

CBPF

**Centro Brasileiro de
Pesquisas Físicas**

Tese de Doutorado

**Abundância de Aglomerados de Galáxias
como Observável Cosmológico:
Aplicações aos Levantamentos Fotométricos DES e SDSS**

Mariana Penna Lima

Orientador:

Dr. Martín MAKLER

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Mariana Penna Lima

*Abundância de Aglomerados de Galáxias
como Observável Cosmológico:
Aplicações aos Levantamentos
Fotométricos DES e SDSS*

Tese apresentada ao Programa de Pós- Graduação em Física, do Centro Brasileiro de Pesquisas Físicas para a obtenção do título de Doutor.

Orientador:
Martín Makler

CENTRO BRASILEIRO DE PESQUISAS FÍSICAS
PROGRAMA DE PÓS-GRADUAÇÃO

Rio de Janeiro, 30 de Setembro de 2010

Agradecimentos

Ao longo desses quatro anos encontrei novos amigos que muito me ajudaram nesta etapa tão importante da minha vida. Gostaria de agradecer aos amigos Álvaro Gomes, Bruno Mota, Cristina Furnalietto, Fernando Simoni, Gabriel Caminha, Gustavo Catão, Habib Montoya, Maria Elidaiana e Stella Pereira por todo o apoio e incentivo.

Agradeço ao Marcos Lima, Michael Busha, Eduardo Rozo e Juan Estrada pelas discussões.

Agradeço às pessoas do DES-Brazil pela oportunidade de participar de um projeto

Agradeço a Stella Pereira, Álvaro Gomes e Marcos Lima pela revisão do texto, comentários e sugestões.

Agradeço aos funcionários do CBPF. Um agradecimento especial à Cláudia, Luciene, Elisete, Elisabete e Ricardo e aos professores Nelson Pinto Neto e Ivan dos Santos Oliveira Júnior.

Agradeço ao CNPq e à CAPES pelo apoio financeiro.

Agradeço à minha família por todo suporte.

Agradeço ao professor Martín Makler que além de orientador tornou-se um amigo

Resumo

Os aglomerados de galáxias são uma fonte de informação para estudar modelos cosmológicos, em especial, a energia escura. Neste trabalho abordamos a abundância de aglomerados de galáxias como observável cosmológico. Introduzimos uma nova verossimilhança para a abundância de aglomerados, a partir da qual obtemos estimadores para os parâmetros cosmológicos. Além de não requerer a utilização de *bins* no desvio para o vermelho z , essa nova verossimilhança possui a vantagem conceitual de permitir a inclusão das incertezas em z e na massa M para cada aglomerado de forma individual. Utilizamos o método de Monte Carlo para obter o viés e a variância dos estimadores e verificamos que a verossimilhança proposta fornece resultados idênticos ou até ligeiramente melhores que a verossimilhança de Poisson em *bins* (e claramente superior ao método χ^2). Utilizamos o nosso método no contexto dos projetos *Dark Energy Survey* (DES) e *Sloan Digital Sky Survey III* (SDSS-III). Nós aplicamos esse método para obter previsões para os limites que podem ser impostos nos parâmetros cosmológicos a partir de amostras de aglomerados no SDSS-III. Também aplicamos a metodologia descrita para validar os catálogos de halos do DES. Finalmente, incluímos as distribuições de erros do desvio para o vermelho fotométrico e da relação massa-observável na nova verossimilhança e mostramos a importância de considerar corretamente essas distribuições na modelagem teórica para limitar a energia escura em futuros levantamentos fotométricos de grande área.

Abstract

Galaxy clusters are one of the most promising sources of information about cosmological models, and specially about dark energy. In this work we analyze the use of the galaxy cluster abundance as a probe of cosmological parameters. To this end we introduce a new likelihood for the cluster abundance, from which we obtain the maximum likelihood estimators. The proposed likelihood does not depend on any redshift binning, and provides a way of accounting for the redshift and mass observables uncertainties of each cluster. Using the Monte Carlo method to study the bias and variance of each estimator we find that our proposed likelihood provides comparable and slightly better results than the binned Poisson likelihood (and clearly better than the binned χ^2). We apply the proposed method in the context of the projects Dark Energy Survey (DES) and Sloan Digital Sky Survey III (SDSS-III). We predict the expected constraints on the cosmological parameters from the samples of the SDSS-III. We also apply the new method to validate the halo catalogs from DES. Finally, we include the photometric redshift error and also the mass-observable uncertainty in the new likelihood and show the importance of correct modeling these distributions on parameter estimation from the future photometric dark energy surveys.

Sumário

Resumo	ii
Abstract	iii
1 Introdução	1
1.1 Dark Energy Survey - DES	3
1.2 Simulações no DES - os desafios de dados	5
1.3 Sloan Digital Sky Survey - III	7
2 Função de Massa de Grupos e Aglomerados de Galáxias	8
2.1 O formalismo de Press & Schechter e extensões	9
2.2 Variância da flutuação do campo de densidade de matéria	12
2.3 Espectro de potência linear	14
2.4 Função Crescimento	15
2.5 Função de massa a partir de simulações de N-corpos	16
2.6 Conectando halos e aglomerados de galáxias	18
2.6.1 Desvio para o vermelho fotométrico	19
2.6.2 Relação Massa—Observável	19
2.7 Contagem de aglomerados de galáxias	20
2.8 Parametrização do modelo de energia escura	22
3 Verossimilhanças para Abundância de Aglomerados de Galáxias	24
3.1 Verossimilhanças binadas no desvio para o vermelho	24

3.2	Estudo do viés na equação de estado da energia escura para os métodos binados	26
3.2.1	Geração de uma realização da distribuição de aglomerados	26
3.2.2	Método de Monte Carlo para a abundância de aglomerados	27
3.2.3	Estimadores para w_0 nas estatísticas binadas	28
3.3	Verossimilhança para a abundância sem binagem no desvio para o vermelho	30
3.3.1	Comparação do método de Monte Carlo com a Matriz de Fisher e o Perfil de Verossimilhança	33
3.4	Limites para os parâmetros cosmológicos nos métodos de Poisson e ULC .	36
3.5	Verossimilhança para a abundância de aglomerados sem binagem na massa	38
4	Estudo da Combinação de Possíveis Catálogos de Aglomerados no SDSS	42
4.1	Comparação dos limites em Λ CDM para três amostras fictícias de aglomerados no SDSS	42
4.2	Combinação de duas amostras fictícias de aglomerados no SDSS	44
5	Validação do Catálogo de Halos do <i>Dark Energy Survey</i>	48
5.1	<i>Visible Cosmology Challenge</i>	48
5.2	Catálogo de halos	49
5.3	Completeza do catálogo de halos	51
5.4	Cosmologia com o catálogo de halos	54
6	Inclusão de Efeitos Observacionais nos Estimadores da Contagem de Aglomerados de Galáxias	60
6.1	Inclusão da incerteza do desvio para o vermelho fotométrico	60
6.2	Inclusão da incerteza na massa no cálculo da contagem de aglomerados . .	61
6.3	Inclusão de incertezas na verossimilhança	64
6.3.1	Efeito da incerteza no desvio para o vermelho	65
6.3.2	Efeito da incerteza na massa nos parâmetros cosmológicos	66

6.3.3	Efeito da combinação de erros nos parâmetros cosmológicos	67
7	Conclusões e Perspectivas Futuras	70
	Apêndice A – Métodos Estatísticos Utilizados	73
A.1	Método da Máxima Verossimilhança	73
A.2	Teste da Razão de Verossimilhança	75
A.3	Perfil da Verossimilhança	76
A.4	Matriz de Fisher	77
	Apêndice B – Otimização, Controle e Validação dos Códigos	80
B.1	Interpolação por Spline	81
B.2	Cálculo Numérico da Variância	82
B.2.1	Soluções Analíticas	85
B.3	Cálculo Numérico da Abundância de Aglomerados	87
B.4	Spline Bidimensional	88
B.5	Validação e Controle do Código	94
	Referências	98

