  $\text{area} = \pi \left(\frac{\theta_G}{2}\right)^2$ .

### Radio proyectado

$$r_p = s \tan(\theta_G^{\text{rad}}) \quad (\text{B.12})$$

donde  $\theta_G^{\text{rad}}$  es el diámetro angular  $\theta_G$  en radianes y  $s$  es la distancia física al centro del grupo.

**Estadísticos de Tremaine-Richstone** : calculados usando la fórmula de Tremaine & Richstone (1977).

$$T1 = \frac{\sigma(M_1)}{\langle M_2 - M_1 \rangle} \quad (\text{B.13})$$

$$T2 = \frac{1}{\sqrt{0,677}} \frac{\sigma(M_2 - M_1)}{\langle M_2 - M_1 \rangle} \quad (\text{B.14})$$

donde  $\sigma(X)$  es la desviación estandar de la variable  $X$ , y  $\langle X \rangle$  es la mediana de la variable.

# Apéndice C

## Tablas de galaxias del catálogo SDSS

En este apéndice proporcionamos más información sobre las galaxias utilizadas en el capítulo 4:

- En la tabla C.1 mostramos los números de identificación de objetos SDSS (ObjID) de 21 galaxias en el catálogo Tempel et al. (2017) los cuales fueron clasificados visualmente como parte de galaxias (part of a galaxy (PofG)).
- Seleccionamos galaxias a partir del SDSS DR14 usando la SDSS-III imaging pipeline. El pedido es el siguiente:

```
1 SELECT px.objID,px.ra,px.dec,px.modelMag_u,px.modelMag_g,px
   .modelMag_r,px.modelMag_i,px.modelMag_z,px.modelMagErr_u
   ,px.modelMagErr_g,px.modelMagErr_r,px.modelMagErr_i,px.
   modelMagErr_z,px.extinction_u,px.extinction_g,px.
   extinction_r,px.extinction_i,px.extinction_z,pxx.
   petroRad_r,pxx.petroRadErr_r,pxx.petroR50_r,pxx.
   petroR50Err_r,pxx.petroR90_r,pxx.petroR90Err_r,pz.z,pz.
   zErr,pz.nnCount,pz.photoErrorClass
2 FROM GalaxyTag as px,
3     Photoz as pz,
4     PhotoObjAll as pxx
5 WHERE px.modelMag_r<=17.77
6 AND px.type=3
7 AND px.mode=1
8 AND ((px.flags_r & 0x10000000) !=0)
9 AND ((px.flags_r & 0x8100000c00a0)=0)
10 AND (((px.flags_r & 0x4000000000000)=0)
11 OR (px.psfmagerr_r<=0.2))
```

SDSSObjID		SDSSObjID	
1	1237651823246573749	12	1237654601570451465
2	1237661873468080155	13	1237654601570451467
3	1237654879128060006	14	1237658491206041662
4	1237651822173552811	15	1237667915952816188
5	1237651252561379426	16	1237658424093769746
6	1237657857676607536	17	1237668273507729574
7	1237661816558714927	18	1237665025444282416
8	1237648704047284236	19	1237651251478724811
9	1237665330921013326	20	12376654415119647233
10	123766106926173412	21	12376719913395160477
11	123764872230141965		

**Tabla C.1:** Lista de galaxias en el catálogo Tempel et al. 2017 alrededor de potenciales grupos compactos que fueron visualmente clasificadas como partes de galaxias (PofG).

```

12 AND ((px.flags_r & 0x100000000000)=0)
13 OR (px.flags_r & 0x1000)=0)
14 AND px.ObjID=pz.ObjID
15 AND pxx.ObjID=px.ObjID
16 AND px.specObjID=0

```

*El método usado para calcular los redshifts fotométricos estimados recuperados con este pedido es descrito por Beck et al. (2016).*

- *A partir de la muestra de galaxias sin redshift descrita en los items anteriores, hemos seleccionado aquellas galaxias que satisfacen  $r > 0$ ,  $g-r < 3$ . Las magnitudes fueron corregidas por extinción y transformadas al sistema AB (Doi et al., 2010). Hemos encontrado que algunas de estas galaxias tienen sus redshifts disponibles en la NED. En la tabla C.2 citamos los redshifts encontrados para las galaxias fotométricas alrededor de candidatos a CGs. Hemos corregido los redshifts al sistema en reposo del CMB usando la ecuación estándar:*

$$z_{\text{CMB}} = z_{\text{NED}} + z_a [\sin b \sin b_a + \cos b \cos b_a \cos (l_a - l)],$$

