



UNIVERSIDADE ESTADUAL DE SANTA CRUZ  
PRO-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM COMPUTACIONAL  
EM CIÊNCIA E TECNOLOGIA

**ALLAN SÉRGIO GONÇALVES ALVES**

**MODELAGEM DINÂMICA DA DISTRIBUIÇÃO DE GALÁXIAS EM  
AGLOMERADOS**

**ILHÉUS-BA  
2018**

**ALLAN SÉRGIO GONÇALVES ALVES**

**MODELAGEM DINÂMICA DA DISTRIBUIÇÃO DE  
GALÁXIAS EM AGLOMERADOS**

Dissertação apresentada ao Programa de Pós-Graduação  
em Modelagem Computacional em Ciência e Tecnologia  
da Universidade Estadual de Santa Cruz, como parte  
das exigências para obtenção do título de Mestre em  
Modelagem Computacional em Ciência e Tecnologia.

Orientador: Prof. Dr. André Luis Batista Ribeiro

ILHÉUS-BA  
2018

A474

Alves, Allan Sérgio Gonçalves.

Modelagem dinâmica da distribuição de galáxias em aglomerados / Allan Sérgio Gonçalves Alves. - Ilhéus : UESC, 2018.

67f. : il.

Orientador : André Luís Batista Ribeiro.

Dissertação (Mestrado) – Universidade Estadual de Santa Cruz. Programa de Pós-graduação em Modelagem Computacional em Ciência e Tecnologia.

Inclui referências e apêndices.

1. Galáxias – Aglomerados. 2. Estrelas – Aglomerados. 3. Cosmologia - Métodos estatísticos. 4. Simulação (computadores)  
I.Ribeiro, André Luís Batista. II. Título.

CDD – 523.112

ALLAN SÉRGIO GONÇALVES ALVES

**MODELAGEM DINÂMICA DA DISTRIBUIÇÃO DE  
GALÁXIAS NO ESPAÇO DE FASE DE  
AGLOMERADOS**

Ilhéus-BA, 27/02/2018

Comissão Examinadora

Prof. Dr. André Luis Batista Ribeiro  
UESC  
(Orientador)

Prof. Dr. Francisco Bruno Souza Oliveira  
UESC

Prof. Dr. Sandro Barboza Rembold  
UFSM

À minha mãe, Tânia (*In memoriam*), que sempre esteve comigo durante toda a jornada, dedico.

## **Agradecimentos**

- Ao Prof. André Ribeiro, meu orientador, pela atenção, paciência e oportunidade de aprendizado.
- À CAPES, pelo apoio financeiro.
- À UESC e ao PPGMC, pela estrutura e o apoio sempre concedido.
- Aos meus colegas de mestrado, principalmente, Messias Novaes, Bruno Calazans, Uebert Moreira e Romilson Souza, pelo companheirismo.
- Ao Alisson Costa, do mestrado em Física, por enfrentar, juntamente comigo, os perfis de densidade e a anisotropia.
- À Fabiana, minha namorada, que me apoiou nos momentos de dificuldade.
- À minha irmã, Pollyana, e sobrinha, Lunna, pelo apoio e carinho.

*"The Cosmos is all that is or was or ever will be. Our feeblest contemplations of the Cosmos stir us – there is a tingling in the spine, a catch in the voice, a faint sensation, as if a distant memory, of falling from a height. We know we are approaching the greatest of mysteries."*

Carl Sagan, *Cosmos*

# MODELAGEM DINÂMICA DA DISTRIBUIÇÃO DE GALÁXIAS EM AGLOMERADOS

## Resumo

Os aglomerados de galáxias são as maiores estruturas virializadas e ligadas pela gravidade que existem no universo. Um dos principais problemas no estudo de aglomerados é a determinação da sua aproximação ao estado de equilíbrio, pois perturbações gravitacionais originárias do encontro entre objetos ou a absorção de outras estruturas podem modificar a sua dinâmica. Diversos fenômenos acontecem em aglomerados de galáxias e podem determinar a sua configuração, como os processos de relaxação violenta e a relaxação entre dois corpos. Este trabalho tem como objetivo caracterizar da forma mais ampla possível esses sistemas através da implementação de técnicas para analisar a distribuição de velocidades de galáxias em aglomerados e seus perfis de densidade. Propomos, também, a utilização dos L-Momentos da distribuição gaussiana para recuperar alguns dos aglomerados diagnosticados como indeterminados pela ferramenta GNG, descrita em [Ribeiro et al. \(2013\)](#). As ferramentas desenvolvidas foram aplicadas aos dados oriundos das simulações EAGLE e Millennium. Para os dados do Millennium, encontramos resultados inconsistentes se comparados aos encontrados por [Ribeiro et al. \(2013\)](#). Encontramos um suave aumento na fração de objetos com massa  $13.75 \leq \log_{10} M_{200} < 14.25$ , classificados como gaussianos entre o redshift  $z = 0.5$  e  $z = 0.0$ . Além disso, ao empilharmos os raios projetados das galáxias pertencentes a aglomerados classificados como G e NG, notamos que a pilha de não-gaussianos apresenta uma queda menos acentuada no perfil de densidade superficial em comparação a pilha de gaussianos. Para os dados do EAGLE, encontramos a tendência de leve queda dos valores da Hellinger Distance com o decorrer das épocas. Vimos também que o corte em magnitude faz com que apresentem-se algumas diferenças no diagnóstico do GNG para os mesmos objetos.

**Palavras-chave:** Aglomerados de galáxias. Simulação computacional. Métodos estatísticos. Máxima verossimilhança.

## Abstract

Galaxy clusters are the largest gravitationally bound structures virialized in the universe. One of the main problems in the study of clusters of galaxies is the determination of their dynamic state, since gravitational perturbations from the encounter between objects or the capture of other structures can modify their dynamics. Several phenomena occur in clusters of galaxies and can determine their configuration, such as violent relaxation and relaxation between two bodies. This work aims to characterize as widely as possible these systems through the implementation of techniques to analyze the velocity distribution of galaxies in clusters and their density profiles. We propose the use of the L-Moments of the Gaussian distribution to recover some systems diagnosed as Indeterminate by the GNG tool, described in [Ribeiro et al. \(2013\)](#). The tools developed were applied to data from the EAGLE and Millennium simulations. For Millennium simulation, we found inconsistent results compared to those found by [Ribeiro et al. \(2013\)](#). We found a slight increase in the fraction of objects with mass  $13.75 \leq \log_{10}M_{200} < 14.25$ , classified as Gaussian between redshift  $z = 0.5$  and  $z = 0.0$ . Also, when we stack the galaxies belonging to clusters classified as G and NG, we note that the stack of Non-Gaussian presents a profile with a less pronounced drop in the surface density profile when compared to the Gaussian stack. For the simulation of the EAGLE, we find the trend of decreasing values of Hellinger Distance with the passage of time. We also saw that the cut in magnitude allows some differences in the GNG diagnosis for the same systems.

**Keywords:** Galaxy Cluster. Computational simulation. Statistical methods. Maximum likelihood.

# Sumário

<b>1 – INTRODUÇÃO</b>	.	.	.	.	.	<b>1</b>
1.1	Contextualização	.	.	.	.	1
1.2	Composição dos aglomerados	.	.	.	.	2
1.2.1	Galáxias	.	.	.	.	3
1.2.2	Meio Intra-aglomerado	.	.	.	.	3
1.2.3	Matéria Escura	.	.	.	.	3
1.3	Estado de equilíbrio	.	.	.	.	4
1.3.1	Relaxação entre dois corpos	.	.	.	.	5
1.3.2	Relaxação violenta	.	.	.	.	5
1.4	Justificativa	.	.	.	.	5
1.5	Objetivos	.	.	.	.	6
1.5.1	Objetivos específicos	.	.	.	.	6
1.6	Estrutura do trabalho	.	.	.	.	7
<b>2 – DADOS</b>	.	.	.	.	.	<b>8</b>
2.1	Simulação e dados do projeto Millennium	.	.	.	.	8
2.2	Simulação e dados do projeto EAGLE	.	.	.	.	9
<b>3 – MÉTODOS</b>	.	.	.	.	.	<b>12</b>
3.1	O ambiente estatístico R	.	.	.	.	12
3.2	Métodos	.	.	.	.	12
3.2.1	Método da rejeição	.	.	.	.	12
3.2.2	Método da máxima verossimilhança	.	.	.	.	15
3.2.3	Método do chi-quadrado	.	.	.	.	17
3.2.4	Método do bootstrap	.	.	.	.	17
3.3	Ferramentas	.	.	.	.	17
3.3.1	Conversão de coordenadas	.	.	.	.	17
3.3.2	Medindo desvios de gaussianidade	.	.	.	.	21
3.3.3	Método dos L-Momentos	.	.	.	.	22
3.3.4	Distribuição q-Gaussianiana	.	.	.	.	25
3.3.5	Perfis de densidade	.	.	.	.	28
3.3.5.1	Perfil de NFW	.	.	.	.	28
3.3.5.2	Perfil de Hernquist	.	.	.	.	29
3.3.5.3	Perfil de King	.	.	.	.	30
3.3.5.4	Ajuste dos perfis de densidade	.	.	.	.	30

<b>4 – RESULTADOS E DISCUSSÕES</b>	33
4.1 Resultados do Millennium	34
4.1.1 Evolução dos aglomerados com o redshift	34
4.1.2 Perfis de densidade superficial	38
4.1.3 Estimação da q-Gaussiana	38
4.2 Resultados do EAGLE	41
<b>5 – CONCLUSÕES E PERSPECTIVAS</b>	52
5.1 Síntese e conclusões	53
5.2 Perspectivas	54
<b>Referências</b>	56
<b>Apêndices</b>	60
<b>APÊNDICE A – Distribuição q-Gaussiana</b>	61
<b>APÊNDICE B – Método Brent</b>	62
<b>APÊNDICE C – Principais pacotes utilizados</b>	65
<b>APÊNDICE D – Termos astronômicos</b>	66

# 1 INTRODUÇÃO

## 1.1 Contextualização

A matéria existente no Universo não possui uma distribuição homogênea, se organizando em uma rede de filamentos com comprimento de centenas de Mpc<sup>1</sup> (Mega parsec) (LIMOUSIN et al., 2013). Cerca de 90% das galáxias são encontradas nestes filamentos, sendo que em seus nós localizam-se os aglomerados de galáxias massivos (vide Figura 1). Com massas típicas da ordem de  $10^{14}$  até  $10^{15} M_{\odot}$  e apresentando picos de densidade de matéria em grandes escalas, os aglomerados de galáxias são as maiores estruturas próximas do equilíbrio no Universo (NETO et al., 2014), sendo ótimos laboratórios para o estudo da evolução das estruturas em grande escala. Grupos de galáxias comumente contém algumas dezenas de objetos dentro de um raio de 0.5 a 1.0 Mpc, enquanto aglomerados são sistemas maiores, podendo conter centenas ou até milhares de objetos concentrados em um raio característico entre 1.0 e 2.0 Mpc (VOIT, 2005).

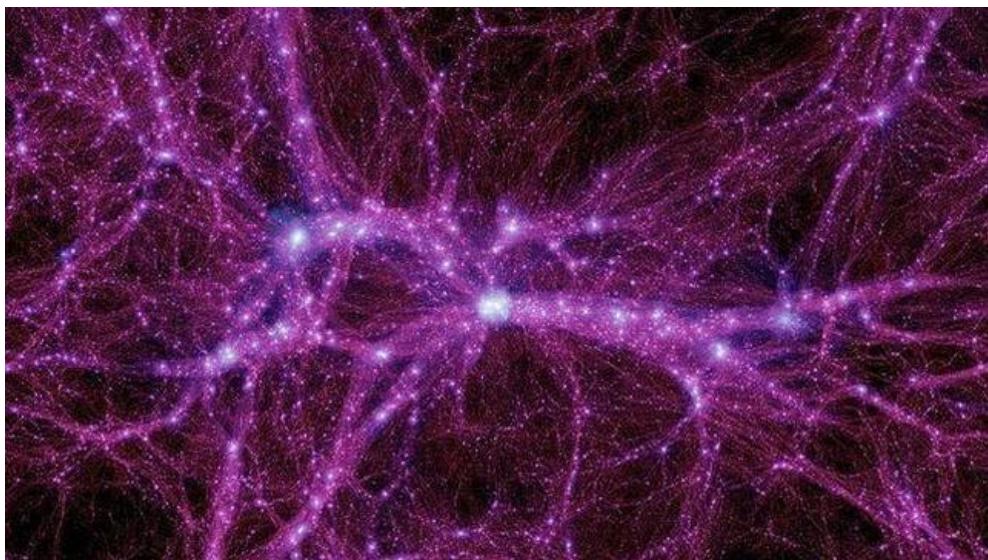


Figura 1 – Representação da distribuição de matéria escura no Universo. Os nós apresentam uma maior concentração de matéria, onde estão localizados os aglomerados de galáxias mais massivos. Fonte: (Millennium Simulation)

As primeiras indicações da existência de aglomerados de galáxias foram feitas por Charles Messier e William Herschel, que publicaram os primeiros catálogos desses objetos. Edwin Hubble revolucionou a astrofísica na década de 1920 ao mostrar que existem galáxias espirais e elípticas localizadas a grandes distâncias de nós, o que clarificou

---

<sup>1</sup>1.0 Mpc equivale a  $3.086 \times 10^{22}$  metros

o fato de que aglomerados de galáxias são estruturas físicas reais que possuem grandes dimensões. Em 1933, Fritz Zwicky encontrou as primeiras evidências da existência de matéria escura, ao estudar a dispersão de velocidades das galáxias no aglomerado de Coma. Ele percebeu que a dispersão estimada implicava em uma massa muito superior a soma da massa contida em estrelas do aglomerado, o que lhe fez concluir que existia muito mais matéria do que podia se observar diretamente ([KRAVTSOV; BORGANI, 2012](#)).

A formação dos aglomerados de galáxias é geralmente descrita dentro do cenário hierárquico de formação de estruturas. Nesse cenário, o *halo* de matéria escura se forma primeiro, originado por flutuações gravitacionais existentes no Universo primordial; em seguida ocorre a acreção de pequenas estruturas que vão se juntando para formar sistemas cada vez maiores ([VOIT, 2005](#)). Os aglomerados de galáxias são estruturas jovens, sendo as últimas das estruturas colapsadas a se formar ([KRAVTSOV; BORGANI, 2012](#)), muitas vezes são considerados estruturas em formação, tendo em vista que ainda existe acreção de matéria oriunda dos filamentos que os envolvem ([NETO et al., 2014](#)).

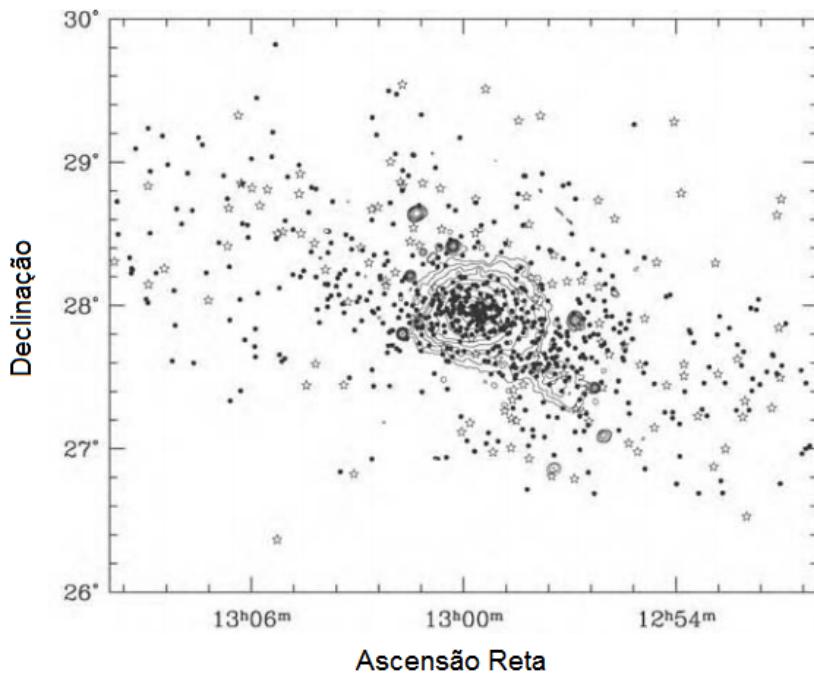


Figura 2 – O aglomerado de Coma projetado no plano do céu: os pontos sólidos representam galáxias elípticas; estrelas vazadas são galáxias espirais; Os contornos representam a intensidade da emissão de Raios-X. Fonte: ([SPARKE; III, 2007](#)).

## 1.2 Composição dos aglomerados

Na composição dos aglomerados, gás e galáxias correspondem a cerca de 20% do total de massa, sendo o restante preenchido pela matéria escura que, apesar de

não poder ser diretamente observada, tem a presença deduzida pelos seus efeitos gravitacionais (REMBOLD, 2007; NETO et al., 2014).

### 1.2.1 Galáxias

A distribuição das galáxias nos aglomerados varia fortemente, podendo ser descrita em aglomerados mais ricos através do perfil de densidade numérica projetada, usualmente modelada por uma esférica isotérmica, assemelhando-se aos aglomerados estelares (REMBOLD, 2007). Predominam nos aglomerados populações de galáxias elípticas e lenticulares, em oposição ao campo onde encontra-se uma maior quantidade de galáxias espirais (NETO et al., 2014). Tal fato se deve aos processos físicos, de natureza ambiental, que ocorrem dentro dos aglomerados (GUNN; GOTTFRIED, 1972; CAVA et al., 2017). As galáxias seguem o seguinte esquema de evolução: acompanham o fluxo de Hubble, ou seja, a expansão do Universo; desacoplam-se dele, atraídos pela gravidade do aglomerado; atingem o raio de virada, onde viram-se em direção ao aglomerado; colapsam, como uma "chuva" de galáxias para dentro do halo de matéria escura, no chamado *infall*; e, por fim, possivelmente virializam (GUNN; GOTTFRIED, 1972).

### 1.2.2 Meio Intra-aglomerado

Uma das mais importantes formas de obter dados de aglomerados é através das observações em raios-X (REMBOLD, 2007), tendo como principal fonte de emissão o gás ionizado encontrado no meio intra-aglomerado (NETO et al., 2014), o que permite a observação de objetos em redshift  $z > 1$  (KRAVTSOV; BORGANI, 2012). Composto principalmente de hidrogênio e hélio sintetizados logo após o Big Bang, é enriquecido com o material ejetado pelas explosões de supernovas, produzindo os metais (silício, níquel, ferro e cálcio) que podem ser detectados pela emissão de raios-x. Possui temperatura no intervalo entre  $10^7$  e  $10^8$  K e densidade na ordem de 1 partícula a cada  $100\text{cm}^3$ .

### 1.2.3 Matéria Escura

A quantidade de matéria escura no Universo é um problema que vem recebendo bastante atenção desde o trabalho de Zwicky (1933). Alguns modelos são propostos para tentar explicar as observações realizadas. Nos modelos cosmológicos mais aceitos, a matéria escura é tratada como uma componente acolisional e de natureza não-bariônica, ou seja, não é formada por prótons e nêutrons. Por compor cerca de 80% da massa do aglomerado, a matéria escura domina o potencial gravitacional do sistema. Apesar de não poder ser diretamente observada, diversas evidências apontam para a sua existência, como o *Bullet Cluster* (vide Figura 3), as curvas de rotação das galáxias, e os efeitos de lentes gravitacionais (CLOWE et al., 2006; RUBIN; FORD JR., 1970; MANDELBAUM

[et al., 2010](#)). Sabe-se, então, que a matéria escura cria um *halo*, espécie de envelope invisível, que envolve toda a matéria bariônica em galáxias e aglomerados.

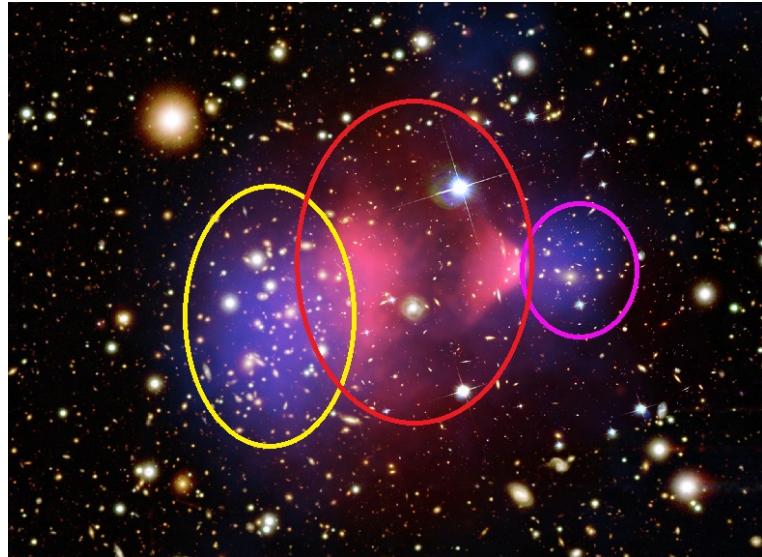


Figura 3 – Bullet Cluster: colisão entre dois aglomerados (destacados pelas elipses amarela e magenta). As galáxias que deveriam seguir o gás, componente bariônico mais massivo (elipse vermelha), seguem um componente invisível.

### 1.3 Estado de equilíbrio

Um dos principais problemas ao estudar a dinâmica dos aglomerados é determinar o quão próximo do estado de equilíbrio ele se encontra. Esse problema é de natureza complexa, tendo em vista que um aglomerado não se encontra em geral isolado no Universo, o que faz com que perturbações referentes ao encontro gravitacional com outras estruturas possam modificar constantemente a dinâmica do sistema. Uma das formas de caracterizar o estado dinâmico do aglomerado é através da sua distribuição de velocidades na linha-de-visada (VLOS), uma análise que vem sendo aplicada sistematicamente pelo menos desde o trabalho pioneiro de [Yahil e Vidal \(1977\)](#). Estudos teóricos e simulações numéricas de N-corpos indicam que uma função de distribuição Maxwell-Boltzmann descreve bem a VLOS de um sistema virializado ([Bernaldo et al., 2013a; LYNDEN-BELL, 1967; Ogorodnikov, 1957; WILLIAMS et al., 2010](#)). Em um sistema relaxado, as galáxias não possuem mais a memória das condições iniciais. Se um sistema é dito virializado, não existem movimentos sistemáticos de expansão ou contração, e as galáxias seguem o teorema do virial ([NATAJARAN et al., 1997](#)) que afirma que em um sistema em equilíbrio, a energia potencial mais o dobro da energia cinética é igual a zero. A seguir serão descritos alguns processos que podem levar o aglomerado a encontrar uma configuração de equilíbrio.

### 1.3.1 Relaxação entre dois corpos

Apesar de ser considerada um componente acolisional, as galáxias podem interagirumas com as outras através dos seus campos gravitacionais, fazendo com que mudem de direção e parte da energia cinética seja dissipada durante o processo. Caso o tempo de transferência de energia entre as galáxias fosse menor que a idade do sistema, esse processo poderia levar os aglomerados a um estado de equilíbrio. Entretanto, estimativas analíticas e simulações de N-corpos mostraram que, através da relaxação entre dois corpos, o tempo total para atingir o estágio de equilíbrio seria muito maior do que a idade do Universo, evidenciando que um outro processo foi o responsável pela configuração observada na maioria dos aglomerados de galáxias ([SARAZIN, 1986](#)).