1 *Introdução*

Os aglomerados de galáxias são “sondas” cosmológicas por excelência já que a sua formação está diretamente ligada à evolução das estruturas em grandes escalas [1, 2]. Por sua dinâmica ser determinada majoritariamente pela matéria escura, é possível estudar vários aspectos desses objetos considerando apenas a interação gravitacional. Esse fato torna relativamente simples modelá-los de “primeiros princípios”, isto é, a partir de simulações de N-corpos e modelagens teóricas. De fato, comparações entre grandes simulações numéricas com distintos métodos mostram que é possível reproduzir com boa precisão muitas características globais dos aglomerados, como a sua abundância em função da massa e a estrutura interna, entre outros [3, 4, 5, 6, 7]. Desse modo, há um mapeamento bem conhecido e controlado, tanto em termos de simulações [8, 9, 10] quanto através de modelagens semi-analíticas [11, 12], entre o modelo cosmológico subjacente e a abundância e a distribuição espacial dos aglomerados, além disso, de modo geral tem havido um acordo entre as características globais dos aglomerados nos dados observacionais e nas simulações [13, 14, 15, 16, 17, 18]. Por essas razões, os aglomerados têm sido considerados ferramentas extremamente úteis para construir observáveis capazes de limitar os modelos cosmológicos, inclusive as densidades de matéria e energia escuras, assim como a equação de estado dessa última. No entanto, para serem úteis para sondar a equação de estado da energia escura, é necessário um grande número de aglomerados ($\gtrsim 10^5$) que permita determinar sua distribuição até desvios para o vermelho relativamente altos ($\gtrsim 1$). Como os aglomerados são objetos raros, isso requer um levantamento com boa profundidade e grande cobertura angular. O *Dark Energy Survey* (DES) [19, 20] (veja a seção 1.1) oferece, pela primeira vez, a combinação desses dois fatores.

Para utilizar a abundância e a distribuição de aglomerados como testes cosmológicos, precisamos fazer a conexão entre a modelagem teórica desses objetos, que supõe o conhecimento preciso das suas massas e dos desvios para o vermelho, com os dados observacionais. Em primeiro lugar, a massa do aglomerado não é medida diretamente nas imagens. Essa medida é feita a partir de observações ou medidas complementares como o fluxo em Raios-

X [21, 22], efeito Sunyaev-Zel'dovich [23], medida do cisalhamento de lente gravitacional [24, 25] e teorema virial [26, 27]. Por isso é necessário obter relações entre a massa e um observável direto no levantamento, como por exemplo, a riqueza¹ [28]. Outro fator está ligado à medida do desvio para o vermelho, já que não é viável obter o espectro de todos os objetos nos levantamentos de grandes áreas. O DES irá utilizar a medida do fluxo em 5 bandas de cada objeto para obter uma estimativa do seu desvio para o vermelho (uma técnica conhecida como desvio para o vermelho fotométrico). Dessa forma, a incerteza dos desvios para o vermelho fotométricos precisa ser modelada e incluída no cálculo da contagem de aglomerados (veja por exemplo as referências [29, 30, 31]). O levantamento *Sloan Digital Sky Survey* (SDSS), por sua vez, além de fornecer o desvio para o vermelho fotométrico, também inclui a medida do desvio para o vermelho espectroscópico de um subconjunto das galáxias observadas, o que é muito útil para uma calibração mais precisa da relação dos desvios para o vermelho espectroscópicos e fotométricos. Uma boa parte desta tese foi voltada para o desenvolvimento de ferramentas a serem utilizadas nos projetos DES e SDSS e a fazer previsões sobre limites cosmológicos a serem impostos pela abundância de aglomerados obtida com os dados desses projetos. Por isso, dedicaremos as seções 1.1 e 1.3 a uma breve revisão de alguns aspectos do DES e SDSS que serão utilizados ao longo desta tese.

Além de uma modelagem que leve em conta incertezas na relação massa-observável e no desvio para o vermelho, precisamos de um conjunto de ferramentas estatísticas que possibilite utilizar os dados da abundância de aglomerados para estimar parâmetros cosmológicos e suas incertezas. Essas ferramentas também são úteis para fazer previsões sobre os limites esperados nos parâmetros de interesse para distintas estratégias/configurações dos levantamentos, podendo ser utilizadas para otimizá-los.

Neste trabalho, introduzimos uma nova verosimilhança para a abundância de aglomerados (baseada na própria curva teórica da distribuição de aglomerados no desvio para o vermelho), a partir da qual obtemos estimadores para os parâmetros cosmológicos. Além de não requerer a utilização de *bins* no desvio para o vermelho, essa nova verosimilhança possui a vantagem conceitual de permitir a inclusão das incertezas no desvio para o vermelho e massa para cada aglomerado de forma individual.

Esta tese está estruturada da forma que segue. Nas seções 1.1 e 1.2, nós descrevemos aspectos gerais do DES e das simulações geradas nesse projeto, respectivamente. Uma breve descrição do SDSS é apresentada na seção 1.3. No capítulo 2, apresentamos

¹Riqueza é o número de galáxias vermelhas, N_{gal} , pertencentes a um aglomerado com luminosidade acima de um limiar $L_*/2$.

os ingredientes necessários para o cálculo da função de massa, isto é, a densidade de aglomerados em função da massa e do desvio para o vermelho, e para a exploração de sua dependência cosmológica. Os capítulos 3, 4, 5 e 6 constituem a parte original deste trabalho. No capítulo 3, revisamos dois métodos estatísticos para restringir parâmetros cosmológicos a partir da abundância de aglomerados de galáxias e, em seguida, propomos uma nova verossimilhança para estimar esses parâmetros. Para comparar os métodos, estudamos o viés e a eficiência dos seus estimadores. Apresentamos, também, algumas previsões para a restrição dos parâmetros cosmológicos esperados para o DES a partir desses métodos. No capítulo 4, aplicamos o novo método para estudar a restrição dos parâmetros cosmológicos esperada para algumas configurações de catálogos de aglomerados do SDSS. No capítulo 5, apresentamos as análises realizadas para validar os catálogos de halos obtidos a partir de simulações produzidas pela colaboração DES. Já no capítulo 6, estendemos a nova verossimilhança para incluir os efeitos das incertezas no desvio para o vermelho fotométrico e na massa. Os resultados e perspectivas futuras são discutidos no capítulo 7. No apêndice A, apresentamos algumas ferramentas de estatística utilizadas neste trabalho. Finalmente, no apêndice B, descrevemos alguns procedimentos realizados para otimizar o tempo de execução dos códigos, apresentamos o controle da precisão dos mesmos e mostramos a validação de alguns códigos desenvolvidos neste trabalho.

1.1 Dark Energy Survey - DES

O projeto *Dark Energy Survey* (DES [20]) é uma colaboração de diversas instituições dos EUA, Espanha, Inglaterra, Brasil e Alemanha, que tem por objetivo sondar a natureza da energia escura. Para alcançar este objetivo, será colocada uma câmera de 500 megapixels, altamente sensível no vermelho, no telescópio Blanco (4m), no *Cerro Tololo International Observatory* (CTIO, Chile). O DES coletará dados durante 5 anos, a partir de 2011, utilizando 30% do tempo do Blanco. O que torna o DES um projeto de fronteira é o fato de mapear uma grande área do céu (5000 deg^2 , ou seja, aproximadamente $1/8$ da esfera celeste), em cinco bandas (g , r , i , z e Y) e com uma profundidade inédita para um levantamento desta cobertura ($g = 24.6$, $r = 24.1$, $i = 24.4$, $z = 23.8$ e $Y = 21.3$). Tudo isso resultará na enorme quantidade de dados da ordem de petabytes (aproximadamente 1.8 terabytes por noite de observação).