*donde  $l$  y  $b$  son las coordenadas galácticas,  $l_a = 264,14^\circ$ ,  $b_a = 48,26^\circ$ , y  $z_a = 371,0/c$  (subíndice  $a$  es la abreviatura de *apex*) y  $c$  es la velocidad de la luz en el vacío km/s.*

- *En la tabla C.3 mostramos las 68 galaxias fotométricas consideradas como fuentes potenciales de contaminación alrededor de los HMCGs.*

ObjID	$z_{\text{NED}}$	$z_{\text{CMB}}$	ObjID	$z_{\text{NED}}$	$z_{\text{CMB}}$		
1	1237665566604787742	0.061476	0.062265	33	1237667323797766247	0.026822	0.027721
2	1237658609296146595	0.056870	0.057201	34	1237667323797766298	0.072250	0.073149
3	1237664130484273189	0.142020	0.143180	35	1237667323797766905	0.022800	0.023697
4	1237664130484273365	0.065240	0.066400	36	1237661387063885956	0.080280	0.080741
5	1237664130484207707	0.068550	0.069710	37	1237658311867105407	0.059360	0.059821
6	1237664130484207928	0.069491	0.070651	38	1237661465452740882	0.048000	0.048192
7	1237661850938572946	0.087930	0.088548	39	1237659329781367066	0.066270	0.066296
8	1237657608567783547	0.057460	0.058040	40	1237664871900643452	0.050148	0.051053
9	1237657595683078689	0.058420	0.058998	41	1237657630057562291	0.061040	0.061713
10	1237657608567783553	0.057460	0.058041	42	1237654880205209753	0.091600	0.092229
11	1237651272419508510	0.048090	0.048565	43	1237662199355146310	0.051142	0.051814
12	1237661850928676959	0.024390	0.025164	44	1237667550347460631	0.025124	0.026180
13	1237661872399646878	0.094300	0.095075	45	1237667550347460635	0.025124	0.026180
14	1237661872399646879	0.094300	0.095076	46	1237662637443842271	0.045111	0.045687
15	1237655373572866101	0.028140	0.028216	47	1237651271888208191	0.029890	0.030349
16	1237661872949297333	0.092020	0.092573	48	1237654604253233186	0.041870	0.043081
17	1237661872949297281	0.090300	0.090854	49	1237665351854915780	0.069428	0.070044
18	1237667734492020911	0.014460	0.015493	50	1237654601018507679	0.028610	0.029593
19	1237665531723317688	0.037763	0.038075	51	1237654604776407247	0.027900	0.028882
20	1237668670253039811	0.071470	0.071932	52	1237655691404968131	0.039110	0.039771
21	1237668294985842855	0.162200	0.163290	53	1237655691404902688	0.072650	0.073313
22	1237651753457483933	0.071500	0.072612	54	1237648722841043108	0.078600	0.079802
23	1237658492269494418	0.047310	0.048481	55	1237667253462892735	0.025657	0.026634
24	1237653617470603328	0.045010	0.045497	56	1237651822717239556	0.043920	0.044412
25	1237655108374757517	0.049380	0.049877	57	1237657630599872534	0.027052	0.027669
26	1237665443126116438	0.051652	0.052523	58	1237674650998145068	0.024140	0.025366
27	1237661416068546653	0.036950	0.037288	59	1237667254015688747	0.012929	0.013876
28	1237657856064618588	0.024980	0.025678	60	1237671751364313102	0.008382	0.008707
29	1237648720164880479	0.082030	0.083096	61	1237662335183225035	0.134000	0.134208
30	1237648720164880703	0.084400	0.085466	62	1237665548887261200	0.029552	0.030330
31	1237654605873348678	0.075669	0.076827	63	1237667541753266247	0.024240	0.025276
32	1237651822716715274	0.119100	0.119615				

**Tabla C.2:** Redshift corregido al sistema en reposos del CMB para 63 galaxias fotométricas en el SDSS DR14 extraídas de la base de datos NASA/IPAC Extragalactic Database (NED).