### 1.3.2 Relaxação violenta

[LYNDEN-BELL \(1967\)](#) propôs que o estado de equilíbrio fosse produzido por efeitos gravitacionais coletivos, devido a não-homogeneidades em pequenas escalas geradas por flutuações no campo gravitacional do aglomerado. Tal processo, conhecido como relaxação violenta, poderia muito rapidamente trazer o aglomerado para uma configuração de quase-equilíbrio, fazendo com que as galáxias viessem a ter, em média, uma distribuição espacial e de velocidades aproximadamente consistente com uma esfera isotérmica. Diversas simulações numéricas confirmaram que esse processo pode levar a uma configuração próxima à observada nos aglomerados de galáxias ([SPARKE; III, 2007; SARAZIN, 1986](#)).

## 1.4 Justificativa

O problema central da construção de modelos dinâmicos para aglomerados está em escolher uma forma adequada para a função de distribuição de probabilidade  $f(E, J)$ , onde  $E$  é a energia total de uma partícula do sistema, e  $J$  seu momento angular. Uma vez que a relaxação de dois corpos não é capaz de dominar a dinâmica de todo o aglomerado, admite-se que a função de distribuição deve refletir o processo de relaxação violenta ([LYNDEN-BELL, 1967](#)), que promoveria uma randomização das velocidades e produziria, ao final do processo, uma distribuição do tipo Maxwell-Boltzmann, sem traços de segregação de massa, na região central do aglomerado. Ao término do período de relaxação violenta, o sistema evolui governado pela interação entre dois corpos, quando objetos mais massivos migram em direção ao centro do sistema, onde se encontram com diversas outras galáxias em um espaço reduzido, algo que pode levar a uma microfísica local, com efeitos sobre as órbitas desses objetos. Ao mesmo tempo, galáxias continuariam sendo atraídas para o aglomerado através de um fenômeno conhecido como *infall* secundário ([ROMERO; LAMBAS, 2012](#)). Este

*infall* produziria uma espécie de *halo* de baixa densidade em torno da região central, tornando ainda mais complexa a modelagem do equilíbrio de aglomerados. Dadas as dificuldades de se restringir as condições de contorno durante (e após) a relaxação violenta, a abordagem tradicional do problema é feita através de várias possíveis funções de distribuição que descrevem os perfis observados das galáxias nos espaços de posição e velocidade. São os chamados perfis de densidade e de distribuição de velocidades, assim como o perfil combinado destes dois observáveis, formando uma espécie de perfil 2D no espaço de fase projetado ([LIMOUSIN et al., 2013](#)).

## 1.5 Objetivos

Este trabalho tem como principal objetivo a modelagem da dinâmica dos aglomerados de galáxias a partir da implementação, em linguagem R, de códigos de ajuste e diagnóstico dos perfis de densidade e distribuição de velocidades, com o intuito de caracterizar, da forma mais ampla possível, o comportamento das galáxias e determinar o seu estágio de aproximação ao equilíbrio.

### 1.5.1 Objetivos específicos

Inicialmente serão desenvolvidos *scripts* para a automatizar a coleta de dados, tanto observacionais quanto de simulações, nos catálogos disponíveis. Para aplicação nos dados simulados, uma ferramenta de projeção deverá ser desenvolvida, tendo em vista que as simulações nos fornecem estruturas tridimensionais, enquanto os dados observados apresentam projeções desses objetos na linha-de-visada (bidimensional) e no plano do céu.

Dado que os perfis de densidade e a distribuição de velocidades podem ser representados por funções de densidade de probabilidade, desenvolveremos rotinas para a realização de variáveis aleatórias que sigam tais distribuições, tendo em vista que seja possível simular de maneira controlada seus parâmetros e comparar a simulação com os resultados obtidos a partir dos dados reais. Além disso, a realização de testes nas rotinas implementadas faz-se indispensável, de modo que explorar métodos numéricos e estatísticos para a estimativa de parâmetros é a melhor forma de verificar a validade dessas ferramentas, além de fornecer um arcabouço necessário para a caracterização dos modelos.

Também examinaremos como a gaussianidade da VLOS se relaciona aos perfis de densidade e à distribuição de galáxias no espaço de fase de aglomerados, visto que mudanças na configuração das distribuições podem promover desvios da gaussianidade do sistema.

## 1.6 Estrutura do trabalho

Este trabalho está estruturado da seguinte forma: o Capítulo 2 trará as informações sobre os dados e como foram obtidos dos catálogos e repositórios de dados; no Capítulo 3 os métodos que foram desenvolvidos para realização das análises serão descritos; o Capítulo 4 apresentará os resultados obtidos e as suas implicações; e finalmente, o Capítulo 5 abordará os objetivos alcançados e as perspectivas para trabalhos futuros.

## 2 DADOS

Tendo como finalidade acompanhar a evolução dos aglomerados em diferentes épocas, utilizamos amostras provenientes de simulações computacionais, pois estas se mostram ótimos laboratórios para o entendimento da dinâmica dos aglomerados. Foram utilizados dados obtidos de duas das mais conceituadas simulações cosmológicas existentes, são elas o EAGLE e o Millennium. As simulações de N-corpos são soluções numéricas que resolvem o problema da interação de diversas partículas interagindo gravitacionalmente, algo impraticável de maneira analítica. Essas soluções são amplamente difundidas na astrofísica, com aplicações desde sistemas com alguns poucos corpos, como o sistema solar, até sistemas com milhares de objetos, como galáxias e aglomerados de galáxias, sendo utilizadas como forma de testar modelos e guiar as observações. Nas seções a seguir, serão detalhadas as características das simulações que foram utilizadas e o método aplicado para extração de dados.

### 2.1 Simulação e dados do projeto Millennium

A *Millennium Simulation* (Simulação do Milênio) é uma simulação de alta resolução, elaborada pelo Virgo Consortium, que utiliza  $N = 2160^3$  partículas para simular a distribuição da matéria escura no Universo, em um cubo com  $500h^{-1}Mpc$  de aresta, onde  $h$  é a constante de Hubble com unidades de  $100kms^{-1}Mpc^{-1}$ . Os parâmetros cosmológicos da simulação são  $\Omega_m = \Omega_{dm} + \Omega_b = 0.25$ ,  $\Omega_b = 0.045$  e  $\Omega_\Lambda = 0.75$ . Seus dados estão dispostos em diversos *snapshots* que variam do *redshift*  $z = 127$  até  $z = 0$  ([Springel et al., 2005](#)). A simulação utiliza sofisticados modelos de formação de estruturas, como o descrito por [De Lucia e Blaizot \(2007\)](#), em uma cosmologia  $\Lambda CDM$  ( $\Lambda$  Cold Dark Matter) com resolução de massa de  $8.6 \times 10^8 h^{-1} M_\odot$ , onde  $M_\odot$  é a massa solar. Uma de suas principais características é a possibilidade de se acompanhar a evolução de objetos observados em diferentes épocas ([Springel et al., 2005](#)). O Millennium apresenta cerca de 49.6% de todas as galáxias incluídas dentro de aglomerados com até 20 objetos. Para nosso propósito, foram selecionados os aglomerados contidos nos redshifts  $z \in \{0.0, 0.5, 1.0\}$ , com massa maior que  $1 \times 10^{13} M_\odot$ , magnitude na banda  $r < 19.0$ , e com mais de 20 objetos dentro de  $R_{200}$  (raio onde a densidade do aglomerado equivale a 200 vezes a densidade crítica do Universo).

Para a coleta dos dados, o Millennium fornece uma interface<sup>1</sup> que possibilita aplicar diversas funções de seleção, utilizando a linguagem SQL (Structured Query Language). A linguagem SQL é a linguagem padrão utilizada em banco de dados rela-

---

<sup>1</sup>Disponível em: <http://gavo.mpa-garching.mpg.de/Millennium/>

cionais, abrangendo instruções para a realização de consultas, atualizações e definições de dados (Elmasri; NAVATHE, 1999). A consulta foi realizada nas tabelas referentes ao modelo semi-analítico de formação de galáxias descrito por De Lucia e Blaizot (2007). O script 2.1 apresenta um exemplo de consulta, em SQL, que retorna as coordenadas de posição e velocidade de todas as galáxias em aglomerados com massa média, em  $R_{200}$ , superior a  $7.3 \times 10^{12} M_{\odot}$ , ordenando os resultados pelo  $R_{200}$  dos aglomerados.

```
SELECT h.fofid , g.x, g.y, g.z, g.velX, g.velY, g.velZ
FROM MPAGalaxies..DeLucia2006a g, MField..FOF h
WHERE h.m_Mean200 > 0.73*1000
AND g.fofId = h.fofId
ORDER BY h.r_Mean200
```

Script 2.1 – Exemplo de consulta SQL realizada na simulação do Millennium

## 2.2 Simulação e dados do projeto EAGLE

O EAGLE (Evolution and Assembly of GaLaxies and their Environments) é um conjunto com diversas simulações hidrodinâmicas que, assim como o Millennium, é mantida pelo Virgo Consortium. Utiliza um padrão de Universo  $\Lambda$ CDM, sendo calibrado para combinar com as observações da função de massa estelar em redshift  $z = 0$ . Cada uma das suas simulações apresenta características diferentes, desde volume e resolução até modelos distintos de formação de galáxias. A simulação foi desenvolvida utilizando uma versão modificada do código GADGET-2, um conjunto de métodos e técnicas que possibilitam a criação de simulações hidrodinâmicas de N-corpos (Springel, 2005b). A cosmologia da simulação é dada pelos parâmetros  $\Omega_m = 0.307$ ,  $\Omega_b = 0.04825$ ,  $\Omega_{\Lambda} = 0.693$  e  $h = H_0/(100kms^{-1}Mpc^{-1})$ . O repositório de dados do EAGLE oferece dados em 29 *snapshots* entre o redshift  $z = 20$  e  $z = 0$ , possibilitando que objetos sejam acompanhados desde a sua formação até a época atual. O repositório disponibiliza uma grande quantidade de dados, sendo possível observar e fazer o *download* até de imagens de galáxias simuladas, como mostrado na Figura 4 (Schaye et al., 2015; CRAIN et al., 2015).

Similar ao Millennium, os dados podem ser obtidos através de uma interface<sup>2</sup> que permite a realização de consultas utilizando a linguagem SQL. As tabelas estão organizadas por grupos, onde cada grupo representa uma característica diferente, como volume, quantidade de partículas ou modelo físico. Foi realizado um mapeamento das tabelas com o objetivo de construir um modelo lógico que melhor representasse a estrutura dos dados. O processo consistiu em buscar relacionamentos entre as tabelas e remontar as ligações lógicas. Na Figura 5, as tabelas que estão representadas

<sup>2</sup>Disponível em: <<http://icc.dur.ac.uk/Eagle/database.php>>



Figura 4 – Imagem gerada para uma galáxia simulada no EAGLE.

do lado "1"(Um) de algum relacionamento (representado pela linha) são referenciadas em campos das tabelas que estão do lado "Estrela"(Muitos). Dessa forma, podemos identificar, por exemplo, todas as galáxias (*SubHalo*) de um determinado aglomerado (*FOF*). Para mais detalhes sobre o modelo lógico de banco de dados, consulte [Elmasri e Navathe \(1999\)](#). A tabela *Snapshot* contém os dados de cada um dos snapshots da simulação, sendo a única comum para todas os grupos. Nela, também, podemos encontrar o redshift referente a cada um dos snapshots. Na tabela *FOF* encontram-se os dados dos aglomerados, enquanto a tabela *Stars* possui as informações sobre as estrelas. Em *SubHalo*, encontramos características das galáxias, como a massa, taxa de formação estelar, e coordenadas cartesianas. As tabelas *Magnitudes*, *Aperture* e *Size*, trazem outros dados complementares das galáxias.

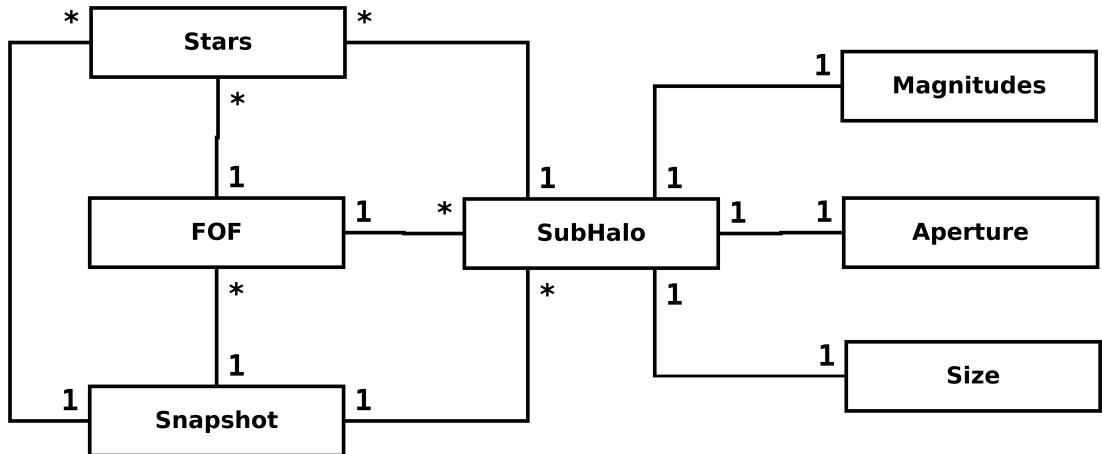


Figura 5 – Modelo lógico de banco de dados das tabelas do EAGLE.

Para este trabalho, utilizamos a simulação '*RefL100N1504*' que representa o modelo físico de referência, em uma caixa de  $100 \text{ Mpc}$  de aresta e  $1504^3$  partículas. Em seguida, selecionamos alguns aglomerados com massa estelar  $13.0 \leq \log_{10} M_{\odot} \leq 15.0$ , e

mais que 20 galáxias dentro de  $R_{200}$ . Identificamos os mesmos objetos nos redshifts  $z \in \{0.37, 0.27, 0.18, 0.10, 0.00\}$ , com o objetivo de acompanhar a evolução do aglomerado e verificar as variações de suas características no decorrer das épocas.

## 3 MÉTODOS

Neste trabalho, algumas ferramentas foram desenvolvidas com o propósito de auxiliar a modelagem dos aglomerados e realizar o tratamento dos dados. Nas seções a seguir, serão descritos os métodos utilizados para ajuste dos modelos e as ferramentas implementadas.

### 3.1 O ambiente estatístico R

Todos os programas apresentados neste trabalho foram desenvolvidos na linguagem R, e são compatíveis com a versão 3.4.0. R é um ambiente estatístico, desenvolvido originalmente por Ross Ihaka e Robert Gentleman, em 1993, e mantido, sob licença GPL (General Public License), pela R Foundation. O R possibilita que sejam realizados diversos tipos de análises estatísticas, além de fornecer poderosos recursos gráficos que permitem uma melhor exposição dos resultados. A relação dos pacotes utilizados neste trabalho pode ser encontrada no Apêndice C. Na astrofísica, o R se mostra superior a muitas outras linguagens e ambientes estatísticos por possuir pacotes especializados para análises de fenômenos astronômicos, como é o caso dos pacotes *astro* e *celestial*. Possui diversos laboratórios independentes para o desenvolvimento de pesquisa na área da astrofísica, como o *Center for Astrostatistics at Penn State*, criado em 2003, que mantém o *The Astrostatistics and Astroinformatics Portal*, uma comunidade de astrônomos e cientistas da computação que promove a difusão do R com o desenvolvimento de diversas ferramentas estatísticas, e fornece um espaço para interação entre pesquisadores de diversos lugares do mundo. Um marco no uso do R em astrofísica foi a publicação, em 2012, do livro "Modern Statistical Methods for Astronomy: With R Applications", de E.D. Feigelson e G.J. Babu, que vem se tornando livro-texto de cursos de astroestatística em todo o mundo.

### 3.2 Métodos

#### 3.2.1 Método da rejeição

Muitas das ferramentas que serão apresentadas nas próximas seções necessitam da geração de amostras que sigam uma determinada distribuição, para que seja realizada a calibração dos métodos e validação da obtenção dos parâmetros estimados a partir dos dados reais ou simulados. Para desenvolver esses geradores, utilizou-se o método da rejeição (ou rejeição-aceitação), que consiste em um procedimento amplamente

utilizado para a simulação de variáveis aleatórias (THOMAS; LUK, 2007).

Considere as funções de distribuição de probabilidade  $F(x)$  e  $G(x)$ . Deseja-se gerar valores com densidade  $F$ , mas é possível apenas gerá-los com a distribuição  $G$ . Se  $X$  é o espaço amostral para as duas distribuições, temos como restrição que

$$F(x) \leq G(x) \quad (1)$$

para todo  $x \in X$ . Para garantir que a restrição imposta pela Equação 1 seja satisfeita, podemos maximizar a função  $\frac{F(x)}{G(x)}$  e obter uma constante  $c$ , de modo que  $c \times G(x)$  funcione como um envelope para  $F(x)$  em todo o espaço amostral.

Uma maneira mais simples de se implementar o método da rejeição é através da utilização da distribuição uniforme  $U$  como envelope. O algoritmo recebe como entrada a função de distribuição desejada e os limites do espaço amostral. Em seguida, procede-se com a maximização de  $\frac{F(X)}{U(X)}$  para encontrar a constante  $c$ . O próximo passo consiste em gerar um par ordenado de valores  $(x, p)$ , onde  $x$  corresponde a uma observação pertencente ao espaço amostral com distribuição  $U(X)$  e  $p$  uma probabilidade com distribuição  $c \times U(X)$ . Testa-se, então, se a probabilidade  $p$  é menor ou igual ao valor de  $F(x)$ . Em caso verdadeiro, o valor de  $x$  é aceito e o algoritmo termina; caso contrário, a execução retorna ao passo da geração de  $x$  e as etapas seguintes devem ser novamente executadas. Uma visão mais didática do algoritmo é encontrada na Figura 6.

Como exemplo, vamos gerar valores que seguem a distribuição gaussiana a partir da distribuição uniforme. Sabemos que a distribuição gaussiana (ou normal) tem função de densidade

$$n(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2},$$

onde  $\mu$  é a média e  $\sigma$  o desvio padrão. Com o auxílio da distribuição uniforme, que tem como função de densidade

$$u(x) = \frac{1}{x_b - x_a},$$

onde  $x_a$  e  $x_b$  são os limites inferior e superior do espaço amostral, respectivamente, podemos encontrar uma constante  $c$ , no intervalo  $x \in [-5, 5]$ , que maximize  $n(x)/u(x)$ , com  $\mu = 0$  e  $\sigma = 1$  sendo os parâmetros de  $n(x)$ , e obtemos o valor  $c = 3.98942$ . São gerados oito pontos com probabilidade  $c \times u(x)$ . Vemos na Figura 7 que alguns pontos (em vermelho) são rejeitados por não terem probabilidade menor que  $n(x)$ , e outros (em verde) são aceitos por satisfazerem a condição. Podemos notar que a escolha da função de distribuição favorece, ou não, a quantidade de pontos que serão aceitos pelo

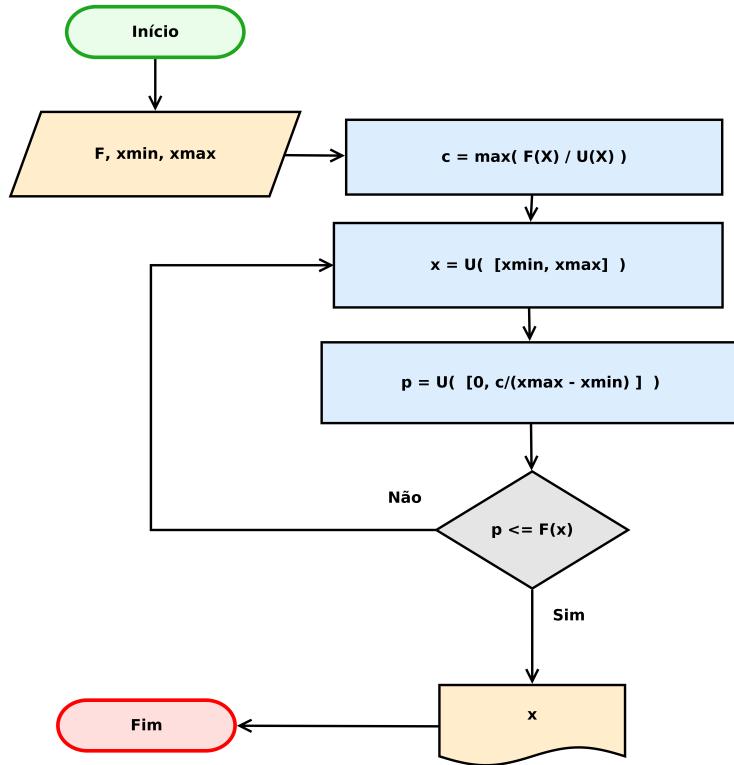


Figura 6 – Algoritmo do método da rejeição utilizando distribuição uniforme.

algoritmo, tendo sua eficiência computacional diretamente associada. Neste trabalho não nos preocupamos com esse detalhe, muito embora seja algo que deva ser apreciado em versões futuras.

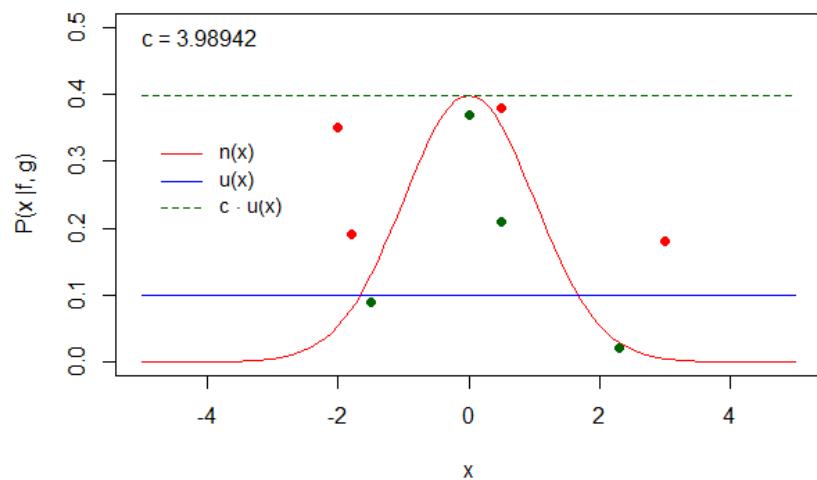


Figura 7 – Exemplo do funcionamento do método da rejeição para simulação da distribuição normal. Os pontos em vermelho representam os valores rejeitados, enquanto os pontos em verde correspondem aos valores aceitos.

### 3.2.2 Método da máxima verossimilhança

A máxima verossimilhança é um método estatístico de estimação que consiste em encontrar o melhor conjunto de parâmetros de um dado modelo para uma determinada amostra. Dada uma distribuição de probabilidade e uma amostra que a siga, utilizam-se todas as observações da amostra para calcular o conjunto de parâmetros que pode ter gerado tais valores. Em outras palavras, sendo  $x_1, x_2, x_3, \dots, x_n$  observações de uma variável aleatória  $X$ , que tiveram origem em uma distribuição de probabilidade  $f(x, \theta)$ , onde  $\theta$  é o conjunto de parâmetros da distribuição, temos

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta). \quad (2)$$

Ao maximizarmos a Equação 2, encontraremos o conjunto de parâmetros mais adequado para gerar os valores da amostra, denotado por  $\bar{\theta}$ . Note que quando  $n \rightarrow \infty$ ,  $\bar{\theta} \rightarrow \theta$ , ou seja, quanto maior o tamanho da amostra, mais precisa é a estimativa.