Após os 5 anos de operação, o DES terá deixado um legado para a comunidade astronômica que consiste de:

- imagens do levantamento fotométrico de uma grande área (5000 deg²), com excelente qualidade de imagem (*seeing* mediano de cerca de 0,8"), fotometria uniforme, com magnitudes calibradas com 1% de precisão;
- catálogos com todo tipo de objetos, desde estrelas e quasares a galáxias e aglomerados de galáxias, que poderão ser utilizados pela comunidade para diversos estudos;
- a câmera instalada no telescópio Blanco ficará montada no mesmo lugar, para ser utilizada pela comunidade após o survey.

A participação do Brasil no DES se dá através de um consórcio de pesquisadores ligados a diversas instituições brasileiras, denominado DES-Brazil. A contribuição do DES-Brazil no DES se dá de três maneiras: contribuição em infra-estrutura, financeira e em ciência. A contribuição em infra-estrutura consiste no desenvolvimento de três ferramentas: i) um *software* de validação dos dados quando são adquiridos pela câmera, que será utilizado no próprio observatório (*Quick Reduce*), ii) redução dos dados de um levantamento em operação que servirá para a calibração do DES (*PreCam*) e iii) desenvolvimento de um Portal científico. A função do Portal é reunir e disponibilizar, em um só local, vários conjuntos de dados do projeto e códigos de análises dos dados, provendo meios de serem executados e concatenados dentro desta infraestrutura, mantendo os resultados das análises e o histórico do que foi feito. A estrutura do Portal é utilizada rotineiramente pelos participantes do projeto, especialmente pelos desenvolvedores de códigos, que contam com uma infra-estrutura consistindo de:

- armazenamento do código, com permissões de leitura e/ou escrita apenas para os usuários/grupos escolhidos;
- sistema de versionamento do código, feito pelo sistema *git* [32], que permite sincronizar o repositório do programa em qualquer computador, desenvolver códigos em colaboração, criar “marcos” (*releases*) do código, sendo sempre possível retornar o código para este estágio. O *git* também permite desenvolver o código paralelamente (em “ramos”), mantendo a versão atual estável (e eventualmente corrigindo *bugs*) no ramo principal. Isto é útil para grandes mudanças, sendo fácil reverter e recuperar modificações. Todo o desenvolvimento do código é registrado, com todas as modificações de cada arquivo, incluindo comentários do desenvolvedor.
- além de armazenar o código é possível disponibilizá-lo para ser executado através do Portal, com o processamento sendo feito por clusters de computadores. Isto permite atingir alto desempenho computacional sem sobrecarregar máquinas individuais;

- sistema de *ticket*, para informar e acompanhar erros, *bugs* e melhorias solicitadas nos sistemas.

Os dados do projeto serão armazenados com cópia no Brasil. Isto facilitará o acesso aos dados pela colaboração brasileira e, após 1 ano, os dados estarão disponíveis para toda a comunidade.

A contribuição científica se dá através de grupos de trabalho, divididos por área. São ao todo 11 grupos de trabalhos científicos: estrutura em grande escala, aglomerados de galáxias, supernovas, efeito fraco de lenteamento, teoria, efeito forte de lenteamento, simulações, quasares, desvio para o vermelho fotométrico, evolução de galáxias e Via Láctea. Estes 11 grupos de trabalho cobrem um amplo espectro da ciência que poderá ser realizada com os dados do DES. Em particular, os grupos de estrutura em grande escala, aglomerados, supernovas e efeito fraco de lenteamento têm como foco principal, de maneiras independentes e complementares, determinar o parâmetro w da equação de estado da energia escura.

1.2 Simulações no DES - os desafios de dados

No projeto DES são geradas, a cada período de aproximadamente um ano, imagens simuladas com o máximo de semelhança possível com as que serão obtidas pelo levantamento, como *seeing*, dimensão do pixel, bandas, magnitudes limites, efeitos de atmosfera, do instrumento e dos detectores, etc. Estas simulações possuem 3 utilidades básicas. A primeira é testar todo o sistema de processamento de dados, o que inclui a sua redução, calibração fotométrica e astrométrica, coadição das imagens, identificação de objetos, armazenamento e distribuição dos produtos associados. Esses testes são fundamentais devido ao grande volume de dados que deverá ser processado pelo projeto. A segunda utilidade é ajudar no desenvolvimento dos códigos de análise científica, que devem ser capazes de reobter os parâmetros fiduciais da simulação, como os do modelo cosmológico e da energia escura. A terceira utilidade é aplicar os códigos de análise, já testados em simulações anteriores, em várias novas séries de simulações com diferentes parâmetros de entrada para reobter os diferentes modelos cosmológicos. Isso levará a determinar os erros estatísticos e sistemáticos de cada uma das 4 técnicas que permitirão obter limites sobre os parâmetros cosmológicos. O DES é o primeiro projeto a realizar simulações baseadas em características do levantamento com esse grau de realismo.

As simulações são realizadas aumentando o realismo a cada novo ciclo, de acordo com

a necessidade dos grupos de trabalho. Há 3 níveis no processo de simulações. No primeiro é elaborado um catálogo com as propriedades astrométricas (posições), fotométricas (magnitude nas bandas do projeto) e morfológicas (coeficientes *shapelets* [33] que descrevem a morfologia) de galáxias. A base da distribuição de galáxias é dada por simulações de N-corpos de matéria escura (*Carmen*, das simulações *LasDamas - Large Suite of Dark Matter Simulations* [34, 35]). As galáxias são distribuídas segundo a densidade local de partículas de matéria escura reproduzindo uma correlação local observada entre luminosidade, cor e densidade [36]. De acordo com as propriedades atribuídas às galáxias (magnitudes, etc), a morfologia destas é obtida a partir de catálogos do *Cosmic Evolution Survey* (COSMOS [37]). A morfologia das galáxias é gravada no catálogo utilizando a decomposição em *shapelets* (por exemplo, para as galáxias, a decomposição vai até ordem $n = 15$, com 136 coeficientes). Galáxias mais fracas (ou de fundo), não baseadas em partículas de matéria escura, são também adicionadas ao catálogo. A inserção de estrelas é feita através de um catálogo produzido pelo DES-Brazil utilizando o código *trilegal* [38]. O catálogo de cisalhamento (devido ao efeito fraco de lente gravitacional) é obtido a partir da distribuição de matéria escura da simulação de N-corpos e interpolado para a posição de cada galáxia. Ainda no nível do catálogo, é feita a convolução com a PSF no espaço de *shapelets*, o que modifica esses coeficientes [39].

No segundo nível das simulações, os catálogos são utilizados para gerar imagens correspondentes aos CCDs, considerando efeitos atmosféricos (transparência, extinção, *seeing*, brilho do céu), da ótica e características do CCD (efeitos de pixels ruins, resposta do CCD, saturação, etc.).

O terceiro nível é a redução das imagens para obter as imagens finais, como serão distribuídas. Esta etapa inclui calibração astrométrica e fotométrica, coadição de imagens, homogeneização da PSF, etc. Após todo o processo de redução, as imagens são combinadas e separadas em regiões quadradas com área de $0,723 \text{ deg}^2$. As imagens são gravadas no formato FITS [40].

Parte deste trabalho, especialmente o capítulo 5, foi desenvolvida no contexto do grupo de trabalho de aglomerados do DES, em particular no que se refere à validação do catálogo de halos², a ser utilizado nesse processo de comparação entre a análise nos dados simulados e a “tabela verdade” das simulações. Na realidade, o grupo de aglomerados está trabalhando atualmente com uma escala intermediária de realismo, sem incluir o processamento de imagens, como será descrito no capítulo 5.

²Nesta tese *halo* corresponde a uma estrutura formada somente por matéria escura, tendo uma relação de um para um com um aglomerado de galáxia.