SDSSObjID		SDSSObjID	
1	1237651272419509016	35	1237661850928611367
2	1237660635999764658	36	1237665442601107677
3	1237661971727515740	37	1237665442601107712
4	1237661416068546838	38	1237665442601107767
5	1237662199356850525	39	1237665548349931560
6	1237665442601107571	40	1237662336256442628
7	1237667781744328983	41	1237662337865482596
8	1237651252027326626	42	1237662337865482480
9	1237653617470603455	43	1237662636911034773
10	1237651822716649609	44	1237665179521319111
11	1237651753457418421	45	1237665533325541581
12	1237654948378771564	46	1237665533325541619
13	1237657857144717348	47	1237665429168783623
14	1237660412650193181	48	1237665179522760752
15	1237654605873348872	49	1237661949196763146
16	1237658493347758315	50	1237664295293812934
17	1237655693557891393	51	1237667538550456389
18	1237659162276135058	52	1237668273507729489
19	1237658492280242289	53	1237668272440672468
20	1237661971727515766	54	1237668495776219326
21	1237660671429967980	55	1237667448873549879
22	1237660671429967981	56	1237667782270648349
23	1237661387617730866	57	1237667911138279447
24	1237662199355146533	58	1237668298738172072
25	1237662528995852513	59	1237668289620672636
26	1237661086961107151	60	1237668611197567138
27	1237662636370100371	61	1237668336862691636
28	1237665429170618580	62	1237668336862691739
29	1237661418747134094	63	1237671991336239245
30	1237665548351438938	64	1237654390570680607
31	1237665442601042133	65	1237658205048930492
32	1237662194533531763	66	1237665548351438942
33	1237665533338255410	67	1237665567153127563
34	1237662337865417030	68	1237668272440672437

**Tabla C.3:** Lista de galaxias en el catálogo fotométrico SDSS DR14 alrededor de grupos compactos que son potenciales fuentes de contaminación.

# Bibliografia

- Abazajian K., et al., 2004, AJ, 128, 502*
- Abazajian K. N., et al., 2009, ApJS, 182, 543*
- Abolfathi B., et al., 2018, ApJS, 235, 42*
- Alam S., et al., 2015, ApJS, 219, 12*
- Allam S. S., Tucker D. L., 2000, Astronomische Nachrichten, 321, 101*
- Baier F. W., Tiersch H., 1975, Astrofizika, 11, 221*
- Baier F. B., Tiersch H., 1976, Astrofizika, 12, 7*
- Baier F. W., Tiersch H., 1978, Astrofizika, 14, 279*
- Baier F. W., Tiersch H., 1979, Astrofizika, 15, 33*
- Baier F. W., Petrosyan M. B., Tiersch H., Shakhbazyan R. K., 1974, Astrofizika, 10, 327*
- Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945*
- Baldry I. K., et al., 2012, MNRAS, 421, 621*
- Barton E., Geller M., Ramella M., Marzke R. O., da Costa L. N., 1996, AJ, 112, 871*
- Bayer F. W., Tiersch H., 1976, Astrofizika, 12, 409*
- Beck R., Dobos L., Budavári T., Szalay A. S., Csabai I., 2016, MNRAS, 460, 1371*
- Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32*
- Berlind A. A., et al., 2006, ApJS, 167, 1*
- Bertone S., De Lucia G., Thomas P. A., 2007, MNRAS, 379, 1143*
- Bianchi L., 2014, Ap&SS, 354, 103*
- Blanton M. R., Lupton R. H., Schlegel D. J., Strauss M. A., Brinkmann J., Fukugita M., Loveday J., 2005, ApJ, 631, 208*



- Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, *MNRAS*, 398, 1150
- Chilingarian I. V., Zolotukhin I. Y., 2012, *MNRAS*, 419, 1727
- Coenda V., Muriel H., Martínez H. J., 2015, *A&A*, 573, A96
- Cohen M., Wheaton W. A., Megeath S. T., 2003, *AJ*, 126, 1090
- Colless M., et al., 2001, *MNRAS*, 328, 1039
- Colless M., et al., 2003, *ArXiv Astrophysics e-prints*,
- Conselice C. J., 2006, *MNRAS*, 373, 1389
- Cora S. A., et al., 2018, *MNRAS*, 479, 2
- Corwin Jr. H. G., Buta R. J., de Vaucouleurs G., 1994, *AJ*, 108, 2128
- Cousins A. W. J., 1976, *MmRAS*, 81, 25
- Davis M., Efsthathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- De Lucia G., Blaizot J., 2007, *MNRAS*, 375, 2
- De Lucia G., Boylan-Kolchin M., Benson A. J., Fontanot F., Monaco P., 2010, *MNRAS*, 406, 1533
- Díaz-Giménez E., Mamon G. A., 2010, *MNRAS*, 409, 1227
- Díaz-Giménez E., Zandivarez A., 2015, *A&A*, 578, A61
- Díaz-Giménez E., Mamon G. A., Pacheco M., de Oliveira C. M., Alonso M. V., 2012, *Monthly Notices of the Royal Astronomical Society*, 426, 296
- Díaz-Giménez E., Zandivarez A., Taverna A., 2018a, *VizieR Online Data Catalog*, pp J/A+A/618/A157
- Díaz-Giménez E., Zandivarez A., Taverna A., 2018b, *A&A*, 618, A157
- Djorgovski S. G., Gal R. R., Odewahn S. C., Brunner R. J., de Carvalho R. R., 1998, in *American Astronomical Society Meeting Abstracts*. p. 13.01
- Doi M., et al., 2010, *AJ*, 139, 1628
- Duplancic F., O'Mill A. L., Lambas D. G., Sodr e L., Alonso S., 2013, *MNRAS*, 433, 3547
- Duplancic F., Coldwell G. V., Alonso S., Lambas D. G., 2018, *MNRAS*, 481, 2458
- Eisenstein D. J., et al., 2011, *AJ*, 142, 72
- Eke V. R., et al., 2004, *MNRAS*, 348, 866

- Farhang A., Khosroshahi H. G., Mamon G. A., Dariush A. A., Raouf M., 2017, *ApJ*, 840, 58
- Focardi P., Kelm B., 2002, *Astronomy and Astrophysics*, 391, 35
- Guo Q., et al., 2011, *MNRAS*, 413, 101
- Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, *MNRAS*, 428, 1351
- Hearin A. P., et al., 2017, *AJ*, 154, 190
- Henriques B. M. B., White S. D. M., Lemson G., Thomas P. A., Guo Q., Marleau G.-D., Overzier R. A., 2012, *Monthly Notices of the Royal Astronomical Society*, 421, 2904
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, 451, 2663
- Hernández-Fernández J. D., Mendes de Oliveira C., 2015, *MNRAS*, 453, 1965
- Hickson P., 1982, *ApJ*, 255, 382
- Hickson P., Mendes de Oliveira C., Huchra J. P., Palumbo G. G., 1992, *ApJ*, 399, 353
- Hollowed J., 2019
- Huchra J. P., Geller M. J., 1982, *ApJ*, 257, 423
- Huchra J. P., et al., 2012, *ApJS*, 199, 26
- Hwang H. S., Elbaz D., Lee J. C., Jeong W.-S., Park C., Lee M. G., Lee H. M., 2010, *A&A*, 522, A33
- Iovino A., 2002, *AJ*, 124, 2471
- Iovino A., de Carvalho R. R., Gal R. R., Odewahn S. C., Lopes P. A. A., Mahabal A., Djorgovski S. G., 2003, *AJ*, 125, 1660
- Jarrett T. H., Chester T., Cutri R., Schneider S. E., Huchra J. P., 2003, *AJ*, 125, 525
- Johnson H. L., Morgan W. W., 1953, *ApJ*, 117, 313
- Kampstra P., 2008, *Journal of Statistical Software, Code Snippets*, 28, 1
- Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, *MNRAS*, 457, 4340
- Komatsu E., et al., 2011, *ApJS*, 192, 18
- Krzywinski M., Altman N., 2014, *Nature Methods*, 11, 119
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64

- Lange R., et al., 2015, *Monthly Notices of the Royal Astronomical Society*, 447, 2603
- Lee B. C., et al., 2004, *Astron J*, 127, 1811
- Li C., White S. D. M., 2009, *MNRAS*, 398, 2177
- Limber D. N., Mathews W. G., 1960, *ApJ*, 132, 286
- Mamon G. A., 1986, *ApJ*, 307, 426
- Mamon G. A., 1987, *ApJ*, 321, 622
- McConnachie A. W., Ellison S. L., Patton D. R., 2008, *MNRAS*, 387, 1281
- McConnachie A. W., Patton D. R., Ellison S. L., Simard L., 2009, *MNRAS*, 395, 255
- McGill R., Tukey J. W., Larsen W. A., 1978, *The American Statistician*, 32, 12
- Mendes de Oliveira C., Torres-Flores S., Amram P., Plana H., Epinat B., 2015, in Ziegler B. L., Combes F., Dannerbauer H., Verdugo M., eds, *IAU Symposium Vol. 309, Galaxies in 3D across the Universe*. pp 175–177, doi:10.1017/S1743921314009612
- O'Mill A. L., Duplancic F., García Lambas D., Valotto C., Sodr e L., 2012, *MNRAS*, 421, 1897
- Petrosian M. B., 1974, *Astrofizika*, 10, 471
- Petrosyan M. B., 1978, *Astrofizika*, 14, 631
- Planck Collaboration et al., 2015, *A&A*, 580, A22
- Prandoni I., Iovino A., MacGillivray H. T., 1994, *AJ*, 107, 1235
- Pujol A., et al., 2017, *MNRAS*, 469, 749
- R Core Team 2013, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org/>
- Ramella M., Diaferio A., Geller M. J., Huchra J. P., 1994, *AJ*, 107, 1623
- Robinson L. B., Wampler E. J., 1973, *ApJ*, 179, L79
- Rose J. A., 1977, *ApJ*, 211, 311
- Scrucca L., Fop M., Murphy T. B., Raftery A. E., 2017, *The R Journal*, 8, 205
- Seyfert C. K., 1948, *The Astronomical Journal*, 53, 203
- Shakhbazyan R. K., 1973, *Astrofizika*, 9, 495