É mais conveniente trabalhar com o *log* negativo da função de verossimilhança, uma vez que a medida que o tamanho da amostra cresce, o produtório pode gerar valores muito próximos de zero, causando um problema computacional de precisão finita. Assim, temos

$$-\log L(\theta) = -\sum_{i=1}^n \log(f(x_i, \theta)) \quad (3)$$

como função log-verossimilhança negativa. Diferentemente da Equação 2, que é maximizada, ao minimizarmos a Equação 3, encontramos o conjunto de parâmetros  $\bar{\theta}$ .

No R, podemos encontrar algumas funções que trabalham com o método da máxima verossimilhança, como é o caso da função *mle2* contida no pacote *bbmle*. Agora, apresentaremos um exemplo de construção da função de verossimilhança para uma distribuição gaussiana (normal). Em seguida, estimaremos o conjunto de parâmetros para uma amostra simulada. Inicialmente, vamos gerar uma amostra de 500 elementos, com média  $\mu = 3.5$  e desvio padrão  $\sigma = 1.4$ , e criar a função log-verossimilhança negativa.

```
X = rnorm(n=500, mean=3.5, sd=1.4)

minuslog.normal <- function(theta_m, theta_s)
{
  return( - sum( dnorm(x=X, mean=theta_m, sd=theta_s, log=T) ) )
}
```

Depois de criada a função log-verossimilhança negativa, utilizaremos o procedimento *mle2* do pacote *bbmle* para recuperarmos o conjunto de parâmetros que foram utilizados para simular a amostra. O procedimento tem como entrada o nome definido para a função de log-verossimilhança negativa, valores iniciais para os parâmetros e o método utilizado para minimização. Após obtenção do resultado, podemos ver que os parâmetros se aproximam bastante do valor para o qual a amostra foi simulada.

```
mle2(minuslogl = minuslog.normal, start = list(theta_m = 0.5, theta_s = 0.5), method="Nelder-Mead")

Coefficients:
theta_m  theta_s
3.517491 1.421592

Log-likelihood: -885.5
```

Além dos parâmetros, a função ainda retorna o valor minimizado da função de log-verossimilhança negativa. Esse valor pode ser utilizado na adoção de critérios de seleção de modelos como o *Akaike Information Criterion* (AIC) ou *Bayes Information Criterion* (BIC). Os critérios de informação são formas de avaliar qual modelo, dentre vários, melhor descreve um conjunto de dados. Para calcular os critérios de informação leva-se em conta o valor minimizado da função log-verossimilhança negativa e se penaliza a quantidade de parâmetros de cada modelo. Como podemos observar na Figura 8, o estimador de verossimilhança fornece um bom resultado para amostras com um número significativo de observações, muito embora encontrar o tamanho mínimo da amostra que torna o método da máxima verossimilhança confiável nem sempre seja trivial.

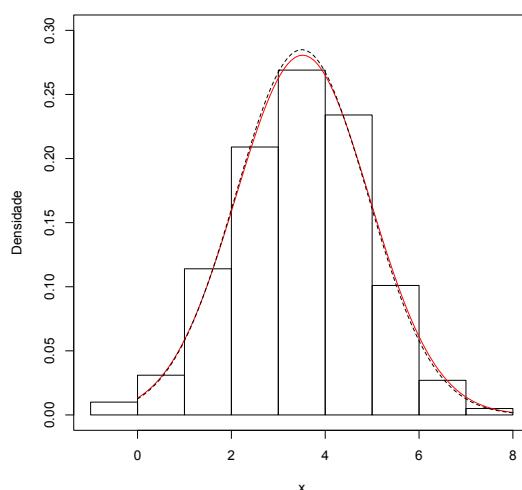


Figura 8 – Histograma representando os dados simulados a partir da distribuição normal. A linha tracejada em preto apresenta a curva para os parâmetros reais, enquanto a linha vermelha apresenta a curva para os parâmetros estimados.

### 3.2.3 Método do chi-quadrado

O método do  $\chi^2$  é um teste estatístico que destina-se a analisar a dispersão entre duas variáveis, verificando a associação entre elas. Consiste em medir a discrepância entre os dados observados e estimados através da equação

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

onde  $O$  corresponde a frequência observada para um grupo, e  $E$  a frequência esperada. Dessa forma, é possível atribuir valores para um determinado parâmetro do modelo e verificar qual dos valores apresenta o menor  $\chi^2$ , em relação aos dados observados. Se o valor observado for muito diferente do esperado, o  $\chi^2$  correspondente será muito grande; caso contrário, o valor de  $\chi^2$  será pequeno.

### 3.2.4 Método do bootstrap

Estabelecer a confiança da análise estatística apresenta-se como um procedimento tão relevante quanto encontrar os seus resultados, pois torna-se possível diferenciar um resultado ruidoso ou tendencioso de outro que realmente reflete uma estatística real. Uma das principais técnicas utilizadas para determinar o intervalo de confiança para um parâmetro de interesse é o bootstrap, que consiste na realização de diversas reamostragens, com reposição, em um conjunto, seguido da estimativa dos parâmetros do modelo para cada uma das reamostragens. Ao final do processo, calcula-se a média dos parâmetros estimados e o desvio padrão. Então, a média corresponde à melhor aproximação para o parâmetro e o desvio padrão delimita a confiança associada. Neste trabalho, utilizaremos o bootstrap tanto para encontrar o intervalo de confiança para os parâmetros estimados quanto na correção do viés para alguns métodos desenvolvidos.

## 3.3 Ferramentas

### 3.3.1 Conversão de coordenadas

Inicialmente, foi desenvolvido uma ferramenta que tem como objetivo transformar dados de objetos tridimensionais, oriundos das simulações, em bidimensionais, na forma que são encontrados em observações do céu, considerando as velocidades e posições dos objetos, em suas componentes  $x$ ,  $y$  e  $z$ . Dessa forma, foi possível comparar os resultados obtidos para dados observacionais e dados de simulação, ao projetar na linha de visada aglomerados (simulados) que possuam características previamente conhecidas.

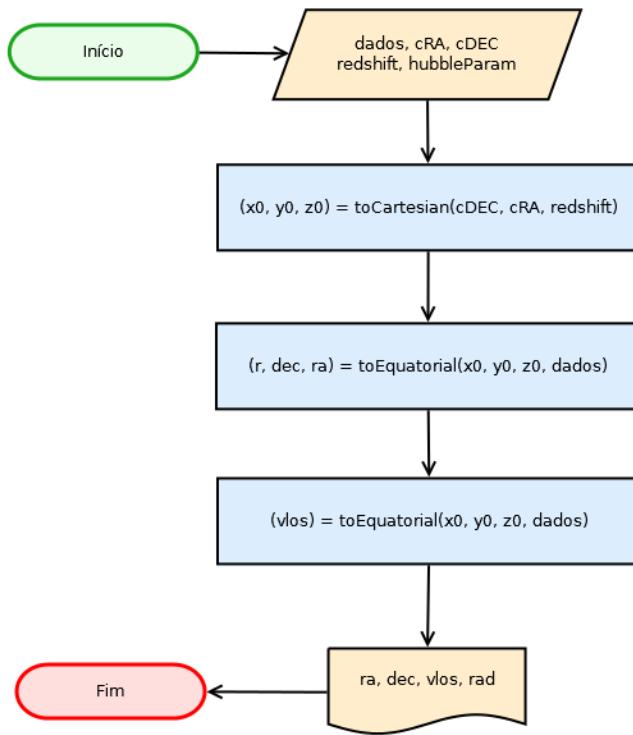


Figura 9 – Algoritmo para conversão de coordenadas cartesianas para esféricas (observacionais).

A ferramenta utiliza como dados de entrada as componentes de posição e velocidade  $x$ ,  $y$  e  $z$ , dados em  $Mpc$  e  $kms^{-1}$ , os ângulos de observação na linha de visada, em graus, o redshift, e os parâmetros que definem a cosmologia. Primeiramente, os ângulos de observação são convertidos de graus para radianos, facilitando, assim, o manuseio dos valores no programa. A posição onde o objeto será colocado foi obtida da seguinte forma:

$$x_0 = \frac{\cos(DEC_{los})\cos(RA_{los})v(z)}{H_0}$$

$$y_0 = \frac{\cos(DEC_{los})\sin(RA_{los})v(z)}{H_0}$$

$$z_0 = \frac{\sin(DEC_{los})v(z)}{H_0}$$

onde  $v(z) = \frac{cz}{1+z} \times [1 + 0.5(1 - q_0)z]$  é o produto de  $cz$  corrigido pelo redshift,  $q_0 = -0.55$ ,  $DEC_{los}$  e  $RA_{los}$  são a declinação e ascensão reta do centro <sup>1</sup>, respectivamente, e  $H_0$  é o parâmetro de Hubble na época atual. Para converter as coordenadas das galáxias,

<sup>1</sup>Ascensão reta e declinação são coordenadas angulares sobre a esfera celeste no sistema de coordenadas equatoriais celestes.

a norma do vetor entre o centro do aglomerado e a posição da galáxia é calculada da seguinte maneira:

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}.$$

Em seguida, obtemos a declinação e a ascensão reta:

$$DEC_{gal} = \arcsen\left(\frac{z - z_0}{r}\right);$$

$$RA_{gal} = \arccos\left(\frac{x - x_0}{\sqrt{(x - x_0)^2 + (y - y_0)^2}}\right).$$

Para a velocidade, temos:

$$VLOS_{gal} = \frac{v_x(x - x_0) + v_y(y - y_0) + v_z(z - z_0)}{r},$$

onde  $v_x$ ,  $v_y$  e  $v_z$  são respectivamente as velocidades nas direções  $x$ ,  $y$  e  $z$ , dadas em  $kms^{-1}$ . Podemos ver na Figura 10 o resultado da projeção de um aglomerado no plano do céu utilizando a ferramenta desenvolvida. Como exemplo, vamos supor que queiramos posicionar, no plano do céu, uma galáxia pertencente a um aglomerado simulado nas coordenadas  $(Ra, Dec) = (30, 60)$ . Dado que a galáxia possui coordenadas tridimensionais  $(x, y, z) = (5.00, 3.00, 8.00)$ , velocidade  $(v_x, v_y, v_z) = (-2.5, 9.7, 1.4)$ , e redshift 0.5, utilizando o Script 3.1, obtemos como resultado  $(Ra, Dec, Vlos) = (30.29001, 58.37281, 2.62562)$ .

```
cluster01 = data.frame(x = 5.00, y = 3.00, z = 8.00, velX = -2.5, velY =
9.7, velZ = 1.4);

result = cartesianToEquatorial(data = cluster01, losDec = 60, losRa = 30,
redshift = 0.5, hubbleParam = 70.2, omegaM = 0.272, omegaL = 0.728,
omegaK = 0);
```

Script 3.1 – Exemplo de utilização da ferramenta de conversão de coordenadas

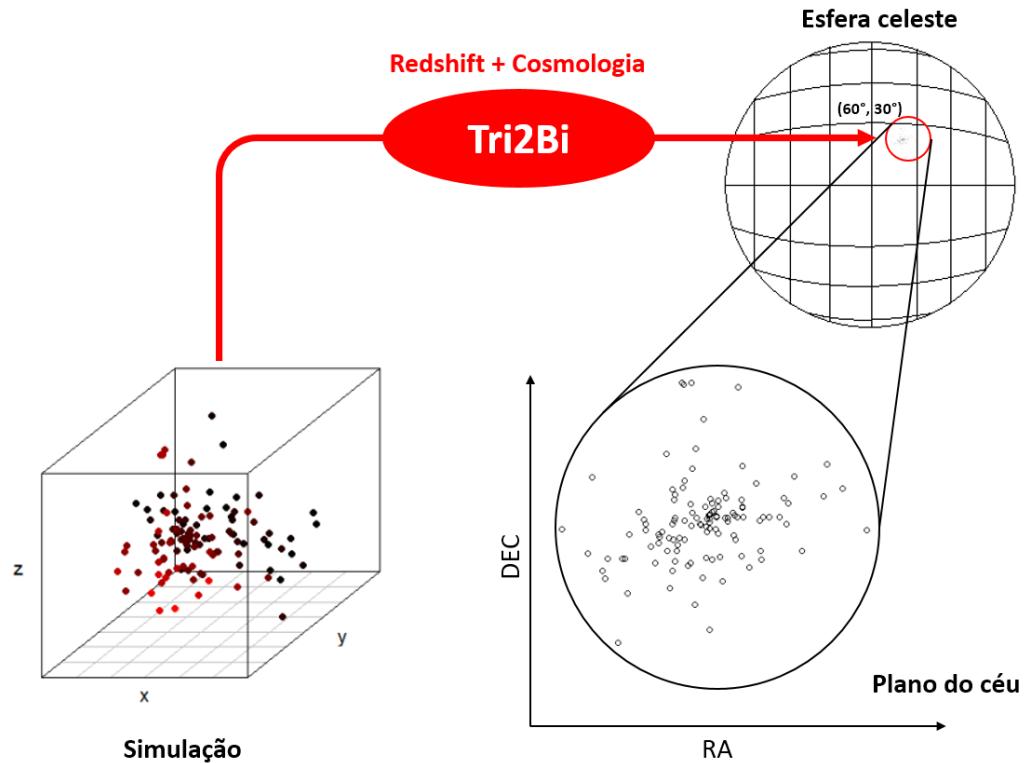


Figura 10 – Projeção de um aglomerado simulado no plano do céu utilizando a ferramenta desenvolvida.

Agora, podemos comparar o histograma de velocidades tridimensionais normalizadas com a projeção da velocidade das galáxias na linha de visada. Na Figura 11, podemos observar a grande diferença entre as duas distribuições de velocidades e, por essa razão, um método inadequado pode trazer resultados divergentes para um mesmo conjunto de dados.

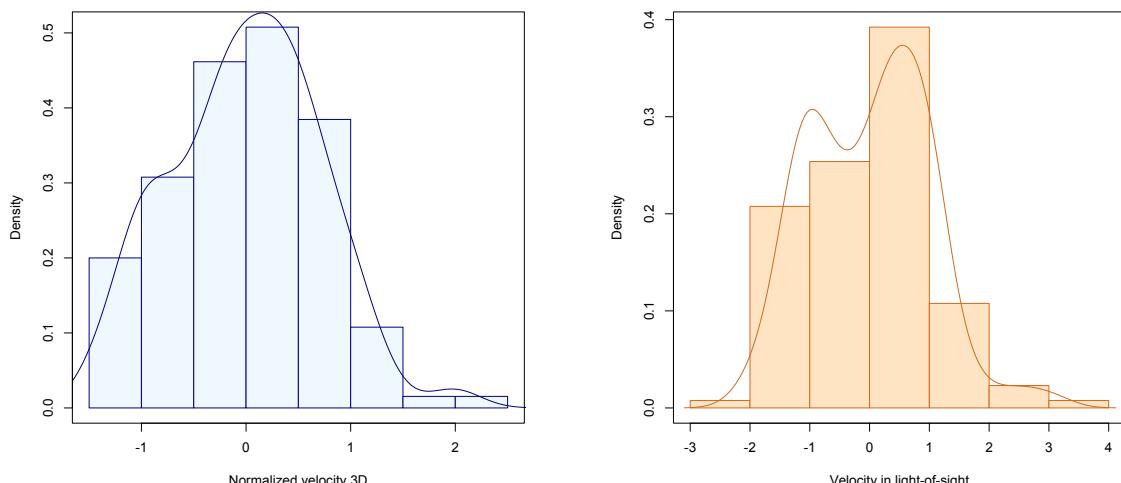


Figura 11 – Exemplo de histograma das velocidades tridimensionais das galáxias de um aglomerado (esquerda), e sua versão projetada na linha de visada (direita).

### 3.3.2 Medindo desvios de gaussianidade

Ribeiro et al. (2013) introduziram a distância de Hellinger (HD) como método para estimar a distância entre a distribuição empírica de velocidades para um aglomerado de galáxias e a distribuição Gaussiana (teórica). A distância de Hellinger, proposta originalmente por Anil Kumar Bhattacharya, é caracterizada em um espaço discreto pela função

$$HD^2(p, q) = 2 \sum_{x \in X} \left[ \sqrt{p(x)} - \sqrt{q(x)} \right]^2,$$

onde  $p$  e  $q$  são distribuições de probabilidade, empíricas ou teóricas, e  $x$  é uma variável aleatória.

A calibração do método se faz necessária, uma vez que o tamanho da amostra pode enviesar os resultados de HD. Então, para amostras de tamanho  $N$ , realizam-se 1000 (mil) reamostragens e se calcula HD para cada uma delas. Como é visto na Figura 12, ao final do processo, um viés sistemático é introduzido nos dados. Para corrigir o viés, desconta-se a mediana dos valores obtidos para HD, acrescido de  $3\sigma$ , onde  $\sigma$  consiste no desvio padrão robusto encontrado com as reamostragens.

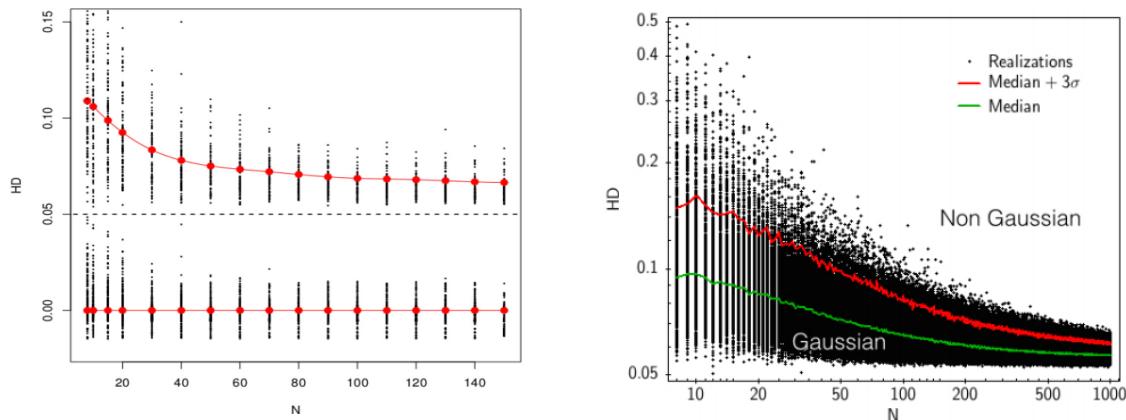


Figura 12 – Calibração inicial de HD definida em Ribeiro et al. (2013) (esquerda) e a versão atualizada proposta em de Carvalho et al. (2017) (direita).

Estudos teóricos, fenomenológicos e simulações de N-corpos sugerem que o estado de equilíbrio virializado em sistemas gravitacionais é bem descrito por uma função de distribuição do tipo Maxwell-Boltzmann, que pode ser traduzida na linha de visada (VLOS) como uma distribuição gaussiana (Beraldo et al., 2013a; WILLIAMS et al., 2010). Em dados simulados, onde a VLOS não está disponível, utiliza-se a ferramenta descrita na secção 3.3.1 para realizar a conversão de coordenadas e obter VLOS. Então, o algoritmo proposto por Ribeiro et al. (2013) e aprimorado por de Carvalho et al. (2017), chamado GNG (Gaussian Non-Gaussian), classifica os aglomerados em três grupos

distintos: G; NG; e I. O grupo G apresenta mais de 70% das reamostragens classificadas como *Gaussianas*, enquanto o grupo NG apresenta o mesmo percentual *Não-Gaussianas*. Quando nenhum dos grupos atingem mais que 70% das reamostragens, diz-se, então, que o diagnóstico do aglomerado é *Indeterminado*. É importante notar que há uma diferença nas calibrações entre os trabalhos de 2013 e 2017, sendo que no primeiro a linha de corte separando sistemas gaussianos de não-gaussianos é uma constante, enquanto no segundo a curva descresce com o aumento do número de objetos nos aglomerados.

### 3.3.3 Método dos L-Momentos

Ainda que o GNG apresente bons resultados ao identificar aglomerados gaussianos e não-gaussianos, ele introduz uma quantidade significativa de aglomerados classificados como indeterminados. Por exemplo, em ([de Carvalho et al., 2017](#)), foram estudados 319 aglomerados, mas 142 tiveram diagnóstico indeterminado, o que corresponde a 44% da amostra. Dessa forma, parte significativa da amostra teve que ser ignorada por não se ter uma definição sobre a qual grupo ela pertence. Por outro lado, a introdução de um diagnóstico indeterminado mostra uma consciência de que muitos objetos que são observados possuem uma classificação difícil, algo que em contrapartida torna os diagnósticos para grupos G e NG mais precisos.

Apresentaremos nesta seção uma possível alternativa, que pode ser combinada com o método GNG, para recuperar muitos dos objetos diagnosticados como indeterminados. Essa alternativa consiste na utilização do Método dos L-Momentos, que consegue descrever a forma de distribuições de probabilidade em termos de momentos com probabilidades ponderadas. Os L-Momentos de uma distribuição foram definidos por [Hosking e Wallis \(1997\)](#) como

$$\lambda_r = \int_0^1 x(u) P_{r-1}^*(u) du$$

onde

$$P_r^*(u) = \sum_{k=0}^r P_{r,k}^* u^k, r \in \mathbb{N}$$

e

$$P_{r,k}^* = \frac{(-1)^{r-k}(r+k)!}{(k!)^2(r-k)!}$$

sendo seus quatro primeiros L-Momentos

$$\begin{aligned}\lambda_1 &= \beta_0 \\ \lambda_2 &= 2\beta_1 - \beta_0 \\ \lambda_3 &= 6\beta_2 - 6\beta_1 + \beta_0 \\ \lambda_4 &= 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0\end{aligned}$$

onde  $\beta_r$  são seus momentos ponderados, definidos por

$$\beta_r = \int_0^1 x(u)u^r du, r \in \mathbb{N}.$$

Relacionamos os coeficientes de L-Skewness e L-Kurtosis, que correspondem respectivamente a  $\tau_3 = \frac{\lambda_3}{\lambda_2}$  (3º L-Momento) e  $\tau_4 = \frac{\lambda_4}{\lambda_2}$  (4º L-Momento), para obter o diagrama de L-Momentos. Muitas distribuições de probabilidade possuem a relação de L-Skewness e L-Kurtosis conhecida, o que torna possível utilizar o diagrama para determinar a proximidade dos objetos com uma distribuição. Para a distribuição Gaussiana, definida com  $\mu = 0$  e  $\sigma = 1$ , temos que o ponto que relaciona L-Skewness e L-Kurtosis (*Ponto Gaussiano*, daqui em diante) é dado por  $\tau_3 = 0$  e  $\tau_4 = 0.1226$ .

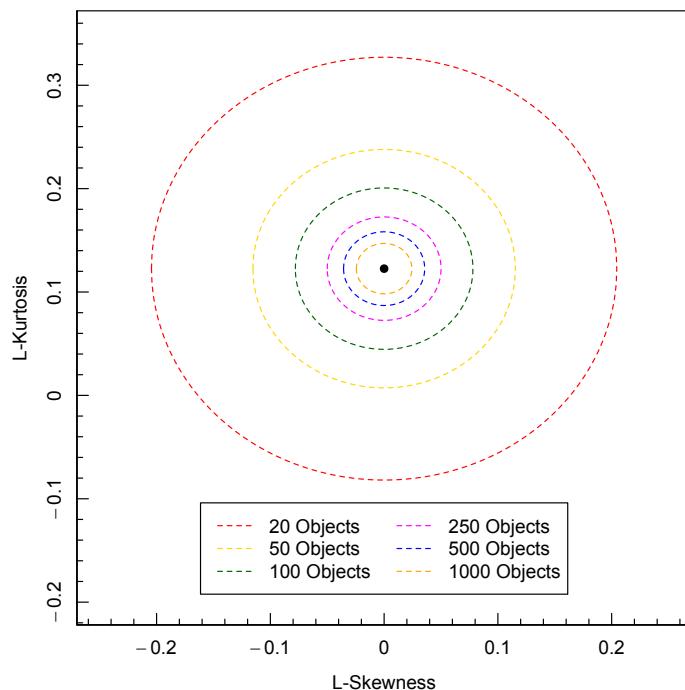


Figura 13 – Limites da região de aceitação definida em torno do Ponto Gaussiano.