1.3 Sloan Digital Sky Survey - III

O SDSS é atualmente o maior levantamento fotométrico e espectroscópico existente, tanto em área quanto em número de objetos. Ele foi iniciado no ano 2000 e continua em operação, entrando agora na fase III, com novos espectrógrafos em funcionamento. Nas fases I e II do projeto foi desenvolvido um programa extremamente bem sucedido de aglomerados de galáxias, levando à descoberta de milhares desses objetos, a diversas formas de calibração da massa e ao uso da abundância de aglomerados como um observável cosmológico [41, 42, 43].

As etapas I (2000-2005) e II (2005-2008) cobriram uma área de aproximadamente 8000 deg^2 e desvios para o vermelho até $\simeq 0,3$ fornecendo mapas que contém cerca de 930.000 galáxias [44]. O SDSS-III (2008-2014) consiste de 4 levantamentos sendo o *Baryon Oscillation Spectroscopic Survey* (BOSS) o levantamento (espectroscópico) com o foco na cosmologia [45]. Ela irá complementar a parte de imageamento do SDSS, estendendo uma área no sul de cerca de 3000 deg^2 . Além disso, serão obtidos novos desvios para o vermelho na área total $\sim 10600 \text{ deg}^2$ com um novo espectrógrafo focando em galáxias luminosas vermelhas e em quasares distantes. A amostra de galáxias consistirá de aproximadamente 1,5 milhões de objetos até $z \simeq 0,7$.

Através do *Brazilian Participation Group* do SDSS-III, temos acesso aos novos dados de imageamento e além dos dados de espectroscopia que já somam da ordem de 200.000 desvios para o vermelho (cerca de 10% do total a ser obtido nos próximos 4 anos). Esses dados serão utilizados para identificar novos aglomerados de galáxias e obter suas implicações cosmológicas. No capítulo 4 discutimos alguns limites nos parâmetros cosmológicos esperados dos catálogos de aglomerados e da sua combinação.

2 Função de Massa de Grupos e Aglomerados de Galáxias

A função de massa $n(M)$ é definida de modo que $n(M) dM$ fornece o número de aglomerados de matéria escura gravitacionalmente ligados, por unidade de volume (comóvel), no intervalo de massa entre M e $M + dM$. Esses aglomerados são associados aos halos proto-galácticos e aos grupos e aglomerados de galáxias. A função de massa tem inúmeras aplicações práticas. Como mencionado na introdução, é possível obter limites nos parâmetros cosmológicos, e em particular nos da energia escura, comparando estimativas teóricas da função de massa com os dados observacionais [46, 47, 48]. Além disso, a função de massa também é um teste poderoso para estabelecer limites sobre o espectro primordial das flutuações [49, 50] e pode ser usada também como um teste da teoria da gravitação [51]. Com a função de massa é possível estimar o viés dos aglomerados (relação entre o espectro das flutuações da matéria escura e da densidade de halos) [11, 52], a taxa de fusão dos halos virializados [53], o tempo de formação e o tempo de vida dessas estruturas [54].

Neste capítulo, apresentamos uma breve revisão da modelagem teórica da função de massa, com foco nos resultados de simulações de N-corpos, que são o padrão mais utilizado atualmente para obter a abundância de halos, fornecendo a base para as comparações com os dados observacionais. Discutiremos em detalhes os elementos que entram no cálculo da função de massa a partir de um dado modelo cosmológico e a implementação destes utilizada nesta tese. Também especificaremos a parametrização para a energia escura que será utilizada ao longo deste trabalho. Finalmente, mencionaremos dois aspectos fundamentais na conexão entre essa modelagem teórica e os dados observacionais, a saber a utilização de desvios para o vermelho fotométricos e a obtenção de uma relação entre a massa e observáveis de aglomerados.

2.1 O formalismo de Press & Schechter e extensões

Historicamente, o formalismo mais conhecido para o cálculo da função de massa é o proposto por Press & Schechter [55]. Como veremos, ele é baseado em uma aproximação simplista para a formação de halos: o colapso esférico. No entanto, esse foi o modelo que predominou até a década de 1990, quando simulações numéricas de N-corpos de alta resolução permitiram obter parametrizações mais acuradas para a função de massa. O modelo de Press & Schechter tem o mérito de exibir explicitamente uma propriedade de *universalidade* da função de massa em relação ao modelo cosmológico. Essa universalidade serviu de guia para modelar a função de massa obtida a partir de simulações para diversas cosmologias e é um dos fundamentos de quase toda a análise da abundância de aglomerados com fins cosmológicos, em particular a que será adotada nesta tese.

O *Ansatz* de Press & Schechter (PS) consiste em supor que a função de massa $n(M)$ é obtida a partir da fração $F(M; z)$ de objetos colapsados em z com massa maior que M pela relação

$$\frac{dn(M, z)}{dM} = -\frac{\bar{\rho}_m(z)}{M} \frac{dF}{dM}. \quad (2.1)$$

Precisamos agora de uma forma de calcular F . Seguindo o modelo do colapso esférico [56], supõe-se que, se o contraste de densidade médio¹ δ_R numa região de tamanho R for maior que uma certa densidade crítica $\delta_c(z)$, essa região terá colapsado, formando uma estrutura ligada em z . Apesar da dinâmica nessa região ser não-linear, o processo de condensação não alterará a sua massa de maneira significativa. Então, se estimarmos o número de regiões com $\delta_R > \delta_c$ num campo aleatório, seremos capazes de determinar $F(M)$.

É possível mostrar que a distribuição das flutuações na escala R , $P(\delta_R)$ segue a mesma distribuição do contraste de densidade δ filtrada por uma função janela com essa escala. Para uma distribuição gaussiana, a probabilidade de que esse campo tenha um valor δ num dado ponto é

$$P(\delta_R) = \frac{1}{\sqrt{2\pi\sigma_R^2(z)}} \exp\left(-\frac{\delta_R^2}{2\sigma_R^2(z)}\right).$$

O cálculo da variância filtrada na escala R (σ_R^2) a partir do espectro de potência das flutuações é discutido na seção 2.2. O efeito do filtro é fazer uma média sobre as flutuações

¹O contraste de densidade local é definido por

$$\delta(\vec{x}, z) = (\rho_m(\vec{x}, z) - \bar{\rho}_m(z)) / \bar{\rho}_m(z), \quad (2.2)$$

onde $\bar{\rho}_m(z)$ é a densidade média de matéria do universo em z e independe de \vec{x} dada a homogeneidade do espaço em grandes escalas.

menores que R , para associar um valor de δ a essa região de tamanho R . Essa escala é associada à massa M através de R pela relação

$$M = \frac{4\pi}{3} f_W R^3 \bar{\rho}.$$

O fator f_W depende da forma da função filtro, para um filtro tipo cartola (*top-hat*) temos $f_W = 1$. Como a largura da janela no espaço k é da ordem da inversa da largura no espaço físico, grandes massas corresponderão a pequenos valores de σ_R^2 , enquanto pequenas massas correspondem a σ_R^2 grande. Note que a variável adequada para descrever o número de objetos colapsados é mesmo a massa, pois esta é conservada mesmo que o tamanho da região tenha variado enormemente. Não seria possível estimar o número de objetos de tamanho L , por exemplo, no formalismo de PS.