- Shakhbazyan R. K., Petrosyan M. B., 1974, *Astrofizika*, 10, 13
- Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, *MNRAS*, 343, 978
- Shi F., et al., 2016, *ApJ*, 833, 241
- Skrutskie M. F., et al., 2006a, *AJ*, 131, 1163
- Skrutskie M. F., et al., 2006b, *AJ*, 131, 1163
- Sohn J., Hwang H. S., Geller M. J., Diaferio A., Rines K. J., Lee M. G., Lee G.-H., 2015, *Journal of Korean Astronomical Society*, 48, 381
- Sohn J., Geller M. J., Hwang H. S., Zahid H. J., Lee M. G., 2016, *ApJS*, 225, 23
- Spergel D. N., et al., 2003, *ApJS*, 148, 175
- Springel V., Yoshida N., White S. D. M., 2001, *New Astronomy*, 6, 79
- Springel V., et al., 2005a, *Nature*, 435, 629
- Springel V., et al., 2005b, *Nature*, 435, 629
- Stephan C. P., 1877, *Information Bulletin on Variable Stars*, 1350
- Targett T. A., Dunlop J. S., McLure R. J., 2012, *MNRAS*, 420, 3621
- Taverna A., 2015, *Master's thesis, Facultad de Matematica, Astronomía, Física y Computación, Universidad Nacional de Córdoba, Cordoba, Argentina*
- Taverna A., Díaz-Giménez E., Zandivarez A., Joray F., Kanagusuku M. J., 2016, *MNRAS*, 461, 1539
- Tempel E., Twikene T., Kipper R., Libeskind N. I., 2017, *A&A*, 602, A100
- Tremaine S. D., Richstone D. O., 1977, *ApJ*, 212, 311
- Walke D. G., Mamon G. A., 1989, *A&A*, 225, 291
- Wiens C. D., Wenger T. V., Tzanavaris P., Johnson K. E., Gallagher S. C., Xiao L., 2019, *ApJ*, 873, 124
- Yentis D. J., Cruddace R. G., Gursky H., Stuart B. V., Wallin J. F., MacGillivray H. T., Collins C. A., 1992, in MacGillivray H. T., Thomson E. B., eds, *Astrophysics and Space Science Library Vol. 174, Digitised Optical Sky Surveys*. p. 67, doi:10.1007/978-94-011-2472-0\_10
- York D. G., SDSS Collaboration 2000, *AJ*, 120, 1579
- Zandivarez A., Martínez H. J., 2011, *MNRAS*, 415, 2553
- Zandivarez A., Díaz-Giménez E., Mendes de Oliveira C., Ascaso B., Benítez N., Dupke R., Sodr  L., Irwin J., 2014, *A&A*, 561, A71

*de Carvalho R. R., Gonçalves T. S., Iovino A., Kohl-Moreira J. L., Gal R. R., Djorgovski S. G., 2005, AJ, 130, 425*

*de Oliveira C. M., Hickson P., 1991, ApJ, 380, 30*

*de Vaucouleurs G., de Vaucouleurs A., Corwin Jr. H. G., Buta R. J., Paturel G., Fouqué P., 1991, Third Reference Catalogue of Bright Galaxies. Volume I: Explanations and references. Volume II: Data for galaxies between  $0^h$  and  $12^h$ . Volume III: Data for galaxies between  $12^h$  and  $24^h$ .*