Podemos obter uma região em torno do *Ponto Gaussiano* em que um objeto com L-Skewness e L-Kurtosis conhecidos seja considerado Gaussiano. Esta região é conhecida como região de aceitação. [Liou et al. \(2008\)](#) observaram que existe uma dependência entre a dimensão da região de aceitação e o número de objetos contidos na amostra. Dessa forma, para introduzirmos um indicador de gaussianidade para aglomerados classificados como indeterminados, é necessário uma análise mais aprofundada na delimitação da região de aceitação. Para definir a região de aceitação, simulamos 1000 vezes uma variável aleatória X, com distribuição gaussiana. Repetimos o processo para um número de objetos  $N = \{20, 50, 100, 250, 500, 1000\}$ . Definimos a região de aceitação como aquela delimitada pelo raio que contém 90% dos objetos, como podemos ver na Figura 13.

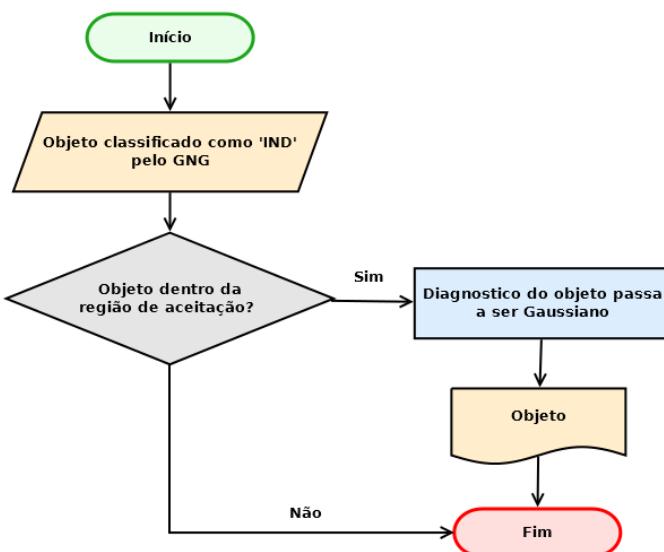


Figura 14 – Algoritmo para recuperação de aglomerados classificados como Indeterminados pelo GNG, utilizando o método dos L-Momentos.

O algoritmo aqui proposto para reclassificação de aglomerados consiste em verificar o L-Skewness e L-Kurtosis do objeto e, em seguida, verificar se ele se encontra na região de aceitação para o número de objetos que possui. Em caso afirmativo, seu diagnóstico é modificado para gaussiano. Note que esse procedimento deve ser aplicado apenas em aglomerados diagnosticados como indeterminados, tendo em vista que aglomerados classificados como não-gaussianos dentro da região de aceitação apresentam-se como casos minoritários, representando menos de 3% das amostras estudadas, e não serão abordados neste trabalho. Em R, podemos utilizar os pacotes *Lmoments* e *lmomco* para trabalhar com os L-Momentos de uma distribuição. Na Figura 14 vemos o algoritmo para realizar a recuperação de indeterminados.

### 3.3.4 Distribuição q-Gaussiana

Diversas simulações numéricas e computacionais mostram que a projeção da distribuição de Maxwell-Boltzmann no espaço de fase apresenta contundentes desvios da gaussianidade (Hansen et al., 2006). Nesse sentido, muitos estudos foram desenvolvidos com o objetivo de aplicar outras funções que melhor ajustam o comportamento da distribuição projetada de velocidades. Bernaldo et al. (2013a) propuseram uma versão anisotrópica da distribuição q-gaussiana, uma generalização da distribuição de Maxwell proposta por Tsallis (1988), que consegue um melhor ajuste para a velocidade projetada e, ademais, possui apenas um parâmetro (para mais detalhes sobre distribuições generalizadas consulte o Apêndice A). Temos,

$$F(u_x, u_y, u_z) = C_q \left[ 1 - (1-q) \frac{D_q}{2} (u_x^2 + u_y^2 + u_z^2) \right]^{\frac{1}{1-q}}$$

onde

$$u_i = \frac{v_i - \bar{v}_i}{\sigma_i}, \quad (4)$$

com  $\bar{v}_i$  e  $\sigma_i$  sendo a média e a dispersão das velocidades na direção  $i$  em coordenadas cartesianas, enquanto  $D_q$  e  $C_q$  são constantes dadas por

$$D_q = \frac{2}{7 - 5q}$$

e

$$C_q = \left( \frac{D_q}{2\pi} \right)^{3/2} \times \begin{cases} (1-q)^{3/2} \frac{\Gamma[1/(1-q)+5/2]}{\Gamma[1/(1-q)+1]}, & (0 < q < 1); \\ (1-q)^{3/2} \frac{\Gamma[1/(q-1)]}{\Gamma[1/(q-1)-3/2]}, & (1 < q < 7/5). \end{cases}$$

Definindo, também, o módulo da função de distribuição de velocidades (SLOS - Speed in line-of-sight), temos que

$$G(u) = 4\pi u^2 C_q \left[ 1 - (1-q) \frac{D_q}{2} u^2 \right]^{\frac{1}{1-q}} \quad (5)$$

onde

$$u = \sqrt{u_x^2 + u_y^2 + u_z^2}. \quad (6)$$

O domínio da Equação 5, para  $q < 1$ , é dado por  $0 \leq u \leq \sqrt{(7-5q)/(1-q)}$ , e para  $q > 1$ ,  $0 \leq u < \infty$ .

A ferramenta descrita nesta seção consiste em um algoritmo que estima o parâmetro  $q$ , em amostras de galáxias oriundas de simulações computacionais. Para isso, utilizou-se o método da máxima verossimilhança, descrito na secção 3.2.2. A função log-verossimilhança negativa é dada por

$$-\ln L(q) = - \sum \ln G(u|q). \quad (7)$$

A minimização da Equação 7 foi realizada utilizando a função *mle2* do pacote *bbmle*, com método proposto por Brent (1973), que consiste na combinação de diversos métodos computacionais para encontrar as raízes de funções, sendo aplicado para otimização de problemas unidimensionais (para mais detalhes, vide Apêndice B).

Para testar o algoritmo, foi desenvolvido um programa que utiliza o método da rejeição, descrito na secção 3.2.1, e simula números pseudo-aleatórios seguindo a distribuição q-gaussiana, como podemos ver na Figura 15. Então, foram geradas diversas amostras com valores distintos para  $q$ . Utilizou-se o método da máxima verossimilhança para recuperar os valores de  $q$ , ao minimizarmos a Equação 7. O script 3.2 exemplifica o procedimento adotado para validar o método.

```
X = qGaussian.random(n = 5000, q = 0.7)

minuslog.qgaussian <- function(theta_q)
{
  return( - sum( qGaussian.density(x=X, q=theta_q, log.p=T) ) )
}

mle2(minuslogl = minuslog.qgaussian, start = list(theta_q = 0.1),
method = "Brent", lower = 0.01, upper = 1.39)

Coefficients:
theta_q
0.7189459

Log-likelihood: -4441.7
```

Script 3.2 – Exemplo da recuperação de parâmetros simulados da distribuição q-gaussiana, utilizando o método da máxima verossimilhança.

Após a estimativa dos parâmetros, investigamos o desvio de  $q$  com relação ao número de objetos da amostra e constatamos que o viés apresenta-se significativo para valores de  $N < 250$  objetos. Desenvolvemos, então, uma rotina para corrigir tal desvio através da calibração dos valores discretos de  $q \in \{0.05, 0.10, 0.15, \dots, 0.95, 1.05, \dots, 1.35\}$ . A calibração consiste em gerar 1.000.000 de objetos com distribuição q-gaussiana para um valor de  $q$  e em seguida realizar 1000 reamostragens, com reposição, para cada  $N$

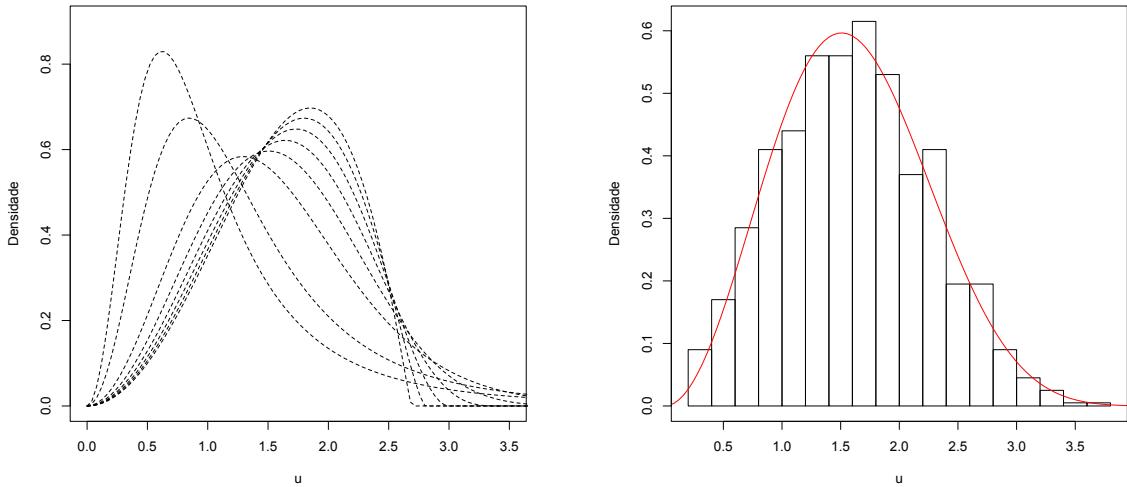


Figura 15 – Curva da distribuição  $q$ -Gaussiana para diversos valores de  $q$  (esquerda). Simulação de valores com distribuição  $q$ -Gaussiana e  $q = 0.9$  (direita).

objetos da amostra. Em cada reamostragem, estimamos o valor de  $\bar{q}_i$ . Por fim, calculamos a média das 1000 estimativas de  $\bar{q}_i$ , como vemos na equação 8.

$$q_N = \frac{\sum q_i}{1000} \quad (8)$$

Dado um  $q$  qualquer, com  $T$  objetos, para encontrar o valor corrigido  $q_c$  inicialmente arredondamos o valor de  $q$  para o seu maior valor discretizado  $q_d$ . Como os valores são subestimados, somamos ao valor de  $q$  a diferença entre  $q_N$  e  $q_d$ , e multiplicamos o resultado por  $\frac{1.25\sigma}{\sqrt{N}}$ , termo que melhor se adequou ao ajuste, sendo encontrado após a realização de exaustivos testes numéricos, como podemos ver na equação 9. Utilizamos a regressão polinomial com mínimos quadrados para ajustar o polinômio para todos os valores de  $N$ .

$$q_c = (q + |q_d - q_N|) \times \frac{1.25\sigma}{\sqrt{N}} \quad (9)$$

Ao observarmos na Figura 16, que apresenta a correção para  $q = \{0.3, 0.8, 1.3\}$ , apenas sistemas com menos que 60 objetos, podemos notar que a qualidade da correção melhora conforme o aumento do valor de  $q$ , isso faz com que a correção seja tolerável para valores de  $q \leq 0.3$  e bastante acurada para  $q \geq 0.8$ . Para sistemas com mais que 60 objetos, a correção se mostra precisa e acurada independente do valor de  $q$ .

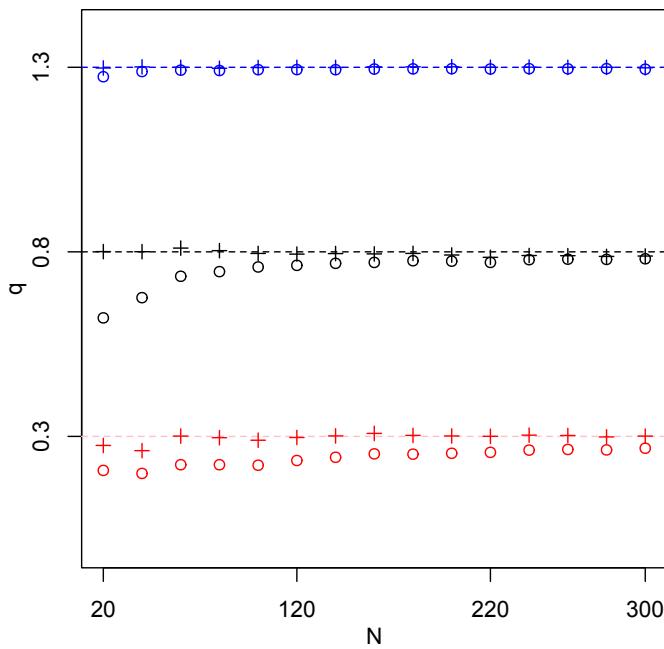


Figura 16 – Correção do viés de  $q$  em relação ao número de objetos da amostra para os valores 0.3, 0.8 e 1.3. Os círculos correspondem aos valores estimados de  $q$ , enquanto as cruzes representam os valores corrigidos.

### 3.3.5 Perfil de densidade

Conforme discutido no Capítulo 1, podemos também caracterizar um aglomerado de galáxias a partir da distribuição dos objetos no espaço, que seguem modelos característicos conhecidos como perfis de densidade. Em acordo com os resultados de simulações numéricas, e confirmado por observações, NAVARRO et al. (1996) propuseram que um sistema em equilíbrio apresenta um perfil de densidade universal, chamado Navarro-Frenk-White (NFW, daqui em diante). É importante notar que mesmo o modelo de NFW apresentando um bom ajuste para o perfil de densidade, alguns desacordos foram encontrados em sua região central (Fukushige; Makino, 2003) e, por essa razão, outros perfis existentes foram investigados.

#### 3.3.5.1 Perfil de NFW

Proposto com o intuito de ajustar dados de simulações de N-corpos, para *halos* de matéria escura, tem a seguinte forma

$$\sigma_{NFW}(r) = \frac{\sigma_0}{(r/r_s)(1+r/r_s)^2} \quad (10)$$

onde  $\sigma_0$  corresponde a densidade central e  $r_s$  o raio de escala. Possui uma forma analítica

para o perfil de densidade superficial  $\Sigma(R)$ , encontrada em [Bartelmann \(1996\)](#), dada por

$$\Sigma_{NFW}(R) = 2\sigma_0 r_s F(R/r_s) \quad (11)$$

onde

$$F(X) = \begin{cases} \frac{1}{X^2-1} \left( 1 - \frac{2}{\sqrt{1-X^2}} \operatorname{arctanh} \sqrt{\frac{1-X}{1+X}} \right), & (X < 1) \\ \frac{1}{3}, & (X = 1) \\ \frac{1}{X^2-1} \left( 1 - \frac{2}{\sqrt{X^2-1}} \operatorname{arctan} \sqrt{\frac{X-1}{X+1}} \right), & (X > 1) \end{cases}$$

Nota-se que esse perfil possui uma divergência não-física na origem, além de tender assintoticamente para zero em suas partes mais externas, o que implica em um perfil de massa total infinita ([Beraldo et al., 2013b](#)).

### 3.3.5.2 Perfil de Hernquist

O perfil de Hernquist ([HERNQUIST, 1990](#)), proposto inicialmente para galáxias esféricas e bojos, tem a seguinte forma analítica

$$\sigma_{Hernquist}(r) = \frac{\sigma_0}{(r/r_s)(1+r/r_s)^3} \quad (12)$$

diferindo do perfil de NFW apenas nas partes exteriores, onde o mesmo varia com  $r^{-4}$ . Tal modelo foi encontrado por meio de expressões analíticas envolvendo o potencial gravitacional, a função de distribuição de energia e a densidade de estados. O perfil de densidade superficial relacionado é dado por

$$\Sigma_{Hernquist}(R) = 2\sigma_0 r_s G(R/r_s) \quad (13)$$

onde

$$G(X) = \frac{[(2+X^2)H(X) - 3]}{2(1-X^2)^2}$$

e

$$H(X) = \begin{cases} \frac{1}{\sqrt{1-X^2}} \operatorname{sech}^{-1}(X), & (X < 1) \\ 1, & (X = 1) \\ \frac{1}{\sqrt{X^2-1}} \operatorname{sec}^{-1}(X), & (X < 1) \end{cases}$$

o que implica que dadas as condições acima

$$\lim_{X \rightarrow 1} G(X) = \frac{2}{15},$$

em que  $X \equiv R/r_s$ , semelhante ao perfil de NFW.

### 3.3.5.3 Perfil de King

O perfil de King, inicialmente idealizado para sistemas estelares (King, 1962) e posteriormente aplicado ao aglomerado de Coma (King, 1972), é descrito por um perfil espacial que se ajusta razoavelmente bem a regiões centrais dos sistemas aos quais é empregado por meio da seguinte expressão

$$\sigma_{King}(r) = \frac{\sigma_0}{(1 + (r/r_c)^2)^{3/2}}, \quad (14)$$

que quando projetado a linha de visada, adquire a seguinte forma

$$\Sigma_{King}(R) = \frac{\Sigma_0}{(1 + (R/r_c)^2)}, \quad (15)$$

onde  $\sigma_0$  e  $\Sigma_0$  são respectivamente a densidade central tridimensional e projetada, com  $r_c$  equivalendo o raio do core do sistema considerado. As equações 14 e 15 são aproximações para o perfil de King, tendo em vista que não existe uma solução analítica exata para o modelo.

### 3.3.5.4 Ajuste dos perfis de densidade

O método usualmente empregado para ajustar o perfil de densidade consiste em realizar a divisão dos dados em anéis e, em seguida, obter o ajuste a partir da minimização do  $\chi^2$  (SARAZIN, 1986), como descrito na secção 3.2.3. Uma outra forma, reside em realizar o ajuste dos dados binados, através do método dos mínimos quadrados. A binagem dos dados, como utilizada, introduz uma imprecisão nos resultados do ajuste, tendo em vista que a escolha do anel pode não considerar pequenas estruturas, se muito grande, ou tornar maiores as incertezas estatísticas, se muito pequeno (SARAZIN, 1986).

Conforme vem sendo aplicado em outras ferramentas, utilizou-se o método da máxima verossimilhança para ajustar os perfis de densidade, combinando-o com a minimização do  $\chi^2$ . Para validar o teste, foi desenvolvido um programa para gerar amostras, utilizando o método descrito na secção 3.2.1, que seguissem os perfis apresentados.

Como os perfis não correspondem a distribuições de probabilidade, foi utilizada função proposta por [Mamon et al. \(2013\)](#), que tem como forma

$$p(R|b) = \frac{2\pi R[\Sigma(R) + \Sigma_b]}{N_p(R_{max}) - N_p(R_{min}) + \pi\Sigma_b(R_{max}^2 - R_{min}^2)}, \quad (16)$$

onde  $\Sigma(R)$  e  $N_p$  são, respectivamente, o perfil de densidade superficial e o número de objetos dentro de um raio  $R$ , enquanto  $\Sigma_b$  é a constante de *background* da densidade superficial. São traduzidos, dessa forma, os perfis de densidade superficial em funções de distribuição, o que permite a aplicação de métodos estatísticos como é o caso dos métodos da rejeição e máxima verossimilhança.

Sabemos que o *background*, para simulações computacionais, tem valor nulo, por ser obtido através de algoritmos como o *FoF*, e em dados observacionais, o seu valor é da ordem de  $4 \times 10^{-5}$  *gals/arcsec*<sup>2</sup> ([Adami et al., 1998](#)). Além disso, o  $\Delta N_p$  pode ser substituído por ([Mamon et al., 2013](#))

$$\Delta N_p = 2\pi \int_{R_{min}}^{R_{max}} R\Sigma(R)dR,$$

tornando possível reescrever a Equação 16 na forma simplificada:

$$p(R) = \frac{2\pi R\Sigma(R)}{\Delta N_p}. \quad (17)$$

Utilizamos a Equação 17 para realizar a simulação das amostras com densidade superficial estabelecida em  $\Sigma(R)$ . O raio de escala  $r_s$  dos perfis de densidade superficial pode ser estimado com o método da máxima verossimilhança. O erro associado a estimação é obtido através do método do *bootstrap*.

No próximo capítulo, apresentaremos os resultados da aplicação das ferramentas aqui descritas nos dados retratados no Capítulo 2.

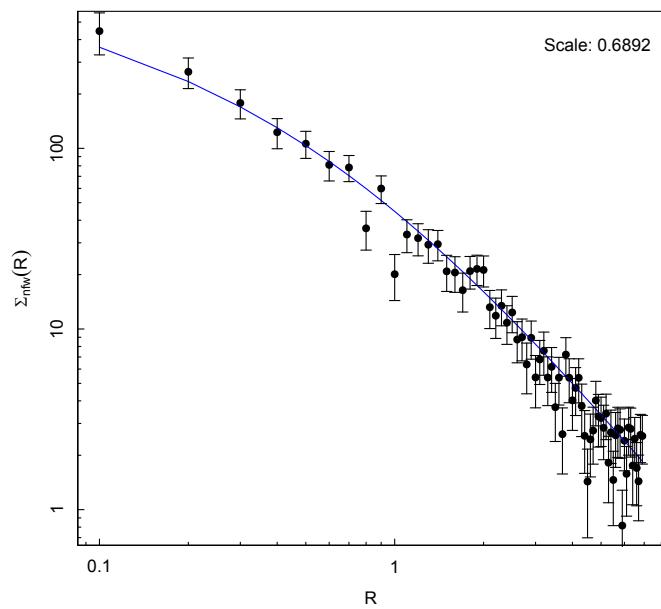


Figura 17 – Exemplo de geração de dados simulados e ajuste do perfil de NFW. Para o valor  $r_s = 0.7$  simulado, foi recuperado  $r_s = 0.6892$ .

## 4 RESULTADOS E DISCUSSÕES

Neste capítulo apresentaremos os resultados encontrados com a aplicação das ferramentas e apresentadas no Capítulo 3 aos dados descritos no Capítulo 2. É importante salientar que o conjunto de tais ferramentas, por si só, já constitui um resultado deste trabalho. Para um melhor entendimento da suas funções, a seguir nomearemos e descreveremos os códigos desenvolvidos.

- A ferramenta *Tri2Bi v1.0* recebe como entrada dados tridimensionais de simulações numéricas e realiza a projeção desses na linha de visada e no plano do céu. No desenvolvimento desta ferramenta, levamos em conta a cosmologia e o redshift em que o objeto se encontra, fazendo com que a conversão seja mais próxima possível de dados observados.
- Como a ferramenta *GNG* introduz uma grande quantidade de diagnósticos indeterminados, desenvolvemos uma alternativa, denominada *Análise-L v1.0*, que utiliza os L-Momentos da distribuição de velocidades para reconsiderar a classificação "indeterminada" de um aglomerado gerada pelo *GNG*. Uma região de aceitação é definida para o tamanho da amostra e, então, verifica-se se o par ordenado (L-Skewness, L-Kurtosis) se encontra dentro dessa região. Caso verdadeiro, o aglomerado é reclassificado como gaussiano.
- Para tentar medir os desvios de gaussianidade que podem representar estados estacionários do sistema, utilizamos a distribuição generalizada q-gaussiana. Desenvolvemos uma ferramenta, denominada *qGaussian v1.0*, que implementa a função de densidade da distribuição q-Gaussiana, um gerador de números aleatórios que seguem a distribuição, e um estimador do parâmetro *q* que utiliza o método da máxima verossimilhança.
- Desenvolvemos uma ferramenta para ajustar os perfis de densidade superficial dos modelos de NFW, Hernquist e King. Chamada de *density2D v1.0*, a ferramenta possui, para cada modelo, a função de densidade de probabilidade, um gerador de números aleatórios para a função de densidade, e um estimador dos parâmetros  $R_s$  ( $R_c$  para King) e  $\Sigma_0$ , que correspondem ao raio de escala (raio do core, para King) e a densidade central, respectivamente. Utilizamos os métodos da máxima verossimilhança e  $\chi^2$  para realizar as estimativas.