Supondo que a região com $\delta_R > \delta_c(z)$ vai formar um objeto gravitacionalmente ligado de massa M , a fração de objetos com massa maior que M será

$$\begin{aligned} F(M) &= \int_{\delta_c}^{\infty} P(\delta_R) d\delta = \frac{1}{\sqrt{2\pi\sigma_R^2}} \int_{\delta_c}^{\infty} \exp\left(-\frac{\delta_R^2}{2\sigma_R^2}\right) d\delta \\ &= \frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{\delta_c(z)}{\sqrt{2}\sigma_R(z)}\right) \right], \end{aligned} \quad (2.3)$$

onde $\operatorname{erf}(x)$ é a função erro². Utilizando a equação (2.1), obtemos a função de massa de PS:

$$\frac{dn(M, z)}{dM} = -\frac{\bar{\rho}_m(z)}{M} \frac{1}{\sqrt{2\pi}} \frac{\delta_c(z)}{\sigma_R(z)} \left(\frac{1}{\sigma_R(z)} \frac{d\sigma_R(z)}{dM} \right) \exp\left(-\frac{\delta_c^2(z)}{2\sigma_R^2(z)}\right). \quad (2.5)$$

O critério para determinar se um objeto colapsou formando um objeto ligado é baseado no campo *inicial* de densidades, de modo que $\delta(z)$ e $\sigma_R^2(z)$ são calculados no regime linear. Nesse caso $\delta(z) = D(z)\delta_0$ e $\sigma_R(z) = D(z)\sigma_R^0$, onde $D(z)$ é o fator de crescimento linear normalizado de modo que $D(z=0) = 1$ (veja a seção 2.4), δ_0 denota o valor da perturbação, linearmente extrapolada para $z=0$, e σ_R^0 é a variância filtrada na escala R linearmente extrapolada para $z=0$. Assim, a razão $\delta_c(z)/\sigma_R(z)$ que aparece na função de massa não depende de z . Vamos agora determinar o valor de δ_0^c correspondendo a δ_c .

Press & Schechter basearam-se no modelo do colapso esférico para determinar $\delta_c(z)$. No caso de um Universo plano com $\Omega_m = 1$ (Einstein-de Sitter), é possível mostrar que $\delta_c(z) = 1,686D(z) = \delta_c^0 D(z)$. No entanto, a flutuação de densidade crítica depende

²A função erro é definida como

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt. \quad (2.4)$$

fracamente da densidade de matéria (veja referências [57, 58]) e, conseqüentemente, da cosmologia. Assim, podemos escrever a função de massa de PS somente em função de δ_c^0/σ_R^0 e de $(d\sigma_R^0)/dM)/\sigma_R^0$. Essa é uma propriedade importante da função de massa de PS: a *auto-similaridade no tempo*.

Há ainda um problema fundamental com a equação (2.5). A integral de n sobre todas as massas deveria ser igual a 1, mas é fácil ver que

$$\int_0^\infty n(M) dM = \int_0^\infty dF = \frac{1}{2}.$$

Esse é conhecido como o problema da normalização da função de massa. A solução encontrada por Press & Schechter foi multiplicar a função de massa pelo fator *ad-hoc* de normalização 2.

O valor subestimado de objetos encontrado por eles é devido ao fato das subregiões de regiões menos densas (*clouds in clouds*) não serem levadas em conta corretamente nesse formalismo. Na referência [59], Bond et al. obtiveram a expressão correta considerando a probabilidade de se obter objetos acima de um limiar $\delta_c(z)$ para uma dada escala da massa entre M e $M + dM$, mas abaixo do limiar para os valores maiores de massa (veja também [60, 61]).

Outro problema do formalismo original de PS é a utilização do contraste de densidade δ como único critério para definir se uma certa região colapsará em $z = 0$, usando o modelo do colapso esférico para determinar δ_c . Diversas abordagens para ir além do colapso esférico foram propostas na literatura, além de redefinições do valor de $\delta_c^0(z)$. No entanto, durante a última década, a forma mais acurada para determinar a função de massa tem sido obtida a partir de ajustes (*fits*) às medições da abundância de halos realizadas em simulações de N-corpos. Esses ajustes são largamente inspirados em uma propriedade da função de massa de PS conhecida como *universalidade*.

De fato, olhando para a equação (2.5) vemos que a cosmologia não aparece explicitamente e nem tampouco o desvio para o vermelho z . Podemos escrever a (2.5) na forma

$$\frac{dn(M, z)}{dM} = -\frac{\bar{\rho}_m(z)}{M} f(\sigma_R(z)) \frac{1}{\sigma_R(z)} \frac{d\sigma_R(z)}{dM}, \quad (2.6)$$

onde

$$f_{PS}(\sigma_R) = \sqrt{\frac{2}{\pi}} (\delta_c \sigma_R^{-1}) \exp[-\delta_c^2 \sigma_R^{-2}/2]. \quad (2.7)$$

Toda a dependência do processo de colapso e formação do halo está contida na função f_{PS} (conhecida como *função de multiplicidade*), desacoplando-se da dependência cosmológica

através de ρ_m e σ_R (cujo cálculo veremos em detalhes a seguir).

Essa propriedade de universalidade é observada em outras extensões do formalismo de PS que vão além do colapso esférico e do método estatístico utilizado na abordagem original. Por exemplo, utilizando um método de passeio aleatório com barreira móvel em conjunto com um modelo de colapso elipsoidal, Sheth e Tormen (ST) [11, 12] obtiveram a expressão (2.6), mas com uma forma diferente para a função multiplicidade:

$$f_{ST}(\sigma) = A_{ST} \sqrt{\frac{2a}{\pi}} [1 + (a\delta_c\sigma^{-1})^{-p}] (\delta_c\sigma^{-1}) \exp[-(a\delta_c\sigma^{-1})^2/2], \quad (2.8)$$

onde $A_{ST} = 0,3222$, $a = 0,707$ e $p = 0,3$.

Outras extensões do método de PS envolvendo aproximações para a evolução não linear [62] também exibem a mesma forma da equação (2.6) para a função de massa, mas com outras funções de multiplicidade.

Essa propriedade de universalidade³ permite escalonar uma mesma expressão para diversas cosmologias e é ideal para utilizar a função de massa como observável para limitar modelos cosmológicos. Além dessas e outras abordagens semi-analíticas, as simulações cosmológicas corroboram a expressão (2.6) o que fornece uma base sólida para a sua utilização em cosmologia.

Antes de descrever como é obtida a função de massa em simulações, vamos detalhar alguns ingredientes necessários para o cálculo da função de massa de PS e que também serão fundamentais no cálculo das versões mais atuais da função de massa. Em particular, vamos detalhar como obter a variância filtrada $\sigma_R(z)$, que depende da função janela, do espectro das flutuações primordiais e da função de crescimento.

2.2 Variância da flutuação do campo de densidade de matéria

A função de massa é estimada a partir de propriedades estatísticas das flutuações do campo $\rho_m(\vec{x}, z)$ em um volume de raio comóvel R . Em geral, os modelos supõem que essas flutuações são inicialmente pequenas, $\delta(\vec{x}, z_i) \ll 1$, e se distribuem como uma Gaussiana. Da equação (2.2), vemos que a média do contraste de densidade é zero, $\langle \delta(\vec{x}, z) \rangle = 0$, e, portanto, devido à Gaussianidade, temos que $\delta(\vec{x}, z)$ é determinada completamente pela

³Apesar da função multiplicidade ter essa característica de universalidade, a abundância de halos (ou aglomerados de galáxias) é sensível à cosmologia, como mostrado na seção 2.7.

sua variância,

$$V(\delta(\vec{x}, z)) = \langle [\delta(\vec{x}, z)]^2 - \langle \delta(\vec{x}, z) \rangle^2 \rangle = \langle [\delta(\vec{x}, z)]^2 \rangle. \quad (2.9)$$

Com isso, temos que a variância do contraste de densidade $\delta(\vec{x}, z)$ filtrado em um volume comóvel pela função janela $W(\vec{x}, R)$ é,

$$\sigma^2(R, z) = \int d^3y \int d^3y' W(\vec{y}, R) W(\vec{y}', R) \langle \delta(\vec{x} + \vec{y}, z) \delta(\vec{x} + \vec{y}', z) \rangle. \quad (2.10)$$

Reescrevendo o lado direito da equação (2.10) em termos da transformada de Fourier espacial do contraste de densidade $\delta(\vec{k}, z)$, obtemos

$$\sigma^2(R, z) = \int d^3y \int d^3y' W(\vec{y}, R) W(\vec{y}', R) \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} e^{-i\vec{k}\cdot(\vec{x}+\vec{y})} e^{i\vec{k}'\cdot(\vec{x}+\vec{y}')} \langle \delta(\vec{k}, z) \delta^*(\vec{k}', z) \rangle. \quad (2.11)$$

Dado que a função de correlação de dois pontos no espaço de Fourier é definida como [63, 64, 65, 52]

$$\langle \delta(\vec{k}, z) \delta^*(\vec{k}', z) \rangle = (2\pi)^3 \delta_D(\vec{k} - \vec{k}') P(k, z), \quad (2.12)$$

onde $P(k)$ é o espectro de potências, $k = |\vec{k}|$ e $\delta_D(\vec{k} - \vec{k}')$ é a função delta de Dirac, que pode ser escrita como

$$\delta_D(\vec{k} - \vec{k}') = \int \frac{d^3\vec{x}}{(2\pi)^3} \exp[i\vec{x} \cdot (\vec{k} - \vec{k}')], \quad (2.13)$$

temos que a variância do contraste de densidade na escala R é,

$$\begin{aligned} \sigma^2(R, z) &= \int \frac{d^3k}{(2\pi)^3} P(k) \int d^3y W(\vec{y}, R) e^{-i\vec{k}\cdot\vec{y}} \int d^3y' W(\vec{y}', R) e^{i\vec{k}\cdot\vec{y}'} \\ &= \int_0^\infty \frac{dk}{k} \frac{k^3 P(k, z)}{2\pi^2} |W(k, R)|^2. \end{aligned} \quad (2.14)$$

Neste trabalho, nós usamos a função janela do tipo cartola que corresponde ao cálculo da média do contraste de densidade em um volume esférico de raio R . Com isso, a função cartola no espaço real é dada por

$$W_{TH}(r, R) = \frac{3}{4\pi R^3} \begin{cases} 1 & \leq R \\ 0 & > R. \end{cases}$$

No espaço de Fourier, ela fica

$$\begin{aligned} W_{TH}(k, R) &= \frac{3}{(kR)^3} (\sin kR - (kR) \cos kR) \\ &= \frac{3}{kR} j_1(k, R), \end{aligned} \quad (2.15)$$

onde $j_1(k, R)$ é a função esférica de Bessel de primeiro tipo.

Vale lembrar ainda que, com essa função janela, a relação entre a massa e o tamanho da região é dada por $M = \frac{4\pi}{3}\bar{\rho}_m(z)R^3$.

2.3 Espectro de potência linear

O espectro de potências $P(k, z)$ contém as informações da evolução linear das perturbações do campo de densidade de matéria até z , e pode ser escrito como

$$P(k, z) = Ak^{n_s}T(k)^2D(z)^2, \quad (2.16)$$

onde $T(k)$ é a função transferência, $D(z)$ é a função crescimento (veja a próxima seção), Ak^{n_s} é o espectro primordial de Harrison-Zel'dovich-Peebles e A é a sua normalização. Nós calculamos essa normalização usando o parâmetro⁴ σ_8 , que é o valor do desvio quadrático médio do contraste de densidade na escala $R = 8h^{-1}\text{Mpc}$, a saber

$$A = \frac{\sigma_8^2}{\int_0^\infty \frac{dk}{2\pi^2} k^{(n+2)} T(k)^2 W^2(k, 8)}, \quad (2.17)$$

onde $W(k, R = 8h^{-1}\text{Mpc})$ é a função janela cartola⁵.

A função transferência codifica a evolução dos modos k dos componentes materiais do universo desde o regime dominado pela radiação até o desacoplamento. Para obter $T(k)$ é necessário resolver as equações de Boltzmann acopladas de cada componente, por exemplo, matéria escura, bárions, fótons, neutrinos e perturbações escalares da métrica [67]. Existem códigos numéricos disponíveis, como o CMBFAST [68, 69, 70] e o CAMB, que resolvem esse conjunto de equações. Há também fórmulas ajustadas que foram obtidas para alguns modelos e que são válidas em intervalos bem definidos dos parâmetros [63, 71, 72]. Apesar do CMBFAST e CAMB fornecerem resultados mais precisos, neste trabalho nós usamos a fórmula ajustada de Eisenstein e Hu (EH [72]) por uma questão de otimização na velocidade do ajuste dos parâmetros cosmológicos. Vale ressaltar que a diferença relativa entre usar a função EH e o CMBFAST no cálculo da abundância de aglomerados é $\simeq 0.6\%$ no máximo (veja a figura 32 do apêndice B.5).

⁴ σ_8 é um dos parâmetros que será ajustado ou mantido livre nas análises estatísticas apresentadas nos próximos capítulos.

⁵Mesmo que o cálculo da equação (2.14) seja feito com uma função janela diferente do cartola, a normalização dada por (2.17) é, por definição, calculada com a função janela cartola [66].

2.4 Função Crescimento

Como a normalização do espectro de potências (eq. 2.17) é feita para $z = 0$ (ou seja, para o valor de σ_8 em $z = 0$), nós precisamos reescalonar o espectro para outros valores de z , o que é feito através da função de crescimento (como na equação 2.16). Para tanto, vamos considerar a equação que determina a evolução do contraste de densidade de matéria após a recombinação para os modos dentro do horizonte (e $k \ll k_J$ no caso dos bárions)[67, 73]

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = 4\pi G\rho_m\delta, \quad (2.18)$$

onde a é o fator de escala e $\dot{} = d/dt$. Dado que o fator de escala a está relacionado com o desvio para o vermelho por $1 + z = a_0/a$ de modo que

$$\frac{d}{dt} = -H(z)(1+z)\frac{d}{dz}, \quad (2.19)$$

temos que

$$H^2(z)(1+z)^2 \left[\frac{d^2\delta}{dz^2} + \left(\frac{1}{H(z)} \frac{dH(z)}{dz} - \frac{1}{(1+z)} \right) \frac{d\delta}{dz} \right] = 4\pi G\rho_m\delta, \quad (2.20)$$

onde $H(z)$ é a função de Hubble, dada por⁶

$$H(z) = H_0 \left[\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_{DE}e^{3\int_0^z d\ln(1+z')^{[1+w(z')]} + \Omega_k(1+z)^2 \right]^{1/2}, \quad (2.21)$$

e $H_0 = H(z=0)$. Os símbolos Ω_r , Ω_m , Ω_{DE} e $\Omega_k = 1 - \Omega_r - \Omega_m - \Omega_{DE}$ representam os parâmetros de densidade ($\Omega_x = \rho_x(z=0)/\rho_{crit}$, com $\rho_{crit} = 3H_0^2/8\pi G$) da radiação, matéria, energia escura e curvatura, respectivamente, e $w(z) = p/\rho$ é a equação de estado da energia escura.

Definindo a função crescimento $D(z) = \delta(z)/\delta(0)$ como a razão do contraste de densidade em z relativo ao contraste em $z = 0$ [74, 75], o que implica na condição de normalização $D(0) = 1$, e reescrevendo ρ_m na equação (2.20) como

$$\rho_m = \Omega_m(1+z)^3 \frac{3H_0^2}{8\pi G}, \quad (2.22)$$

temos que a função crescimento satisfaz a equação [76, 77]

$$\frac{d^2D}{dz^2} - \left(\frac{1}{1+z} - \frac{1}{E(z)} \frac{dE}{dz} \right) \frac{dD}{dz} - \frac{3\Omega_m}{2E^2(z)}(1+z)D = 0, \quad (2.23)$$

⁶Supondo um modelo no qual o universo é composto por matéria, radiação e energia escura, sem interação entre elas, e no qual essas componentes são homogêneas em grande escala.

onde $E(z)$ é a função de Hubble normalizada $E(z) = \frac{H(z)}{H_0}$.