Desenvolvemos ainda diversos scripts com o propósito de automatizar a obtenção e o tratamento dos dados. Por não possuírem um caráter de produto final, esses

scripts não foram descritos nesta dissertação, mas são parte integrante da pesquisa. Nas seções a seguir, apresentaremos a aplicação das ferramentas implementadas, para que se tenha uma visão geral de como funcionam e que novas questões podem levantar.

## 4.1 Resultados do Millennium

Nos dados do Millennium, aplicamos a ferramenta *Tri2Bi* e, em seguida, o GNG em todos os aglomerados com mais que 20 objetos dentro do  $R_{200}$ . Dividimos os aglomerados em quatro intervalos de massa:  $\log_{10}M_{200} \leq 13.75$  (Grupo A);  $13.75 \leq \log_{10}M_{200} < 14.00$  (Grupo B);  $14.00 \leq \log_{10}M_{200} < 14.25$  (Grupo C); e  $\log_{10}M_{200} \geq 14.25$  (Grupo D), seguindo o estudo de [Ribeiro et al. \(2013\)](#).

### 4.1.1 Evolução dos aglomerados com o redshift

Uma vez que esperamos que aglomerados mais evoluídos possuam distribuição de velocidades gaussianas, podemos estimar a taxa com que estas amostras A, B, C e D evoluem ao equilíbrio, através da fração de aglomerados gaussianos em cada época. Inicialmente, investigamos os resultados encontrados por [Ribeiro et al. \(2013\)](#) e apresentados na Figura 18, que sugerem que a fração de aglomerados gaussianos aumenta para todos os intervalos de massa com o decorrer das épocas.

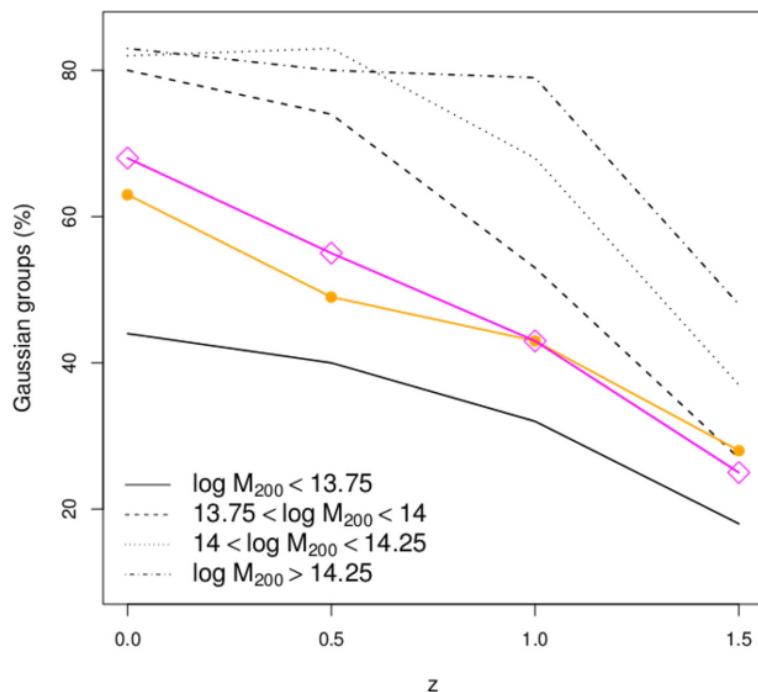


Figura 18 – Evolução de grupos gaussianos na simulação do Millennium, divididos por intervalos de massa. As linhas laranja e magenta representam, respectivamente, a evolução dos sistemas, em todas as massas, para os limites de magnitude  $M_r = -19$  e  $M_r = -20$ . Fonte: ([Ribeiro et al., 2013](#)).

É importante observar que o método aplicado neste trabalho difere do utilizado em [Ribeiro et al. \(2013\)](#), onde as velocidades projetadas foram obtidas utilizando outros métodos de conversão, a ferramenta GNG retornava apenas diagnósticos gaussianos ou não-gaussianos, uma vez que a classificação indeterminada foi introduzido por [de Carvalho et al. \(2017\)](#), e a quantidade mínima de objetos era de apenas oito galáxias em  $R_{200}$ , algo que possibilitava trabalhar com uma maior quantidade de aglomerados.

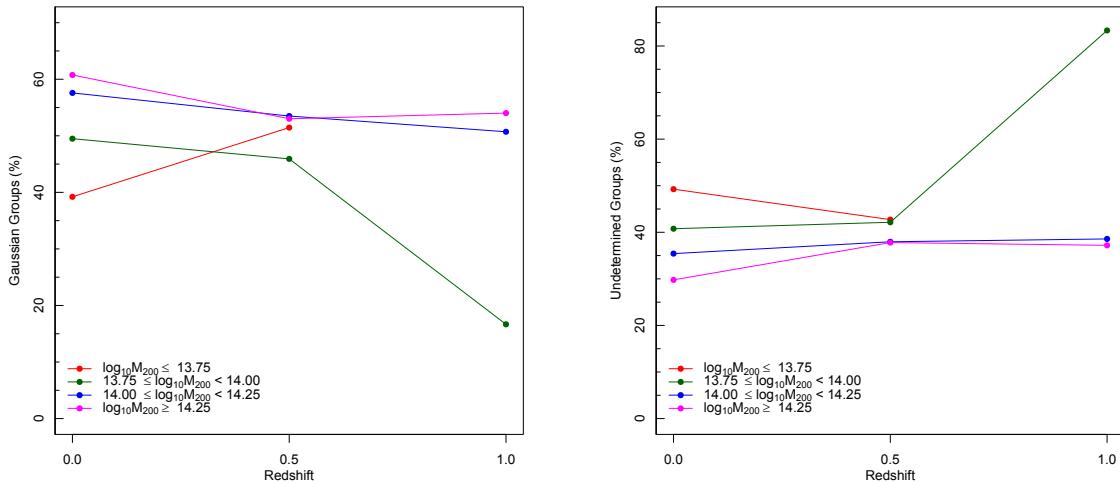


Figura 19 – Fração de aglomerados diagnosticados como Gaussianos (esquerda) e Indeterminados (direita) pelo GNG.

Investigando a fração de aglomerados, encontramos algumas diferenças em relação aos resultados encontrados em [Ribeiro et al. \(2013\)](#), como pode ser visto na Figura 19. Entre os redshifts  $z = 1.0$  e  $z = 0.5$ , observamos um aumento significativo da fração de aglomerados gaussianos para o grupo B, enquanto o grupo C apresenta um leve aumento e o grupo D uma suave diminuição. Entre  $z = 0.5$  e  $z = 0.0$ , o grupo A, menos massivo, aparece na amostra, apresentando mais de 50% de objetos diagnosticados como sendo gaussianos, percentual que diminui para pouco menos de 40% com distribuição gaussiana, em  $z = 0.0$ . Os demais grupos apresentam um suave aumento na fração de objetos em  $z = 0.0$ , se comparado a  $z = 0.5$ . Dentre os aglomerados classificados como sendo indeterminados, destaca-se o grande decrescimento da taxa de aglomerados para o grupo B, entre  $z = 1.0$  e  $z = 0.5$ . Os demais grupos mantém um comportamento estável até  $z = 0.5$ . Entre  $z = 0.5$  e  $z = 0.0$ , os grupos B, C e D apresentam uma suave diminuição, enquanto o grupo A apresenta leve crescimento. É importante notar que em  $z = 0$  os grupos possuem de 30% a 50% de sistemas indeterminados, o que significa uma diminuição considerável da amostra com diagnósticos categóricos, G ou NG.

As frações observadas no lado esquerdo da Figura 19 e na Tabela 1 constituem caminhos evolutivos diferentes daqueles observados na Figura 18. Contudo, para os grupos B, C e D, a tendência geral é de crescimento da fração de sistemas gaussianos, concordando com os resultados de [Ribeiro et al. \(2013\)](#), para estes grupos. A divergência

mais evidente surge para o grupo A, de aglomerados menos massivos, que ao invés de apresentar também uma tendência de aumento da fração de gaussianos, apresenta um comportamento de diminuição desta fração. Devemos notar que este grupo sofre também um aumento dos sistemas classificados como indeterminados. É importante notar que a comparação com o trabalho de Ribeiro et al. (2013) não é direta, dadas as diferenças metodológicas.

Grupo	$z = 0.0$ (%)	$z = 0.5$ (%)	$z = 1.0$ (%)
Grupo A	$39.21 \pm 14.79$	$51.46 \pm 09.03$	-
Grupo B	$49.49 \pm 30.59$	$45.92 \pm 29.72$	$16.67 \pm 07.10$
Grupo C	$57.58 \pm 24.64$	$53.49 \pm 28.56$	$50.71 \pm 34.31$
Grupo D	$60.75 \pm 21.18$	$56.91 \pm 20.74$	$46.15 \pm 31.36$

Tabela 1 – Fração de aglomerados gaussianos e a incerteza associada

Avançamos em seguida a análise, aplicando a *Análise-L*, com a finalidade de recuperar parte dos aglomerados classificados como indeterminados para um diagnóstico gaussiano. Na Figura 20, vemos que alguns objetos classificados como indeterminados (em verde) estão dentro da área com maior predominância de aglomerados gaussianos. Dependendo do tamanho da amostra, esses aglomerados podem ser reclassificados como gaussianos, como abordado na secção 3.3.3.

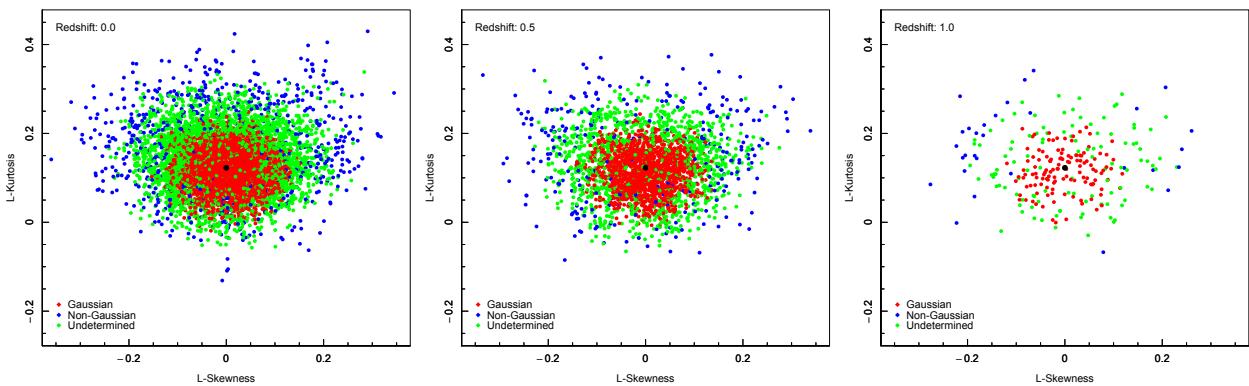


Figura 20 – Diagrama dos L-Momentos para os redshifts  $z = \{0.0, 0.5, 1.0\}$  do Millennium, separados pelo diagnóstico do GNG.

Como podemos ver no exemplo mostrado na Figura 21, em geral, apenas um pequeno desvio da gaussianidade é notado para as velocidades dos aglomerados que são reclassificados pela Análise-L. Isso demonstra que o método utilizado se mostra como uma boa alternativa para recuperar objetos diagnosticados como indeterminados.

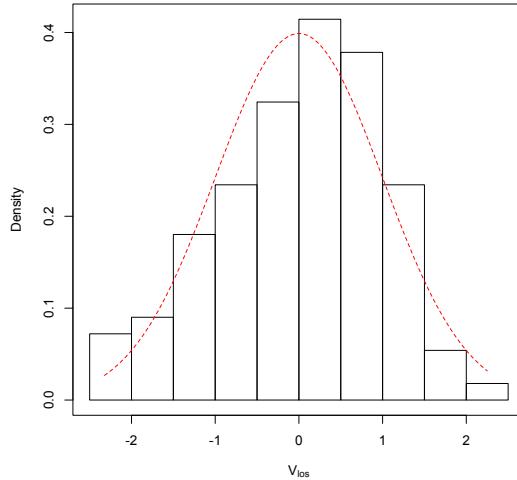


Figura 21 – Exemplo de histograma para VLOS de um aglomerado que foi reclassificado como gaussiano pela Análise-L. A curva tracejada em vermelho representa a distribuição gaussiana com  $\sigma = 1$ .

Refizemos o estudo da evolução de aglomerados com a introdução dos gaussianos reclassificados após a Análise-L, e verificamos que isso altera significativamente o resultado anterior. Em  $z = 1.0$ , todos os aglomerados do grupo B foram diagnosticados como sendo gaussianos, fração que diminui para aproximadamente 84% em  $z = 0.5$  e se mantém estável até  $z = 0.0$ . Diferente do resultado encontrado antes da aplicação da Análise-L, o grupo D apresenta uma fração de gaussianos razoavelmente estável,  $\sim 79\%$  entre  $z = .01$  e  $z = 0.0$ . Ao mesmo tempo, os grupos A e C apresentam comportamento análogo ao visto na Figura 19, mas agora alcançam frações finais mais altas de sistemas gaussianos, 82% e 84%, respectivamente. O caminhos evolutivos após a recuperação de parte dos sistemas indeterminados tornam-se mais distintos daqueles obtidos por Ribeiro et al. (2013), embora as frações finais de sistemas G para os grupos B, C e D agora se aproximam daquelas vistas na Figura 18, enquanto a fração final de sistemas G no grupo A, apesar de ter comportamento de queda, possui fração final mais alta do que a obtida por Ribeiro et al. (2013), quanto aquela obtida em nosso estudo antes da Análise-L. Mais uma vez, enfatizamos que esta comparação tem apenas um caráter descritivo, e ilustra o quanto as metodologias influenciam os resultados.

De fato, a variação observada nos caminhos evolutivos apresentados nos painéis das Figuras 19 e 22 indica a importância dos objetos indeterminados provenientes do programa GNG. O simples reaproveitamento de parte destes sistemas (no presente caso algo em torno de 75% em  $z = 0$ ) pode afetar diretamente a evolução dos aglomerados de galáxias em  $z \leq 1.0$ . Do ponto de vista astrofísico, o fato de que nem todos os grupos evoluam para terem distribuições de velocidades gaussiana (ou que dela se afastem) sugere que os processos de *infall* e fusões podem afetar significativamente a dinâmica desses sistemas.

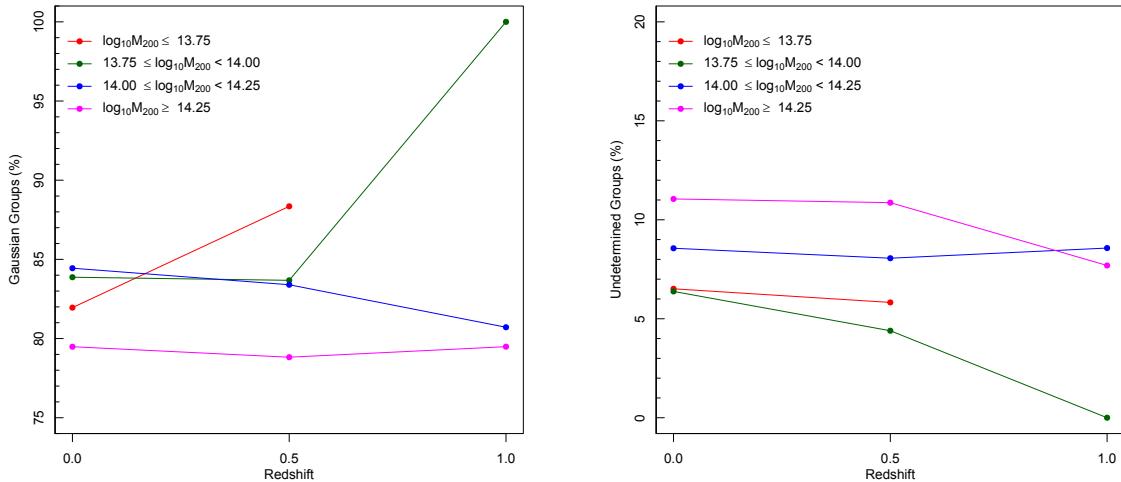


Figura 22 – Fração de aglomerados diagnosticados como Gaussianos (esquerda) e Indeterminados (direita) pelo GNG após a aplicação da Análise-L.

#### 4.1.2 Perfil de densidade superficial

Com o intuito de estabelecer uma relação entre o diagnóstico do GNG e a forma do perfil de densidade superficial, dividimos os aglomerados de acordo com o seu diagnóstico e realizamos o empilhamento das galáxias para cada redshift. Em seguida, ajustamos o perfil de densidade superficial utilizando a ferramenta *density2D*. Como vemos na Figura 23, a pilha de aglomerados não-gaussianos apresenta a curva do perfil de densidade menos inclinada nas regiões mais externas para todos os redshifts, se comparado a pilha de gaussianos. Esse comportamento reflete a menor concentração das galáxias nas redondezas da região central do aglomerado, sugerindo que estes sistemas estejam recebendo maior taxa de *infall* em todo o intervalo de redshift. Percebemos ainda que todas as pilhas apresentaram melhor ajuste para o perfil de NFW, exceto em  $z = 1.0$  da pilha não-gaussiana, que mostrou uma melhor aderência ao perfil de Hernquist.

#### 4.1.3 Estimação da q-Gaussiana

Para finalizar nossas análises com os dados da simulação Millennium, estimamos o valor de  $q$  para os aglomerados com mais de 20 objetos dentro de  $R_{200}$ . Em seguida, calculamos o  $q$  médio para os mesmos intervalos de massa apresentados na seção anterior e, finalmente, corrigimos o viés sobre a estimação de  $q$  como descrito na secção 3.3.4. Podemos ver na Figura 24 que o valor de  $q$  tende a aumentar com o decorrer das épocas, evidenciando a evolução dos aglomerados em direção ao equilíbrio, que aconteceria através de uma sequência de estados estacionários descritos por diferentes valores de  $q$ . Além disso, nota-se que aglomerados com massa intermediária (Grupos B e C) tendem a atingir a gaussianidade mais rapidamente do que aglomerados com

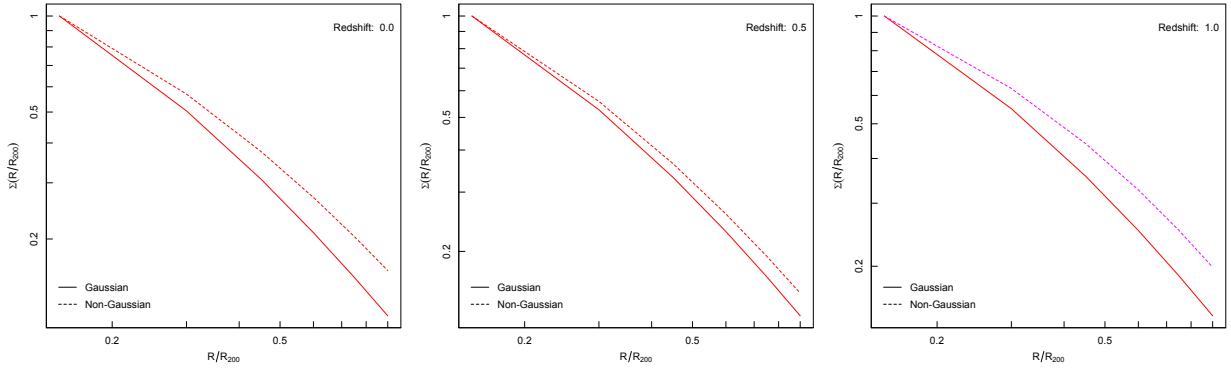


Figura 23 – Perfis de densidade superficial para o empilhamento dos aglomerados do Millennium. A linha vermelha representa o perfil de NFW, enquanto a linha magenta corresponde ao perfil de Hernquist.

alta massa (Grupo D), o que sugere que esses sistemas muito massivos têm uma maior dificuldade de atingir o equilíbrio, possivelmente pelo fato de se localizarem nos nós dos filamentos cósmicos e receberem o *infall* de grandes estruturas constantemente.

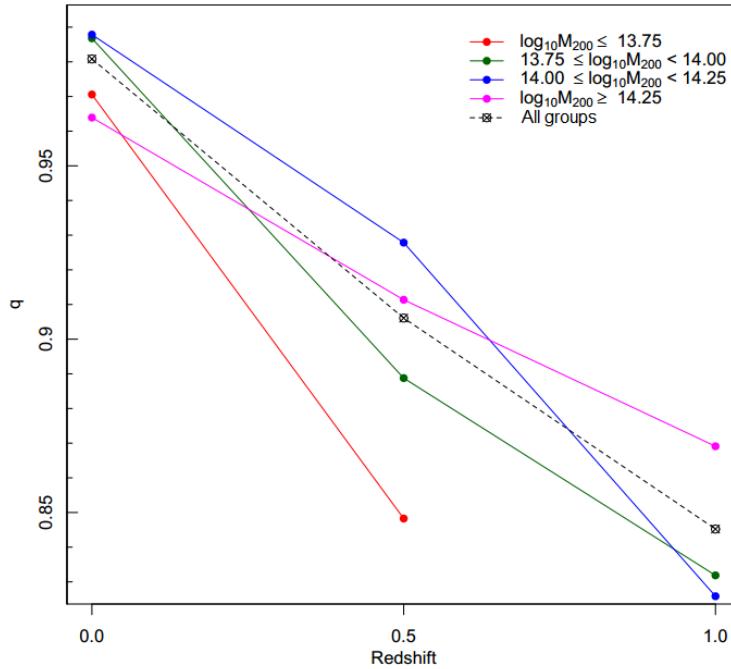


Figura 24 – Valor estimado de  $q$  por intervalos de massa e em diferentes redshifts.

Os aglomerados do Grupo A são os que possuem a maior taxa de crescimento em direção à unidade, que representa a gaussiana. Curiosamente, esse comportamento geral reflete o resultado encontrado por Ribeiro et al. (2013) e discutido na seção anterior. É importante notar que a estimação de  $q$  na presente versão do programa *qGaussian* ainda é aplicada apenas para dados tridimensionais, sendo que utilizamos as equações 4 e 6 para obter o módulo da velocidade que é utilizado na análise. Por outro lado, o

procedimento utilizado por [Ribeiro et al. \(2013\)](#) foi escolher uma das componentes de velocidade como sendo a projeção da velocidade radial na linha de visada. Esta escolha pode ser responsável pelas inconsistências observadas na seção [4.1.1](#), e pela aparente concordância com os resultados da q-gaussiana.

## 4.2 Resultados do EAGLE

As análises que foram realizadas para o Millennium possuem um caráter global, com o objetivo de traçar um cenário da evolução média dos aglomerados nas quatro subamostras definidas por massa. Para as amostras do EAGLE, buscamos analisar características mais específicas, tendo como foco aglomerados individuais. Como o EAGLE apresenta, se comparado ao Millennium, uma maior granularidade com relação aos cortes em redshift disponíveis, isto permite a utilização de seus dados para obter a evolução de sistemas individualmente. Extraímos então dois aglomerados da simulação do EAGLE, ambos com massa  $\log_{10}M_{200} \geq 14.0$ , sendo que um possui um maior grau de isolamento (A0001) enquanto o outro parece interagir gravitacionalmente com outro objeto em  $z = 0.00$ . Dado que temos acesso aos dados dos dois aglomerados em diversos redshifts, podemos observar de que maneira se dá a evolução da distribuição de velocidades e os perfis de densidade dos sistemas em períodos diferentes. Para as duas amostras, as galáxias foram selecionadas realizando-se dois cortes em magnitude na banda  $r$  de  $M_r \leq -19.0$  e  $M_r \leq -17.0$ , o que torna possível verificar se os resultados encontrados sofrem grandes variações ao incluirmos galáxias de baixa luminosidade.

Redshift	$M_r \leq -19$	$M_r \leq -17$	Redshift	$M_r \leq -19$	$M_r \leq -17$
0.37	40	73	0.37	38	84
0.27	45	89	0.27	64	135
0.18	40	79	0.18	78	164
0.10	46	94	0.10	87	193
0.00	54	103	0.00	105	243

Tabela 2 – Quantidade de galáxias dentro de  $R_{200}$  nos aglomerados A0004 (esquerda) e A0001 (direita), com cortes de magnitude  $M_r \leq -17$  e  $M_r \leq -19$ .

Inicialmente, selecionamos os aglomerados em redshift  $z = 0.37$ , por possuírem mais que 20 aglomerados dentro de  $R_{200}$  para o corte  $M_r \leq -19$ , e rastreamos suas posições em  $z = \{0.00, 0.10, 0.18, 0.27\}$ . Em seguida, selecionamos as galáxias inseridas dentro do  $R_{200}$  e retiramos os subhalos "vazios" de matéria escura, ou seja, objetos que possuem massa de matéria escura mas não possuem massa estelar. Finalmente, aplicamos o GNG aliado à *Análise-L*. Redshifts mais altos não foram selecionados por não possuírem aglomerados contendo mais de 20 objetos dentro de  $R_{200}$ .

A Figura 25 apresenta a evolução do diagnóstico do GNG para as duas amostras estudadas, com corte de magnitude em  $M_r \leq -19$ . É possível observar que para ambos os objetos ocorre uma variação no diagnóstico do GNG. Este comportamento demonstra como os aglomerados perdem (ou recuperam) o equilíbrio no decorrer do tempo. Alternativamente, devemos considerar a possibilidade de que estas flutuações de diagnóstico sejam decorrentes das próprias limitações das ferramentas utilizadas.

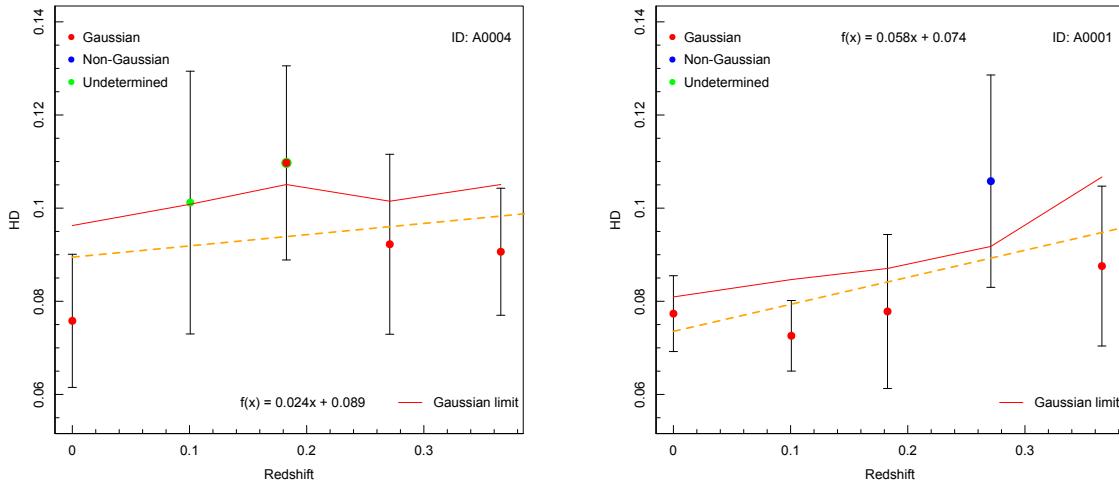


Figura 25 – Diagnóstico de HD para as amostras estudadas (A0004, esquerda; A0001, direita) em diferentes redshifts, com  $M_r \leq -19.0$ . A linha vermelha representa o valor de HD para que o aglomerado venha a ser considerado gaussiano se obtiver mais que 70% dessa classificação nas reamostragens. A linha laranja tracejada corresponde a uma regressão linear para os pontos.

Em A0004, o aglomerado é perturbado em  $z = 0.18$ , de maneira que o diagnóstico do GNG passa para indeterminado, mas após a aplicação da *Análise-L*, ele volta a ser classificado como gaussiano. O mesmo não ocorre em  $z = 0.10$ , onde o GNG identifica o objeto como indeterminado e esta classificação prevalece mesmo após a *Análise-L*. Para A0001, uma perturbação em  $z = 0.27$  faz com que o objeto seja diagnosticado como não-gaussiano pelo GNG. A gaussianidade é recuperada em  $z = 0.18$  e se mantém até  $z = 0.00$ . Devemos notar que, apesar da flutuação de diagnóstico, a tendência geral é a de uma evolução para valores menores de HD, ou seja, para menores desvios de uma distribuição gaussiana, sugerindo que ambos aglomerados evoluem para o equilíbrio.

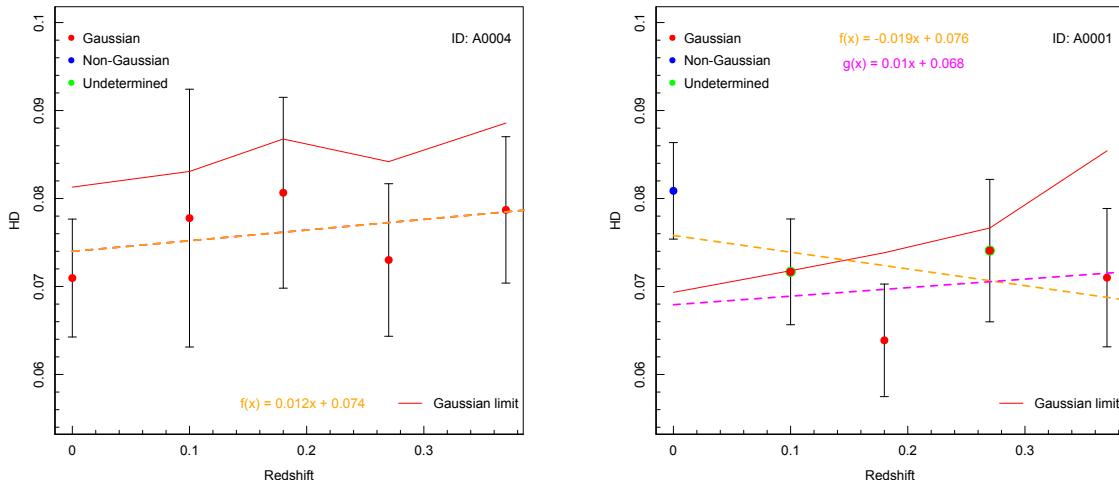


Figura 26 – Diagnóstico de HD para as amostras estudadas (A0004, esquerda; A0001, direita) em diferentes redshifts, com  $M_r \leq -17.0$ .

Realizando o corte de magnitude em  $M_r \leq -17$ , notamos que a amostra A0004 mantém o diagnóstico gaussiano em todos os redshifts, sem sentir as variações apresentadas no corte  $M_r \leq -19$ . A tendência geral é de leve queda dos valores de HD, confirmando o resultado anterior. O aglomerado A0001 mostra diagnóstico indeterminado pelo GNG e recuperado pela *Análise-L*, no redshift  $z = 0.27$ . Essa classificação pode refletir a não-gaussianidade apresentada no corte  $M_r \leq -19$ . Nos redshifts seguintes, com o corte  $M_r \leq -19$ , o diagnóstico é de gaussianidade, inclusive em  $z = 0.0$ , que apresentou classificação não-gaussiana no corte em  $M_r \leq -17$ . Não fosse por este ponto, a tendência geral seria de queda para valores menores de HD.

Investigamos a variação do número de objetos para verificar como a acreção de matéria afeta os diagnósticos (vide Figura 27). Para o objeto A0001, o aumento de número de objetos pode ter influenciado a classificação do aglomerado em  $z = 0.27$ , enquanto que para o objeto A0004, a diminuição do número de objetos pode ter feito com que o aglomerado sofresse com perturbações no diagnóstico em  $z = 0.18$ , tendo seu equilíbrio novamente retomado em  $z = 0$ . Observa-se ainda que, como a quantidade de objetos de A0004 é menor se comparada a A0001, ele se encontra muito mais propenso a variações de diagnóstico em virtude do aumento ou diminuição da quantidade de objetos. A influência do tamanho das amostras sobre o diagnóstico das distribuições de velocidades está relacionada ao fato de que tanto a calibração do método HD quanto a definição de 70% de confiabilidade usado no código GNG têm um caráter estatístico e são direcionadas ao uso em grandes amostras de dados. Além disso, os ingressantes nos aglomerados em diferentes épocas são contabilizados desde que estejam dentro de  $R_{200}$ , a despeito de constituírem objetos virializados ou não.

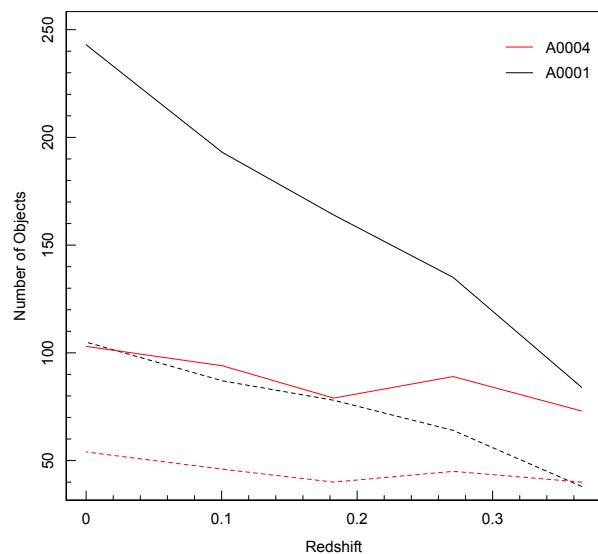


Figura 27 – Evolução da quantidade de objetos em  $R_{200}$ . A linha contínua representa o corte em magnitude  $M_r \leq -17$ , enquanto a linha tracejada corresponde a  $M_r \leq -19$ .

Para ilustrar este último ponto, apresentamos na Figura 28 a evolução das frações de objetos com alta e baixa taxa de formação estelar específica. A divisão foi feita grosseiramente em torno do valor  $\log_{10} sSFR \text{ yr}^{-1} = -11$  (vide (BAUER et al., 2005)). Espera-se que à medida que o sistema evolua para  $z = 0$ , a fração de objetos com baixa (alta) sSFR aumente (diminua). Os comportamentos observados na Figura correspondem a esta expectativa. Contudo, devemos notar que enquanto A0001 tem um crescimento moderado da fração de objetos com baixa sSFR, chegando a algo em torno de 20% em  $z = 0$ , o aglomerado A0004 tem um crescimento pronunciado a partir de  $z = 0.18$ , chegando a uma fração de  $\sim 50\%$  em  $z = 0$ . Objetos com baixa sSFR têm maior probabilidade de terem ingressado no sistema há mais tempo e, portanto, de participarem da virialização central do aglomerado. De forma complementar, a fração de objetos em A0001 com alta sSFR dentro de  $R_{200}$  é muito alta em  $z = 0$ ,  $\sim 80\%$ . O aglomerado A0004 tem uma fração mais moderada, de  $\sim 50\%$ . Devemos notar ainda, na Figura 27, que a grande acreção de objetos em A0001 se dá para objetos de baixa luminosidade,  $-19 \leq M_r \leq -17$ , e que este parece ser o grande diferencial deste sistema. Ou seja, a não-gaussianidade de A0001 em  $z = 0$  provavelmente está relacionada a um grande número de objetos de baixo fluxo com altas taxas de sSFR, que entram em  $R_{200}$  em baixos redshifts, alterando a amostra que é utilizada para estudar a sua distribuição de velocidades. Este resultado reforça a importância da seleção de galáxias a serem utilizadas nos códigos de diagnóstico dinâmico.

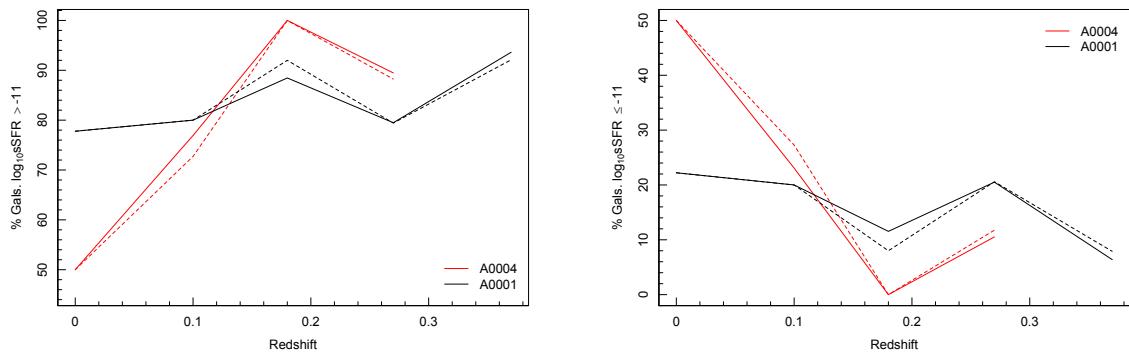


Figura 28 – Evolução da sSFR em  $R_{200}$ . A linha contínua representa o corte em magnitude  $M_r \leq -17$ , enquanto a linha tracejada corresponde a  $M_r \leq -19$ .

Analisamos agora o espaço de fase projetado das duas amostras, buscando encontrar alguma característica que ajude a compreender os resultados anteriores. Traçamos adicionalmente a reta descrita por Oman et al. (2013) que delimita aproximadamente a região onde espera-se que haja equilíbrio virial.

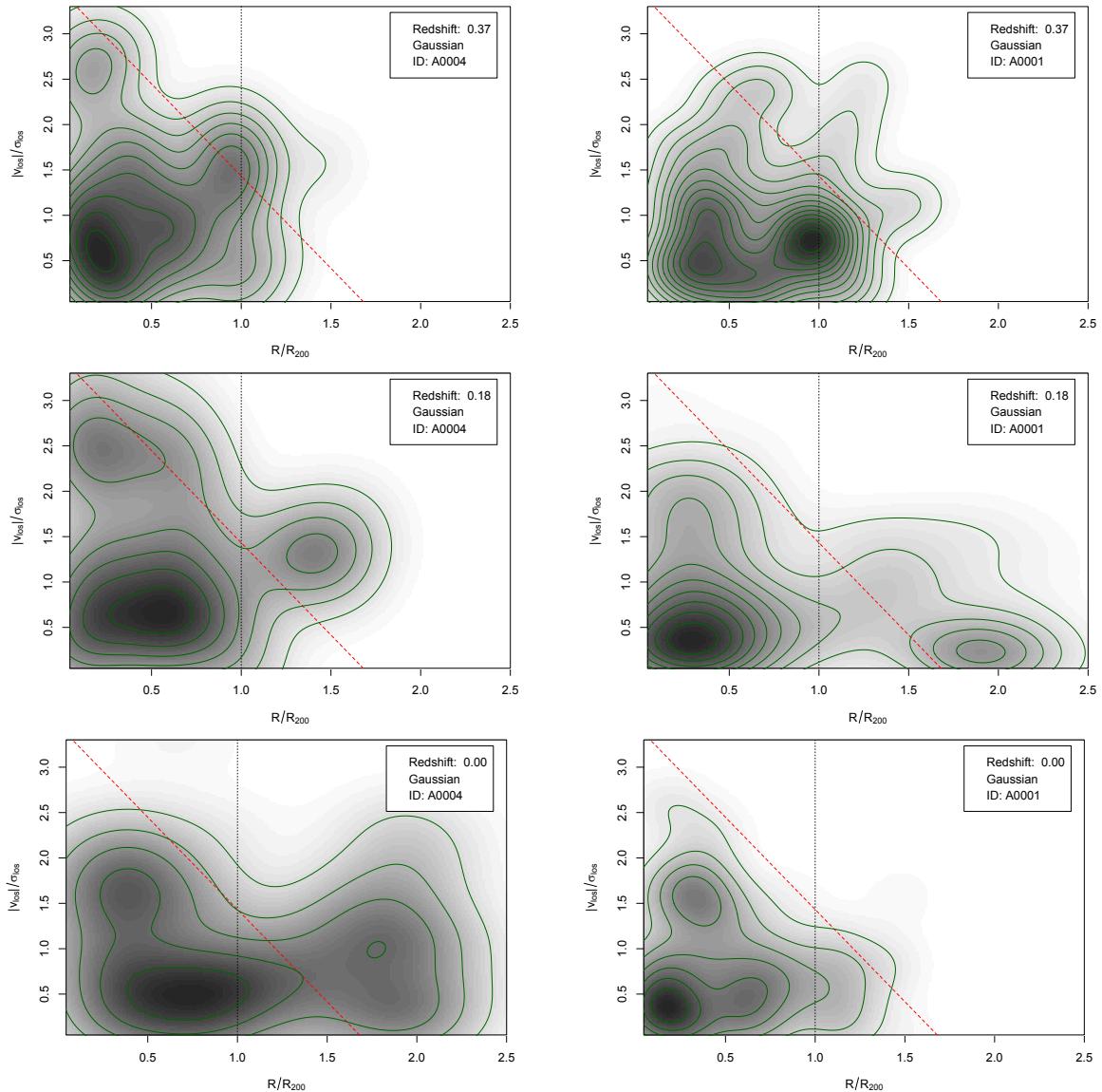


Figura 29 – Evolução do espaço de fase para as amostras A0004 (esquerda) e A0001 (direita), para o corte de magnitude  $M_r \leq -19$ . A linha tracejada em vermelho representa a delimitação de [Oman et al. \(2013\)](#).

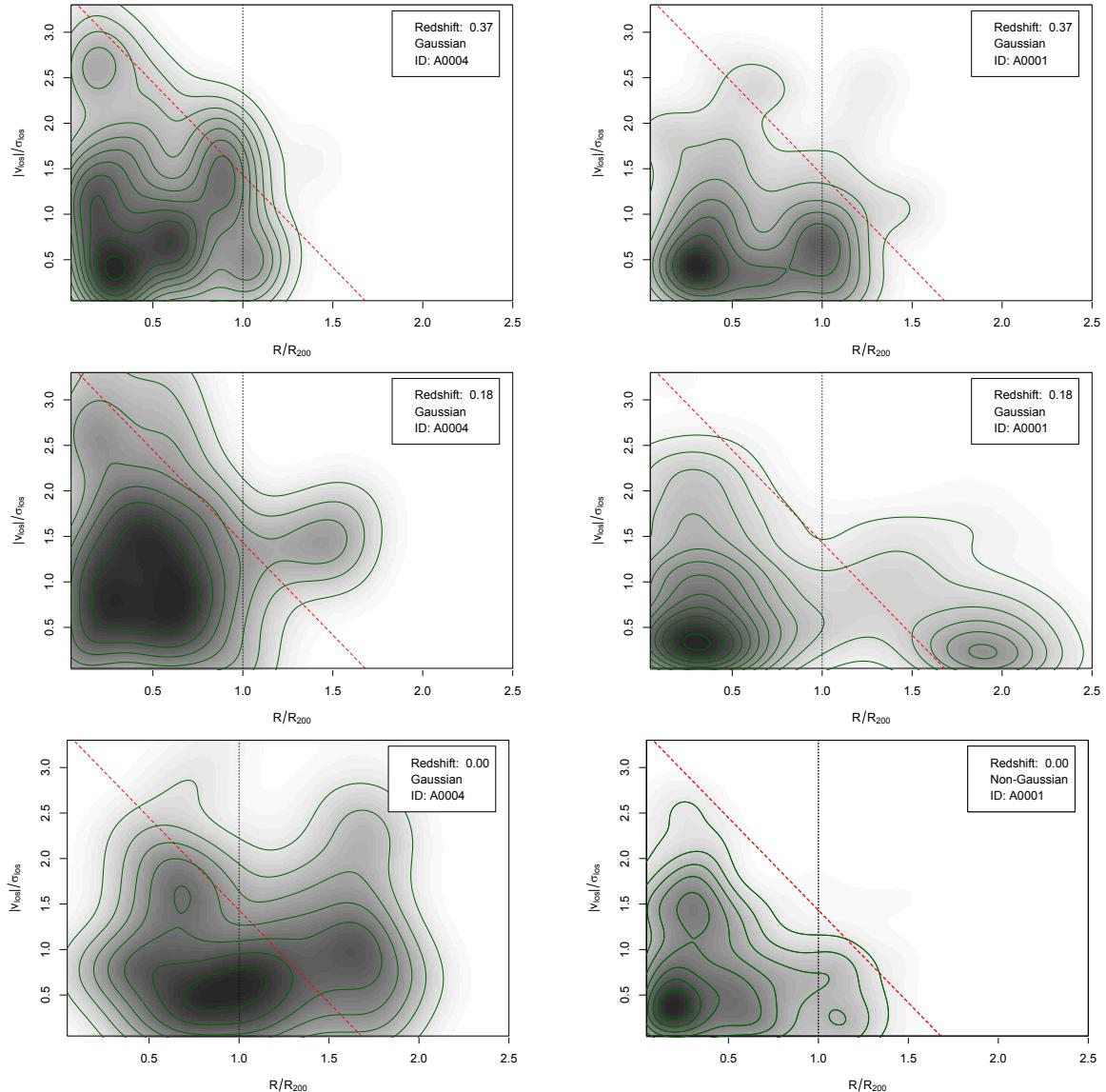


Figura 30 – Evolução do espaço de fase para as amostras A0004 (esquerda) e A0001 (direita), para o corte de magnitude  $M_r \leq -17$ .

Para confrontar a distribuição do espaço de fase para redshifts consecutivos, com o objetivo de verificar se a configuração do espaço de fase é mantida, aplicamos um teste estatístico baseado na distribuição suavizada dos pontos, obtida através de um *kernel* adaptativo que converte a distribuição bivariada de pontos numa distribuição de densidades. O teste, cujos detalhes são descritos em [de Carvalho et al. \(2017\)](#), retorna um valor-p associado à probabilidade de que as amostras tenham sido retiradas da mesma população. Em R, podemos executar o teste usando função *kde.test*, do pacote *ks*. Definimos que se o valor-p for menor que 0.1, as distribuições possuem diferenças significativas. Ou seja, trabalhamos com diferenças significativas para um nível de confiança de 90%.

<i>Redshift</i>	<i>Objeto</i>	<i>valor-p</i>	<i>Objeto</i>	<i>valor-p</i>
0.37 - 0.27	A0004	0.002	A0001	0.101
0.27 - 0.18	A0004	0.051	A0001	0.318
0.18 - 0.10	A0004	0.400	A0001	0.314
0.10 - 0.00	A0004	0.006	A0001	0.156

Tabela 3 – Resultados do *kde.test* para o corte de magnitude  $M_r \leq -19$ .