Na figura 1, mostramos a função crescimento no intervalo $0 \leq z \leq 9$ para quatro cosmologias diferentes. A curva verde é obtida considerando-se um universo plano ($\Omega_k = 0$) e $\Omega_m = 1$, enquanto as curvas vermelha e azul correspondem ao modelo Λ CDM plano ($w = -1$) com $\Omega_\Lambda = 0,7$ e $\Omega_\Lambda = 0,742$ (dados do WMAP5 [78]), respectivamente. A curva amarela é obtida para um universo hiperbólico ($\Omega_k > 1$), $\Omega_\Lambda = 0$ e $\Omega_m = 0,3$. Podemos notar que as estruturas se formam primeiro no universo hiperbólico do que nos modelos de universo plano com e sem constante cosmológica.

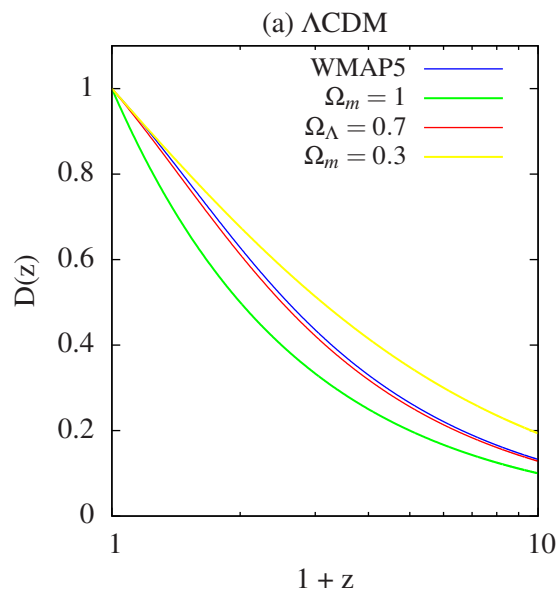


Figura 1: Função crescimento $D(z)$ em 4 diferentes cosmologias. Curva verde: universo plano, $\Omega_m = 1$. Curvas vermelha e azul: universo plano, $\Omega_\Lambda = 0,7$ e $\Omega_\Lambda = 0,742$, respectivamente. Curva amarela: universo hiperbólico, $\Omega_m = 0,3$.

2.5 Função de massa a partir de simulações de N-corpos

Para obter a abundância de halos em simulações de N-corpos é necessário utilizar métodos que identifiquem essas aglomerações e também formas de atribuir uma massa a esses halos. Os métodos mais utilizados para isso são o *friends-of-friends* [79, 80] e *spherical overdensity* [81]:

- *Friends-of-friends* (FoF)

Esse algoritmo consiste em identificar grupos de partículas cujos membros têm pelo menos um outro membro do grupo a uma distância menor ou igual a uma separação limite. Essa distância, chamada de comprimento de correlação, depende somente de um parâmetro b e da densidade média de partículas n , sendo igual a $bn^{1/3}$. Uma característica deste método é que ele não impõe um formato para os halos [8, 4, 10].

- *Spherical Overdensity* (SO)

Esse algoritmo identifica os halos crescendo uma esfera em torno de um centro até essa região atingir uma densidade limiar $\Delta\rho_{ref}$, definindo assim um raio R_Δ (ver equação 2.24). Existem várias maneiras de se escolher esse centro. Uma delas é estimar a densidade local em torno de cada partícula, os centros serão dados pelas regiões mais densas. Na referência [8, 10], após realizar o procedimento descrito, calcula-se o centro de massa da esfera e, a partir desse centro de massa, é crescida a esfera que determinará a massa do halo.

Um procedimento comum é combinar as duas técnicas, aplicando o algoritmo FoF para identificar grupos de partículas e definir o centro a partir do qual será aplicado o método SO para definir o halo e obter a sua massa [82]. A massa do halo é definida como

$$M_\Delta = \frac{4\pi}{3}\Delta\rho_{ref}R_\Delta^3, \quad (2.24)$$

onde Δ é um número que indica quantas vezes mais densa é a região onde se encontra o halo em relação à densidade de referência ρ_{ref} e R_Δ é o raio dessa região esférica. A densidade de referência pode ser a densidade crítica do universo ρ_{crit} , a densidade média do universo ρ_u ou a densidade média de matéria $\bar{\rho}_m$.

Uma vez obtida a distribuição (em número por volume comóvel) de aglomerados em função de sua massa, é possível obter a função de multiplicidade utilizando a equação (2.6)

$$f = -\frac{M}{\bar{\rho}_m(z)}\sigma_R(z) \left(\frac{d\sigma_R(z)}{dM} \right)^{-1} \frac{dn(M, z)}{dM} \quad (2.25)$$

Surpreendentemente, simulações feitas para diversos modelos cosmológicos e com distintos espectros iniciais (ver por exemplo as referências [8, 9]) levam a funções de multiplicidade muito semelhantes (embora distintas da de PS, equação 2.7). Ou seja, a propriedade de universalidade de PS é reproduzida nas simulações.⁷ Simulações mais recentes, explorando um intervalo maior de z mostraram que há uma dependência explícita

⁷Naturalmente, a função multiplicidade obtida nas simulações depende da escolha de Δ e ρ_{ref} assim como do algoritmo de identificação dos halos [83, 84, 85].

na função de multiplicidade em termos do desvio para o vermelho, mas ainda assim essa função depende muito pouco do modelo cosmológico.

O estado da arte em termos de funções de massa obtidas a partir de simulações é dado pela função multiplicidade de Tinker et al. [10], a saber

$$f(\sigma_R(z), z) = A \left[\left(\frac{\sigma_R(z)}{b} \right)^{-a} + 1 \right] \exp(-c\sigma_R^{-2}(z)), \quad (2.26)$$

onde

$$A(z) = A_0(1+z)^{-0,14}, \quad (2.27)$$

$$a(z) = a_0(1+z)^{-0,06}, \quad (2.28)$$

$$b(z) = b_0(1+z)^{-\alpha}, \quad (2.29)$$

e o subscrito 0 significa as quantidades obtidas em $z = 0$. Os valores de A_0, a_0, b_0, α e c (que não dependem de z) são dados na referência [10]. A equação (2.26) foi obtida com o algoritmo SO e para diferentes valores da sobredensidade Δ (equação 2.24) definida em relação à densidade média de matéria do universo em z . Na maior parte deste trabalho utilizaremos a escolha comum $\Delta = 200$, exceto quando precisamos utilizar um valor variável para essa quantidade (capítulo 5). Como mencionado, essa função ainda é universal no que se refere ao modelo cosmológico, mas há uma dependência explícita no desvio para o vermelho como dado pelas equações (2.27) – (2.29).

2.6 Conectando halos e aglomerados de galáxias

As funções multiplicidade obtidas com simulações de N-corpos, que citamos, fornecem ajustes para a função de massa de *halos* de matéria escura, enquanto os dados observacionais podem fornecer a *abundância de aglomerados de galáxias*. Em princípio, cada halo está associado a um aglomerado, de modo que ao calcularmos a função de massa (equação 2.6) podemos obter a abundância de aglomerados de galáxias. Para completar a associação entre os halos de matéria escura das simulações e os aglomerados de galáxias observados ainda é necessário associar uma massa e um desvio para o vermelho a esses últimos.

2.6.1 Desvio para o vermelho fotométrico

Como mencionado na introdução, os levantamentos fotométricos atuais, como o SDSS, e futuros, como o DES, cobrem várias bandas do espectro. Um dos objetivos é o emprego de técnicas de desvio para o vermelho fotométrico [86, 87] que utilizam o fluxo nas várias bandas para estimar o desvio para o vermelho das galáxias. Essas técnicas podem ser baseadas em padrões espectrais de galáxias ou no treinamento de algoritmos de aprendizagem utilizando amostras possuindo desvios para o vermelho espectroscópicos.