Como podemos ver nas Tabelas 3 e 4, aplicamos o teste *kde* nas galáxias dos aglomerados A0004 e A0001, comparando os espaços de fase entre duas saídas subsequentes em redshift. Para o corte  $M_r \leq -19$ , vemos uma intensa evolução na distribuição de galáxias sobre o espaço de fase de A0004, com diferenças significativas sendo detectadas entre  $z = 0.37$  e  $z = 0.27$ , assim como entre  $z = 0.27$  e  $z = 0.18$ , e também entre  $z = 0.10$  e  $z = 0.0$ . Por outro lado, não foram verificadas diferenças significativas nos retratos de A0001 entre uma época e outra. Este resultado parece em desacordo com a inspeção visual de A0001 nos painéis da Figura 29, embora a estatística do teste dependa da integral da diferença quadrática das funções de densidade de probabilidade em todo o domínio dos dados, o que significa que semelhanças em determinadas regiões do espaço de fase possam compensar diferenças em outras, tornando o resultado "pouco visual".<sup>1</sup>

<i>Redshift</i>	<i>Objeto</i>	<i>p-value</i>	<i>Objeto</i>	<i>p-value</i>
0.37 - 0.27	A0004	0.110	A0001	0.043
0.27 - 0.18	A0004	0.128	A0001	0.144
0.18 - 0.10	A0004	0.742	A0001	0.026
0.10 - 0.00	A0004	0.009	A0001	0.029

Tabela 4 – Resultados do *kde.test* para o corte de magnitude  $M_r \leq -17$ .

<sup>1</sup>Em comunicação privada com T. Duong, desenvolvedor do teste *kde*, a sugestão de uso é para preenchimentos mais densos das regiões que serão comparadas. Em nosso caso, isto indicaria melhor aproveitamento da técnica em dados empilhados do que em sistemas isolados.

Quando analisamos os resultados do teste para o corte  $M_r \leq -17$ , vemos uma inversão de comportamento, tendo agora A0004 mais estável, apresentando diferença significativa na distribuição de galáxias apenas entre  $z = 0.1$  e  $z = 0.0$ , época em que o sistema parece iniciar um processo de fusão com outra estrutura e/ou começa a acretar matéria de forma mais acentuada. Ao mesmo tempo, A0001 apresenta diferenças significativas em seu espaço de fase entre  $z = 0.37$  e  $z = 0.27$ ,  $z = 0.18$  e  $z = 0.10$ , assim como entre  $z = 0.10$  e  $z = 0.0$ . Mais uma vez, este sistema parece ser fortemente afetado pelo *infall* de objetos de baixo fluxo. Novamente, a inspeção visual parece indicar também forte evolução em A0004, que não foi detectada pelo teste kde. A dificuldade em se mensurar diferenças em distribuições multivariadas, de modo geral, levanta a questão de como se analisar o espaço de fase de aglomerados com este fim. Um desenvolvimento possível é o estudo da forma do espaço de fase e seu perfil combinado 2D. Este é um dos desenvolvimentos futuros dentro deste projeto.

Uma vez que o teste kde não se mostrou totalmente satisfatório, decidimos investigar a possível variação dos perfis de densidade superficial em relação ao redshift. Dessa forma, podemos ver como a distribuição de galáxias se modifica com o decorrer das épocas, e se há alterações no perfil que melhor ajusta os dados. Inicialmente, apresentamos na Figura 31 os perfis para o corte em  $M_r \leq -19$ . Para a amostra A0004, vemos que o perfil de King melhor ajusta os dados em todas as épocas, exceto em  $z = 0.27$ , onde NFW se mostrou melhor. Nota-se que não há uma organização dos perfis, com curvas mais internas para redshifts mais altos e curvas mais externas para valores menores de  $z$ . Nesse sentido, o aglomerado A0001 apresenta uma maior organização aparente das curvas, o que pode caracterizar um estágio de maior evolução se comparado a A0004. Ainda sobre A0001, para  $z = 0.37$ ,  $z = 0.27$  e  $z = 0.00$ , o perfil de Hernquist mostrou-se o melhor a ajustar os dados. Em ambas as amostras, observamos que existe o aumento da densidade superficial para todo  $R_{200}$ . Apesar da ausência de um padrão claro de evolução, fica claro que há variações significativas (embora difíceis de se distinguir e mensurar) nos perfis dos aglomerados com o redshift que podem sinalizar aproximações ou afastamentos de estados estacionários e/ou do estado de equilíbrio do sistema.

Para o corte  $M_r \leq -17$ , no aglomerado A0004, notamos que NFW se mostrou como a melhor alternativa para os redshifts  $z = 0.00$ ,  $z = 0.18$  e  $z = 0.27$ . Os únicos redshifts que mantiveram o perfil de densidade superficial foram  $z = 0.37$  e  $z = 0.10$ , onde King se mostrou melhor para ambos os cortes de magnitude. Para A0001, apenas os redshifts  $z = 0.10$  e  $z = 0.27$  apresentaram o mesmo perfil que no corte  $M_r \leq -19$ . Como vemos na Figura 32, a mesma e bem definida separação dos perfis de densidade por redshift que é notada para  $M_r \leq -19$ , é vista para  $M_r \leq -17$ .

Para finalizar nossas análises, realizamos o estudo do índice  $q$  da distribuição

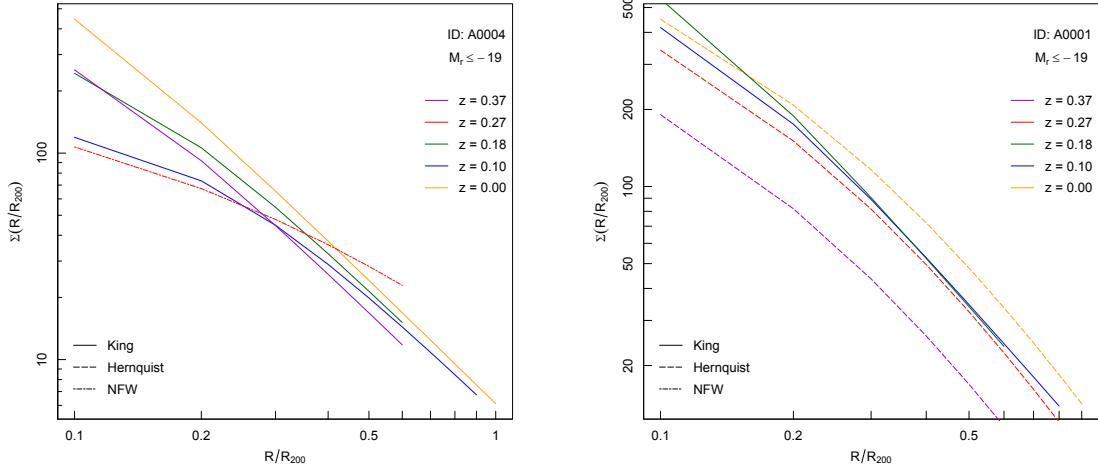


Figura 31 – Variação dos perfis de densidade superficial das amostras A0004 (esquerda) e A0001 (direita), no decorrer das épocas, com corte de magnitude  $M_r \leq -19$ .

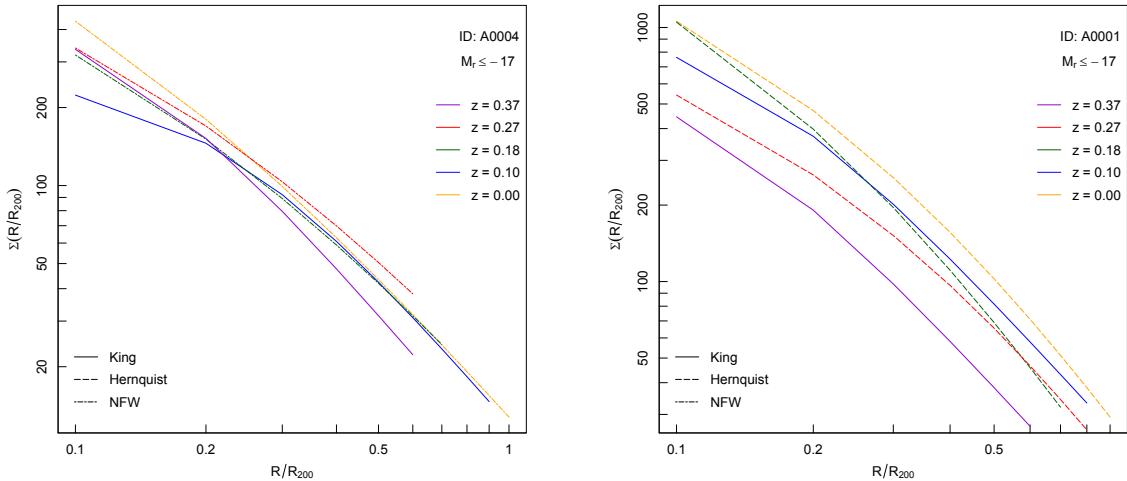


Figura 32 – Variação dos perfis de densidade superficial das amostras A0004 (esquerda) e A0001 (direita), no decorrer das épocas e com corte de magnitude  $M_r \leq -17$ .

$q$ -Gaussiana, com o objetivo de entender melhor a evolução dos aglomerados com o redshift. Podemos observar na Figura 33 que o parâmetro  $q$ , para o aglomerado A0004 no corte  $M_r \leq -19$ , apresenta um decrescimento com o decorrer das épocas, o que indicaria o afastamento da região de equilíbrio e uma convergência para algum estado estacionário dentro da família  $q$ . Para o objeto A0001,  $q$  se mantém estável em torno de  $q = 0.9$ , não refletindo as mudanças dos diagnósticos vistos na Figura 25. Ou seja, enquanto HD indica desvios cada vez menores da gaussiana,  $q$  permanece estável sugerindo a permanência do sistema em algum estado estacionário. Esse comportamento é até certo ponto inesperado e sugere a necessidade de estudos posteriores, utizando um maior número de aglomerados. Devemos lembrar, no entanto, que  $q$  avalia o módulo da velocidade tridimensional, enquanto HD avalia a velocidade projetada na linha de

visada e que a comparação, portanto, não é direta.

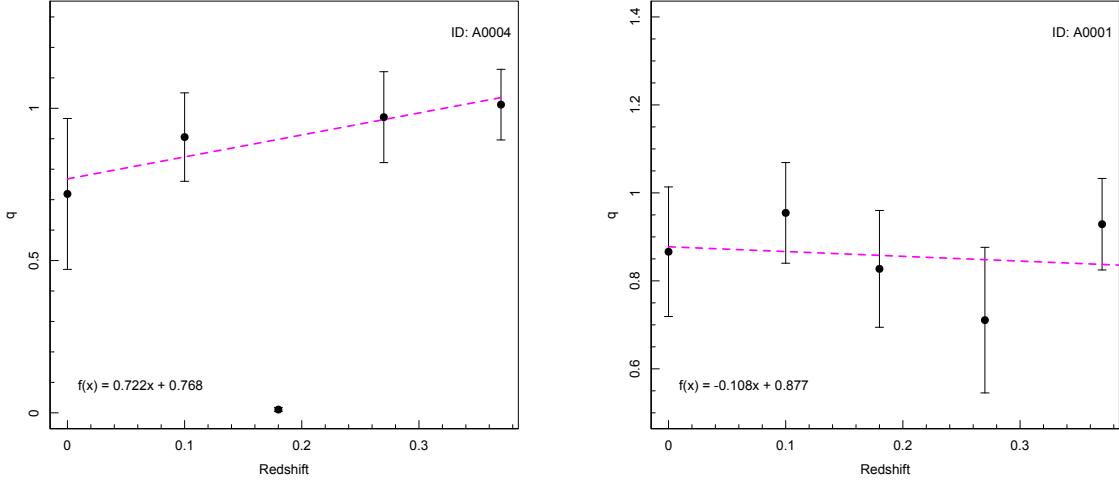


Figura 33 – Evolução do índice  $q$  da distribuição q-Gaussiana para os dados do EAGLE com corte em magnitude  $M_r \leq -19$ .

Após realizarmos o corte de magnitude em  $M_r \leq -17$ , como mostra a Figura 34, percebemos que, para a amostra A0004, o comportamento do parâmetro  $q$  não sofreu grandes modificações, se comparado ao corte analisado anteriormente. Percebemos, inclusive, que a flutuação vista em  $z = 0.18$  ainda é apresentada, mas com um menor impacto. No aglomerado A0001, vemos o decrescimento de  $q$  com o avanço em redshift. Comportamento diferente do que o visto no corte  $M_r \leq -19$ . Em ambos os casos, vemos um afastamento do equilíbrio não detectado por HD, mantendo-se as ressalvas já feitas. Contudo, se levarmos em conta o efeito de *infall* associado aos objetos de baixo fluxo seria de se esperar algum grau de afastamento do equilíbrio do sistema que pode estar sendo descrito pela diminuição dos valores de  $q$  observados.

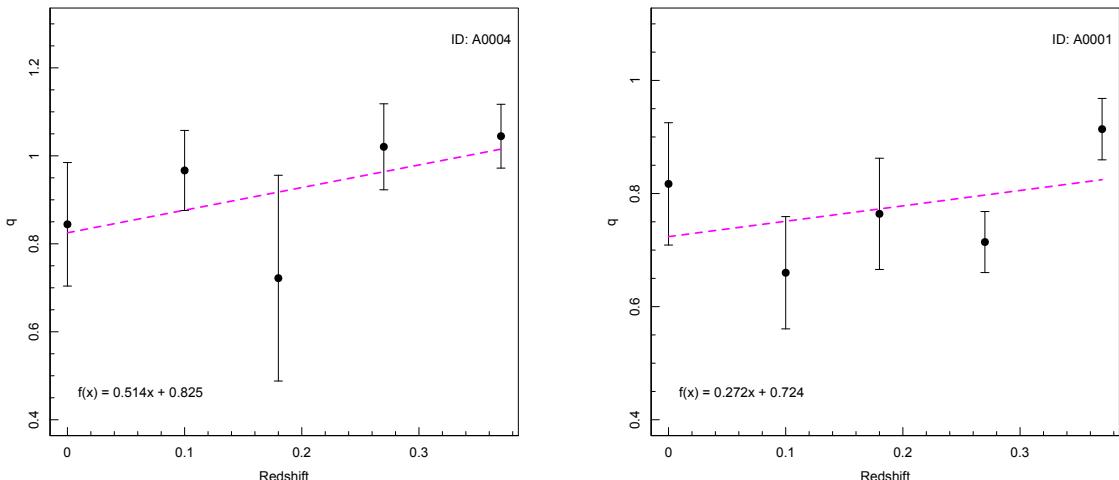


Figura 34 – Evolução do índice  $q$  da distribuição q-Gaussiana para os dados do EAGLE com corte em magnitude  $M_r \leq -17$ .

Nossos resultados, no conjunto, mostram que as ferramentas de análise desenvolvidas neste trabalho podem ser utilizadas em estudos que visem avaliar o estado dinâmico de aglomerados de galáxias. Uma integração maior entre os métodos descritos poderá restringir significativamente a dinâmica desses sistemas. Os resultados apresentados aqui indicam que é possível acessar mudanças importantes nas propriedades cinemáticas e dinâmicas dos aglomerados em diferentes redshifts. Ou seja, objetivo central do trabalho foi alcançado.

## 5 CONCLUSÕES E PERSPECTIVAS

Neste trabalho objetivamos a implementação, em R, de ferramentas que possibilissem o diagnóstico e o ajuste dos perfis de densidade e a distribuição de velocidades das galáxias em aglomerados. Ao final do trabalho, conseguimos implementar, testar e aplicar aos dados todas as ferramentas que tínhamos como meta. Implementamos também scripts para automatizar a obtenção dos dados das simulações Millennium e EAGLE. Tais scripts realizaram diversas consultas, onde foram aplicadas as funções de seleção com o objetivo de separar as amostras conforme as características que foram estudadas. A Figura 35 apresenta o esquema da integração entre ferramentas desenvolvidas e os dados das simulações.

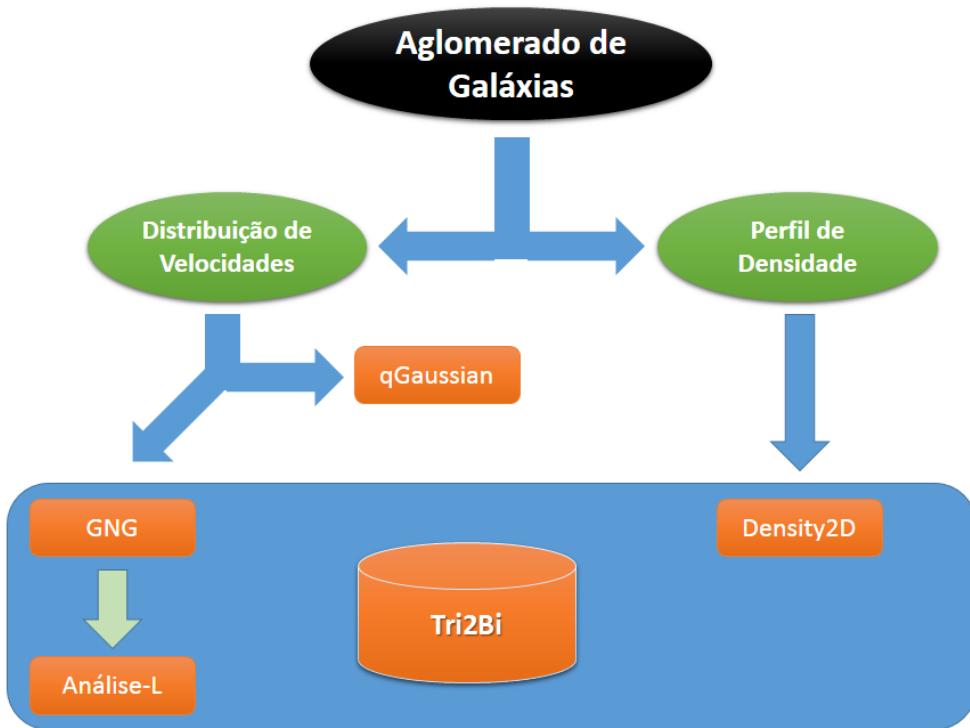


Figura 35 – Esquema da integração das ferramentas desenvolvidas com os dados utilizados.

A seguir fazemos um síntese do trabalho, apresentando também as nossas conclusões.

## 5.1 Síntese e conclusões

Utilizamos, para a distribuição de velocidades, a ferramenta GNG que aplica a distância de Hellinger como métrica para estabelecer a gaussianidade de um aglomerado e, consequentemente, seu estado de equilíbrio. A ferramenta GNG sofreu algumas alterações durante os anos. Em sua proposta inicial, presente em Ribeiro et al. (2013), ela sempre fornecia um diagnóstico de gaussianidade ou não-gaussianidade. Subsequentemente, de Carvalho et al. (2017) introduziram o diagnóstico indeterminado, mostrando a compreensão de que muitos dos aglomerados não possuem um diagnóstico, mas pelo mesmo motivo faz com que, muitas vezes, grande parte da amostra seja ignorada. Aqui propomos uma ferramenta, chamada Análise-L, que utiliza o método dos L-Momentos para recuperar alguns dos aglomerados classificados como indeterminados, algo que possibilita a reintrodução desses objetos na amostra. Por exemplo, quando aplicamos a Análise-L aos dados do Millennium, em  $z = 0.0$ , recuperamos cerca de 75% dos objetos para a classificação gaussiana. Sabemos que a utilização da Análise-L não resolve o problema do diagnóstico indeterminado, mas se apresenta como um grande avanço nesse sentido.

Como utilizamos simulações computacionais para testar os programas, vimos a necessidade de desenvolver uma ferramenta que possibilitasse a conversão de coordenadas tridimensionais, advindas de aglomerados simulados, para bidimensionais, projetando-as na linha de visada e no plano do céu. Dessa forma, foi possível, por exemplo, aplicar o GNG e o density2D para dados simulados, tornando possível realizar análises em um ambiente controlado. A ferramenta desenvolvida considera o redshift e os parâmetros da cosmologia, realizando a projeção, e permitindo comparações diretas com resultados da aplicação do GNG e do density2D a dados oriundos de observação.

Investigamos, também, os desvios da gaussianidade encontrados na distribuição de velocidades de galáxias em aglomerados através da função anisotrópica q-gaussiana, proposta por Bernaldo et al. (2013a). Implementamos um conjunto de programas que permitem estimar o parâmetro  $q$  e simular dados que sigam a distribuição q-gaussiana. Diferentemente do trabalho de (Bernaldo et al., 2013a), analisamos os dados em diversos redshifts, tentando acompanhar a evolução do parâmetro  $q$  com o decorrer das épocas. Os exemplos de aplicação da q-gaussiana ao dados do Millennium e do EAGLE sugerem que esta pode ser uma ferramenta extremamente útil na determinação de estados evolutivos intermediários dos aglomerados. Enquanto o GNG classifica um sistema como gaussiano, não-gaussiano, ou indeterminado (sendo a interpretação usual de um sistema gaussiano como estando em equilíbrio), o código qGaussian ajusta o parâmetro  $q$  da distribuição generalizada, com os valores de  $q$  sendo associados a estados estacionários do sistema.

Para os perfis de densidade, foi desenvolvida uma ferramenta, chamada den-

sity2D, que implementa rotinas para estimar os parâmetros encontrados nos modelos de Navarro-Frenk-White, King e Hernquist, utilizando os métodos da máxima verossimilhança e do  $\chi^2$ . Também foram desenvolvidas rotinas para a geração de raios projetados até o centro do aglomerado seguindo cada um dos perfis. O uso de ajustes de perfis de densidade a aglomearados não é novo, mas a proposta de usá-lo em diferentes redshifts para acessar a evolução dinâmica dos sistema, associado aos diagnósticos vindo de GNG e qGaussian, podem vir a representar um avanço neste tipo de análise. A aplicação aos dados do Millennium sugerem que sistemas não-gaussianos possuem um excesso de densidade nas regiões mais externas em relação a sistemas gaussianos. Este seria um traço de sistemas fora do equilíbrio, e portanto um indicador da dinâmica dos aglomerados. A aplicação aos dados do Eagle mostraram que para um mesmo aglomerado há uma variação significativa de seu perfil de densidade ao longo do tempo. Estudos envolvendo amostras maiores de dados simulados podem revelar possíveis tendências nessas variações, o que tornaria este tipo de análise também um indicador da dinâmica do sistema.

A seguir, relacionamos alguns itens que entendemos como desdobramentos naturais do presente trabalho.

## 5.2 Perspectivas

Como continuidade da pesquisa, sugerimos

- A investigação do perfil de anisotropia das velocidades para aglomerados de galáxias, algo que tornaria possível estabelecer com maior precisão quais galáxias ligadas ao potencial do aglomerado estão em equilíbrio, qual seu tempo dentro do sistema e sua posição no espaço de fase.
- Um estudo mais detalhado sobre os casos onde aglomerados diagnosticados como não-gaussianos que estão inseridos na região de aceitação no diagrama de L-momentos, possivelmente com a aplicação da técnica proposta por [Zabludoff et al. \(1993\)](#) que utiliza a expansão de Gauss-Hermite para medir a distribuição de velocidades, apresentando pouca sensibilidade a galáxias que possuem velocidade distante da média da distribuição.
- Outro passo importante consiste no emprego da função anisotrópica q-gaussiana na linha de visada, possibilitando a estimativa de  $q$  em dados observados, e uma possível integração entre o diagnóstico do GNG com o valor estimado de  $q$ .
- No Capítulo 4, vimos que a função kde.test não apresentou resultados satisfatórios na avaliação do espaço de fase, quando aplicado a aglomerados individuais. Em

decorrência deste problema, sugerimos a obtenção do perfil projetado do espaço de fase utilizando o método proposto por [Wojtak \(2013\)](#), em que temos

$$P_2(\sigma_{los})d\sigma_{los} = \frac{\sigma_{los}}{\sqrt{\sigma_{los}^2(1+2q_v^2) - bq_v^2}} \frac{1+2q_v^2}{\sqrt{b(1-q_v^2)}} d\sigma_{los},$$

onde  $b$  corresponde ao parâmetro de viés,  $\sigma_{los}$  é a dispersão de velocidades normalizada na linha de visada, e  $q_v$  é a taxa de velocidades elipsoidais.

- Para os dados do EAGLE, o processo de rastreamento dos aglomerados em diferentes épocas deve ser otimizado, tendo em vista a grande dificuldade em acompanhar os mesmos objetos por conta dos processos de *infall* e fusões, assim como pela falta de identificação, por parte da simulação, dos aglomerados e seus "progenitores".
- Os dados do EAGLE e Millennium podem e serão utilizados para estudos dos perfis de dispersão de velocidades como proposto por [Costa et al. \(2018\)](#).

## Referências

- Adami, C.; Mazure, A.; Katgert, P.; Biviano, A. The ESO nearby Abell cluster survey. VII. Galaxy density profiles of rich clusters of galaxies. **Astronomy and Astrophysics**, v. 336, p. 63–82, ago. 1998.
- Bartelmann, M. Arcs from a universal dark-matter halo profile. **Astronomy and Astrophysics**, v. 313, p. 697–702, set. 1996.
- BAUER, A. E.; DRORY, N.; HILL, G. J.; FEULNER, G. Specific star formation rates to redshift 1.5. **The Astrophysical Journal Letters**, v. 621, n. 2, p. L89, 2005. Disponível em: <<http://stacks.iop.org/1538-4357/621/i=2/a=L89>>.
- BENTO, E. P. **Fundamentação Cinética da Estatística Não-Gaussiana: Efeitos em Politrópicas**. Tese (Doutorado) — Universidade Federal do Rio Grande do Norte, Natal, Brasil, 2011.
- Bernaldo, L. J.; Mamon, G.; Duarte, M.; Peirani, S.; Boué, G. Anisotropic q -gaussian 3d velocity distributions in lambda cdm haloes. v. 452, 10 2013.
- Bernaldo, L. J.; M.Lima; L.Sodre. Testing phenomenological and theoretical models of dark matter density profiles with galaxy clusters. **Mon. Not. Roy. Astron. Soc.**, v. 436, p. 2616, 2013.
- BRENT, R. P. **Algorithms for Minimization without Derivatives**. Englewood Cliffs, N.J.: Prentice-Hall, 1973.
- CAVA, A.; BIVIANO, A.; MAMON, G. A.; VARELA, J.; BETTONI, D.; D'ONOFRIO, M.; FASANO, G.; FRITZ, J.; MOLES, M.; MORETTI, A.; POGGIANTI, B. Structural and dynamical modeling of WINGS clusters. I. The distribution of cluster galaxies of different morphological classes within regular and irregular clusters. **Astronomy and Astrophysics**, 2017.
- CLOWE, D.; BRADAč, M.; GONZALEZ, A. H.; MARKEVITCH, M.; RANDALL, S. W.; JONES, C.; ZARITSKY, D. A Direct Empirical Proof of the Existence of Dark Matter. **Astrophysical Journal Letters**, v. 648, p. L109–L113, set. 2006.
- COSTA, A. P.; RIBEIRO, A. L. B.; CARVALHO, R. R. de. The shape of velocity dispersion profiles and the dynamical state of galaxy clusters. **Monthly Notices of the Royal Astronomical Society: Letters**, v. 473, n. 1, p. L31–L35, 2018. Disponível em: <<http://dx.doi.org/10.1093/mnrasl/slx156>>.
- CRAIN, R. A.; Schaye, J.; Bower, R. G.; Furlong, M.; Schaller, M.; Theuns, T.; Dalla Vecchia, C.; Frenk, C. S.; McCarthy, I. G.; Helly, J. C.; Jenkins, A.; Rosas-Guevara, Y. M.; White, S. D. M.; Trayford, J. W. The EAGLE simulations of galaxy formation: calibration of subgrid physics and model variations. **Monthly Notices of the Royal Astronomical Society**, v. 450, p. 1937–1961, jun. 2015.
- de Carvalho, R. R.; RIBEIRO, A. L. B.; STALDER, D. H.; ROSA, R. R.; COSTA, A. P.; MOURA, T. C. Investigating the relation between galaxy properties and the gaussianity

- of the velocity distribution of groups and clusters. **The Astronomical Journal**, v. 154, n. 3, p. 96, 2017. Disponível em: <<http://stacks.iop.org/1538-3881/154/i=3/a=96>>.
- De Lucia, G.; Blaizot, J. The hierarchical formation of the brightest cluster galaxies. **Monthly Notices of the Royal Astronomical Society**, v. 375, p. 2–14, fev. 2007.
- Elmasri, R. A.; NAVATHE, S. B. **Fundamentals of Database Systems**. 3rd. ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1999. ISBN 0805317554.
- Fukushige, T.; Makino, J. Structure of dark matter halos from hierarchical clustering. ii. dependence of cosmological models in cluster-sized halos. **The Astrophysical Journal**, v. 588, n. 2, p. 674, 2003. Disponível em: <<http://stacks.iop.org/0004-637X/588/i=2/a=674>>.
- GUNN, J. E.; GOTTF III, J. R. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. **The Astrophysical Journal**, v. 176, p. 1, ago. 1972.
- Hansen, S. H.; Moore, B.; Zemp, M.; Stadel, J. A universal velocity distribution of relaxed collisionless structures. **Journal of Cosmology and Astroparticle Physics**, v. 2006, n. 01, p. 014, 2006. Disponível em: <<http://stacks.iop.org/1475-7516/2006/i=01/a=014>>.
- HERNQUIST, L. An analytical model for spherical galaxies and bulges. v. 356, p. 359–364, 06 1990.
- Hosking, J. R. M.; Wallis, J. R. **Regional Frequency Analysis: An Approach Based on L-Moments**. Cambridge University Press, 1997. ISBN 0521430453. Disponível em: <[http://www.amazon.de/Regional-Frequency-Analysis-Approach-L-Moments/dp/0521430453/ref=sr\\_1\\_2?ie=UTF8&qid=1317235930&sr=8-2](http://www.amazon.de/Regional-Frequency-Analysis-Approach-L-Moments/dp/0521430453/ref=sr_1_2?ie=UTF8&qid=1317235930&sr=8-2)>.
- King, I. The structure of star clusters. I. an empirical density law. **The Astrophysical Journal**, v. 67, p. 471, out. 1962.
- King, I. R. Density Data and Emission Measure for a Model of the Coma Cluster. **Astrophysical Journal Letters**, v. 174, p. L123, jun. 1972.
- KRAVTSOV, A. V.; BORGANI, S. Formation of Galaxy Clusters. **The Annual Review of Astronomy and Astrophysics**, n. 50, p. 353–409, 2012.
- LIMOUSIN, M.; MORANDI, A.; SERENO, M.; MENEGHETTI, M.; ETTORI, S.; BARTELMANN, M.; VERDUGO, T. The three-dimensional shapes of galaxy clusters. **Space Sci Rev**, p. 155–194, 2013.
- Liou, J.-J.; Wu, Y.-C.; Cheng, K.-S. Establishing acceptance regions for L-moments based goodness-of-fit tests by stochastic simulation. **Journal of Hydrology**, v. 355, p. 49–62, jun. 2008.
- LYNDEN-BELL, D. Statistical mechanics of violent relaxation in stellar systems. **Mounthly Notices of Royal Astronomical Society**, v. 136, p. 101–121, 1967.
- Mamon, G. A.; Biviano, A.; Boué, G. Mamposst: Modelling anisotropy and mass profiles of observed spherical systems – i. gaussian 3d velocities. **Monthly Notices of the Royal Astronomical Society**, v. 429, n. 4, p. 3079–3098, 2013. Disponível em: <<http://dx.doi.org/10.1093/mnras/sts565>>.

- MANDELBAUM, R.; SELJAK, U.; BALDAUF, T.; SMITH, R. E. Precision cluster mass determination from weak lensing. **Monthly Notices of the Royal Astronomical Society**, v. 405, n. 3, p. 2078–2102, 2010. Disponível em: <<http://dx.doi.org/10.1111/j.1365-2966.2010.16619.x>>.
- NATAJARAN, P.; HJORTH, J.; van KAMPEN, E. Distribution functions for clusters of galaxies from N-body simulations. **Monthly Notices of the Royal Astronomical Society**, v. 286, p. 329–343, abr. 1997.
- NAVARRO, J. F.; FRENK, C. S.; WHITE, S. D. M. The Structure of Cold Dark Matter Halos. **Astrophysical Journal Letters**, v. 462, p. 563, maio 1996.
- NETO, G. B. L.; LAGANA, T. F.; ANDRADE-SANTOS, F.; MACHADO, R. E. G. Structure in galaxy clusters. **ArXiv e-prints**, 2014.
- Ogorodnikov, K. F. On the Principles of Statistical Mechanics of Stellar Systems. **sovast**, v. 1, p. 787, dez. 1957.
- Oman, K. A.; HUDSON, M. J.; BEHROOZI, P. Disentangling satellite galaxy populations using orbit tracking in simulations. v. 431, 01 2013.
- REMBOLD, S. B. **Propriedades Dinâmicas e Populacionais do Aglomerado de Galáxias LCDCS-S001**. Tese (Doutorado) — Instituto de Física, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil, 2007. Disponível em: <<http://hdl.handle.net/10183/11190>>. Acesso em: 07 sep. 2017.
- Ribeiro, A. L. B.; de Carvalho, R. R.; Trevisan, M.; Capelato, H. V.; La Barbera, F.; Lopes, P. A. A.; Schilling, A. C. SPIDER - IX. Classifying galaxy groups according to their velocity distribution. **Monthly Notices of the Royal Astronomical Society**, v. 434, p. 784–795, set. 2013.
- ROMERO, M. J.; LAMBAS, D. An improved method for the identification of galaxy systems: measuring the gravitational redshift by dark matter haloes. **Monthly Notices of the Royal Astronomical Society**, v. 427, p. L6–L10, 2012.
- RUBIN, V. C.; FORD JR., W. K. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. **The Astrophysical Journal**, v. 159, p. 379, fev. 1970.
- SARAZIN, C. L. X-ray emission from clusters of galaxies. **Reviews of Modern Physics**, APS, v. 58, n. 1, p. 1, 1986.
- Schaye, J.; Crain, R. A.; Bower, R. G.; Furlong, M.; Schaller, M.; Theuns, T.; Dalla Vecchia, C.; Frenk, C. S.; McCarthy, I. G.; Helly, J. C.; Jenkins, A.; Rosas-Guevara, Y. M.; White, S. D. M.; Baes, M.; Booth, C. M.; Camps, P.; Navarro, J. F.; Qu, Y.; Rahmati, A.; Sawala, T.; Thomas, P. A.; Trayford, J. The EAGLE project: simulating the evolution and assembly of galaxies and their environments. **Monthly Notices of the Royal Astronomical Society**, v. 446, p. 521–554, jan. 2015.
- SPARKE, L. S.; III, J. S. G. **Galaxies in the Universe: An Introduction**. 2. ed. Cambridge University Press, 2007. ISBN 0521855934,9780521855938,0521671868,9780521671866. Disponível em: <<http://gen.lib.rus.ec/book/index.php?md5=490F4B2966A6E4749C9BF68F01A0AC87>>.

- Springel, V. Cosmological simulation code gadget. v. 364, p. 1105–1134, 12 2005b.
- Springel, V.; White, S. D. M.; Jenkins, A.; Frenk, C. S.; Yoshida, N.; Gao, L.; Navarro, J.; Thacker, R.; Croton, D.; Helly, J.; Peacock, J. A.; Cole, S.; Thomas, P.; Couchman, H.; Evrard, A.; Colberg, J.; Pearce, F. Simulations of the formation, evolution and clustering of galaxies and quasars. **Nature**, 2005.
- THOMAS, D. B.; LUK, W. Non-uniform random number generation through piecewise linear approximations. In: **IET Computers and Digital Techniques**. [S.l.: s.n.], 2007.
- TSALLIS, C. Possible generalization of boltzmann-gibbs statistics. **Journal of Statistical Physics**, v. 52, n. 1, p. 479–487, Jul 1988. ISSN 1572-9613. Disponível em: <<https://doi.org/10.1007/BF01016429>>.
- VOIT, G. M. Tracing cosmic evolution with clusters of galaxies. **Reviews of Modern Physics**, v. 77, p. 207–258, 2005.
- WILLIAMS, L. L. R.; HJORTH, J.; WOJTAK, R. Statistical Mechanics of Collisionless Orbits. III. Comparison with N-body Simulations. **apj**, v. 725, p. 282–287, dez. 2010.
- WOJTAK, R. Phase-space shapes of clusters and rich groups of galaxies. v. 559, 10 2013.
- Yahil, A.; Vidal, N. V. The Velocity Distribution of Galaxies in Clusters. **apj**, v. 214, p. 347–350, jun. 1977.
- Zabludoff, A. I.; Franx, M.; Geller, M. J. A Technique for Detecting Structure in Cluster Velocity Distributions. **apj**, v. 419, p. 47, dez. 1993.
- ZWICKY, F. Die Rotverschiebung von extragalaktischen Nebeln. **Helv. Phys. Acta**, v. 6, p. 110–127, 1933. [Gen. Rel. Grav.41,207(2009)].

## Apêndices

## APÊNDICE A – Distribuição q-Gaussiana

Diversos sistemas possuem comportamentos que não seguem a entropia de Boltzmann-Gibbs, o que pode ser resultado de interações de longo alcance dentro de sistemas em evolução. Assim, surgem as estatísticas generalizadas que tem como proposta trabalhar com sistemas com essas características através da utilização da mecânica estatística. Dentro deste grupo podemos citar o trabalho de [Tsallis \(1988\)](#) que lida com a entropia parametrizada através de um índice generalizado  $q$ .

Descreveremos, então, a distribuição q-gaussiana aplicada para modelar as moléculas de um gás. Para as componentes de velocidade, temos que a função de distribuição q-gaussiana é dada por ([BENTO, 2011](#))

$$f_q(v_i) = A_q \left[ 1 - (q-1) \frac{mv_i^2}{2k_B T} \right]^{\frac{1}{q-1}}$$

onde  $i$  corresponde as componentes  $x, y$  e  $z$ ,  $k_B$  é a constante de Boltzmann,  $m$  é a massa das partículas,  $T$  é a temperatura e  $A_q$  é dado por

$$A_q = \left( \frac{1+q}{2} \right) \frac{\Gamma(\frac{1}{2} + \frac{1}{q-1})}{\Gamma(\frac{1}{q-1})} \sqrt{\frac{m(q-1)}{2\pi k_B T}}$$

sendo a função  $\Gamma$  definida como

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt.$$

A distribuição de Maxwell é recuperada quando  $q \rightarrow 1$ :

$$F_q(v) = B_q \left[ 1 - (q-1) \frac{mv_i^2}{2k_B T} \right]^{\frac{1}{q-1}},$$

com

$$B_q = (q-1)^{1/2} \frac{(3q-1)}{2} \left( \frac{1+q}{2} \right) \frac{\Gamma(1/2 + 1/(q-1))}{\Gamma(1/(q-1))} \left( \frac{m}{2\pi k_B T} \right)^{3/2}.$$

## APÊNDICE B – Método Brent

Proposto por [Brent \(1973\)](#), o método Brent consiste em um método numérico para encontrar raízes de funções unidimensionais, combinando os métodos da bisseção, secante e interpolação quadrática inversa. Dentre suas vantagens, temos que sua convergência é sempre rápida e garantida, além de não necessitar de derivadas. Em contrapartida, não apresenta-se de maneira genérica, uma vez que é necessário que seja fornecido um intervalo que contenha a raiz.

O método recebe como entrada a função para se encontrar a raiz, os limites superior e inferior do intervalo que delimita o espaço de busca, a tolerância e o número máximo de iterações. Inicialmente, o método utiliza a interpolação quadrática inversa, enquanto não há sobreposição das aproximações intermediárias. Quando acontece a sobreposição, utiliza o método da secante para aproximar a raiz. Como os métodos anteriores não apresentam garantia de convergência, o algoritmo aplica o método da bisseção para novamente encontrar o zero da função. O Script [B.1](#) apresenta a implementação do método em linguagem R.

```

1 brent <- function(f, xl, xr, errto=0.001, imax=500)
2 {
3   errno = errto + 1
4   iterCount = 0
5   x = xl
6   y = xr
7   xrold = 0
8   mflag = True
9   delta = xr
10
11  if(f(xl) * f(xr) >= 0.0)
12  {
13    return(NA)
14  }
15
16  if(abs(f(xl)) < abs(f(xr)))
17  {
18    xr = xl
19    xl = xr
20  }
21
22  while(errno > errto && iterCount < imax)
23  {
24    f1 = f(xl)

```

```

25     fr = f(xr)
26     fx = f(x)
27
28 # try Inverse Quadratic Interpolation
29 if(f1 != fx && fr != fx)
30 {
31     delta = c()
32     delta[1] = f1 - fx
33     delta[2] = f1 - fr
34     delta[3] = fx - f1
35     delta[4] = fx - fr
36     delta[5] = fr - f1
37     delta[6] = fr - fx
38
39 disturb = function(ecs)
40 {
41     if(ecs == 0)
42     {
43         return errto
44     } else {
45         return ecs
46     }
47 }
48
49 delta = disturb(delta)
50
51 s = fx * fr * x1 / (delta[1] * delta[2]) +
52     f1 * fr * x / (delta[3] * delta[4]) +
53     f1 * fx * xr / (delta[5] * delta[6])
54
55 # try Secant's Rule
56 } else {
57     s = xr - fr*(xr - x1)/(fr - f1)
58 }
59
60 if (((!(s > (3 * x1 + xr)/4 && s < xr)) ||
61       (mflag && abs(s - xr) >= abs(xr - x)/2) ||
62       (!mflag && abs(s - xr) >= abs(x - y)/2) ||
63       (mflag && abs(xr - x) < abs(delta))) ||
64       (!mflag && abs(x - y) < abs(delta)))
65 {
66     # bisection method
67     s = (xr + x1)/2
68     mflag = True
69 } else {
70     mflag = False
71 }
```

```
72      fs = f(s)
73      y = x
74      x = xr
75
76      if (f1 * fs < 0.0)
77      {
78          xr = s
79      } else {
80          xl = s
81      }
82
83      if (abs(f(xl)) < abs(f(xr)))
84      {
85          xr = xl
86          xl = xr
87      }
88
89      if (fs == 0 || f(xr) == 0)
90      {
91          break
92      }
93
94      if (abs(xr - xrold) < delta)
95      {
96          delta = abs(xr - xrold)
97      }
98
99      xrold = xr
100     iterCount = iterCount + 1
101     errno = fabs(xr - xl)
102 }
103
104 return(xr)
105 }
```

Script B.1 – Implementação do método Brent em linguagem R.

## APÊNDICE C – Principais pacotes utilizados

Neste trabalho, utilizamos uma série de pacotes para os mais diversos propósitos. A seguir, listaremos os principais pacotes utilizados, além de descrever como foram aplicados no desenvolvimento do trabalho.

- astro
  - comovdist.los: Utilizado para converter o redshift em distância comóvel, na ferramenta *Tri2Bi*;
- bbmle
  - mle2: Aplicado na estimativa de parâmetros nos perfis de densidade superficial e distribuição q-Gaussiana;
- Lmoments
  - Lcoefs: Utilizado na ferramenta *Análise-L* para obter os L-momentos da distribuição gaussiana;
- distrEx
  - GLIntegrate: Aplicado no desenvolvimento da ferramenta *density2D*, realiza a integral numérica utilizando a quadratura gaussiana;
- minpack.lm
  - nlsLM: Aplicado no desenvolvimento da ferramenta *density2D*, realiza o ajuste não-linear robusto;
- cosmoFns
  - D.A: Utilizado na ferramenta *Tri2Bi* para computar a distância do diâmetro angular;

## APÊNDICE D – Termos astronômicos

- **Big Bang:** Inicialmente conhecida como a hipótese do átomo primordial, a teoria do Big Bang consiste na ideia de que o Universo originou-se a partir de um estado extremamente quente e denso, uma singularidade, há cerca de 13.8 bilhões de anos, e vem se expandindo desde então. Proposta por Georges Lemaître, baseando-se na teoria geral da relatividade, teve suas equações formuladas por Alexander Friedmann. Edwin Hubble observou em 1929 que as galáxias mais distantes estavam se afastando umas das outras mais rapidamente do que as mais próximas, corroborando a ideia de Lemaître sugerida em 1927. Dessa forma, em um momento inicial, toda a matéria presente no universo deveria estar comprimida em um único ponto (vide Figura 36);

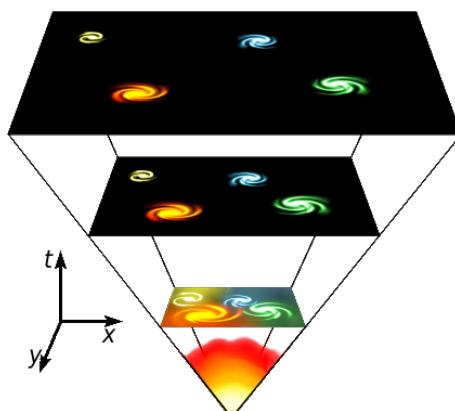


Figura 36 – Esquema da expansão do universo a partir de um ponto extremamente quente e denso. Fonte: Wikipedia

- **Lei de Hubble-Humason:** Consiste na relação observada entre nós e uma fonte emissora e a sua velocidade de afastamento. Edwin Hubble verificou que a maioria das galáxias apresentava um desvio para o vermelho em suas linhas espectrais e que esse desvio se tornava maior nas galáxias mais distantes. Então, ele propôs a equação

$$v = H_0 \times d,$$

onde  $v$  é a velocidade da fonte emissora em  $Kms^{-1}$ ,  $d$  é a distância dada em  $Mpc$  e  $H_0$  é o parâmetro de Hubble, em unidades de  $kms^{-1}Mpc^{-1}$ . O valor de  $H_0$  vem sendo constantemente calibrado em acordo com as observações da radiação cósmica de fundo. Os resultados encontrados pelo satélite Plank, em 2013, apontam o valor de  $H_0$  em  $67.9 \pm 1.5 kms^{-1}Mpc^{-1}$ .

- **Parâmetros Omega:** Diversos estudos mostram que o universo possui densidade muito próxima da densidade crítica, que poderia causar a sua eterna expansão. A densidade total do universo pode ser modelada através de uma única equação, dada por  $\Omega = \Omega_m + \Omega_{rel} + \Omega_\Lambda$ , onde  $\Omega$ ,  $\Omega_m$ ,  $\Omega_{rel}$  e  $\Omega_\Lambda$  representam, respectivamente, a densidade total do universo, densidade de massa, densidade de massa para partículas relativísticas, e densidade de massa para energia escura;
- **Redshift:** Ao observarmos as linhas espectrais obtidas de uma fonte emissora, é possível notar desvios em direção ao vermelho (redshift) ou ao azul (blueshift). Esses desvios podem ser causados por dois fenômenos distintos: o efeito Doppler, gerado pelas mudanças cinemáticas da fonte emissora; e o Redshift, causado pela expansão do universo. Quando o espectro da fonte emissora desloca-se para o azul, temos que há uma aproximação entre a fonte emissora e o observador; quando desloca-se para o vermelho, há o afastamento entre ambos. Tal afastamento pode ser resultado também da expansão do universo. Como apresentado anteriormente, quanto mais distante a fonte emissora encontra-se do observador, mais deslocado para o vermelho é seu espectro;