Seja qual for o método, a técnica de desvio para o vermelho fotométrico é probabilística, ou seja, a cada galáxia é atribuído um z^p que está associado a um z real por uma distribuição de probabilidade. Alguns métodos fornecem a distribuição $P(z^p|z)$ para cada galáxia, outros apenas um valor de z^p associado à galáxia. Dado um sub-conjunto de objetos com z medido espectroscopicamente, é possível determinar a distribuição de probabilidade de z^p , que normalmente é modelada por uma gaussiana (veja, por exemplo a seção 6.1). Por exemplo, para a amostra total de galáxias no DES, espera-se que a largura dessa gaussiana seja $\sigma_z = \sigma_0(1 + z)$, com $\sigma_0 \simeq 0,03$.

Os aglomerados de galáxias são compostos majoritariamente por galáxias vermelhas, que possuem características espectrais bem definidas, o que faz com que a incerteza no z^p seja menor que para a população global de galáxias. Além disso, os próprios métodos de identificação de aglomerados levam a uma determinação mais precisa do desvio para o vermelho fotométrico. Assim, para os aglomerados detectados no DES espera-se ter $\sigma_0 \simeq 0,01$.

Como veremos no capítulo 6, é fundamental modelar os efeitos da distribuição do desvio para o vermelho fotométrico na abundância de aglomerados para recuperar corretamente os parâmetros cosmológicos.

2.6.2 Relação Massa—Observável

Um aspecto fundamental da análise da abundância de aglomerados é a necessidade de ter estimativas para a massa desses objetos. Estas podem ser obtidas a partir de dados da fotometria no próprio projeto, como no caso das estimativas com o efeito fraco de lente gravitacional [24, 25], ou por observações em outros comprimentos de onda, como em raios-x [21, 22] ou através do efeito Sunyaev-Zel'dovich [23].

Para esses observáveis, é possível ter medidas individuais ou agrupadas (isto é, para

um conjunto de objetos) da massa. Por exemplo, para os aglomerados mais massivos, a distorção das galáxias de fundo produzida pelo lenteamento é suficiente para poder ser detectada em um único aglomerado [88, 89, 90]. Já para aglomerados menos massivos, é necessário combinar as distorções das imagens de fundo para um grande número de aglomerados para ter uma boa razão sinal-ruído [91]. Para fazer essa combinação, agrupam-se aglomerados com propriedades similares, por exemplo todos os aglomerados com uma mesma riqueza ótica, obtendo-se uma *massa média* para esses objetos. Dado o elevado número de aglomerados nos grandes levantamentos, essas medidas combinadas podem ser bastante precisas. Fazendo uma análise desse tipo em intervalos de riqueza, é possível obter relações de escala do tipo $M_{obs} = AN_{gal}^\beta$. Esse tipo de estudo pode ser realizado com os vários observáveis mencionados acima.

No entanto, a relação entre riqueza e massa não é unívoca. Dado um aglomerado com uma massa definida, há uma probabilidade de que ele contenha um certo número de galáxias. Portanto, as relações entre massa e observável (como a riqueza) são intrinsecamente probabilísticas e possuem uma dispersão. Em geral supõe-se que a relação massa-observável segue uma distribuição log-normal (ou seja, normal em $\log M$), como discutido na seção 6.2. Mesmo para os casos em que é possível ter uma estimativa para a massa de modo individual, ainda é preciso levar em conta as incertezas na medida da massa.

A relação média, como a lei de potência discutida acima, pode ser obtida combinando-se uma grande amostra de aglomerados. Já para estudar a largura da distribuição de probabilidade é necessário utilizar outros métodos (veja, por exemplo a discussão na seção 7). Aqui utilizaremos algumas estimativas para $\sigma_{\ln M}$ para algumas medidas típicas da massa, como será mencionado na seção 6.2, na qual modelamos os efeitos da distribuição massa-observável na abundância de aglomerados.

2.7 Contagem de aglomerados de galáxias

A função de massa como definida na equação (2.6) não é um observável direto, pois depende do volume que depende do modelo cosmológico. Uma quantidade que pode ser medida de forma independente do modelo cosmológico é a abundância de aglomerados em uma dada área, em um dado intervalo de z e para um dado intervalo de massas. Em particular, é muito usual considerar a contagem de aglomerados com massa acima de um limiar M_{\min} (que pode estar associado, por exemplo, ao limite de detecção ou de

completeza de um catálogo).

Como a função de massa fornece a densidade média de aglomerados por unidade de massa, temos, portanto, que integrar essa função na massa a partir de um limiar M_{\min} e em uma região do céu com volume comóvel V para obtermos o número médio de objetos com massa maior que M_{\min} no volume V , i.e.,

$$\frac{dN(M \geq M_{\min}, V)}{dz} = \frac{dV}{dz} \int_{M_{\min}}^{\infty} dM \frac{dn(M, z)}{dM}. \quad (2.30)$$

Em coordenadas esféricas, temos $dV = r_A(z)^2 d\Omega dr_A$, onde $d\Omega = \sin\theta d\theta d\phi$ e $r_A(z)$ é a distância de diâmetro-angular [92, 64]. Essa última pode ser escrita em termos da distância comóvel $r(z)$ como

$$r_A(z) = \begin{cases} \frac{\sin(\sqrt{-\Omega_k} R_H r(z))}{\sqrt{|\Omega_k|} R_H} & \Omega_k < 0 \\ r(z) & \Omega_k = 0 \\ \frac{\sinh(\sqrt{\Omega_k} R_H r(z))}{\sqrt{|\Omega_k|} R_H} & \Omega_k > 0, \end{cases}$$

onde $R_H = c/H_0$, c é a velocidade da luz e

$$r(z) = \int_0^z dz' \frac{c}{H(z)}. \quad (2.31)$$

Restringindo a análise ao caso do universo plano ($\Omega_k = 0$), temos que a conversão do elemento de volume para o espaço do desvio para o vermelho é dado por

$$\frac{dV}{dz} = \frac{c}{H(z)} r^2(z) d\Omega. \quad (2.32)$$

Como a função de massa e a distância comóvel independem dos ângulos θ e ϕ , temos que a integral no ângulo sólido corresponderá à área angular do levantamento, A_{survey} . Portanto, temos que a abundância de aglomerados em função do desvio para o vermelho é

$$\frac{dN}{dz} = A_{\text{survey}} \frac{dV}{dz d\Omega} \int_{M_{\min}}^{\infty} dM \frac{dn(M, z)}{dM}, \quad (2.33)$$

e, conseqüentemente, a sua abundância no intervalo $z_{\min} \leq z \leq z_{\max}$ é

$$N(M \geq M_{\min}, V) = A_{\text{survey}} \int_{z_{\min}}^{z_{\max}} dz' \left[\frac{dV}{dz' d\Omega} \int_{M_{\min}}^{\infty} dM \frac{dn(M, z')}{dM} \right]. \quad (2.34)$$

Neste ponto fica explícito que a distribuição de aglomerados depende do modelo cosmológico através do elemento de volume e também da função crescimento (2.23), através de $\sigma_M(z)$.

2.8 Parametrização do modelo de energia escura

Um aspecto central desta teste é estudar limites que podem ser impostos na energia escura a partir da abundância de aglomerados de galáxias. É comum definir a *equação de estado da energia escura* $w(z)$ por

$$p = w(z)\rho. \quad (2.35)$$

Ao invés de considerar modelos específicos de energia escura, nos restringiremos a uma parametrização que pode ser considerada uma expansão de w em termos do fator de escala a em torno de $a = 1$ (ou seja $z = 0$) [93, 94], temos que

$$w(z) = w_0 + w_1(1 - a) = w_0 + w_1 \frac{z}{1 + z}. \quad (2.36)$$

Com isso podemos levar em conta uma possível variação da equação de estado com z . Qualquer desvio significativo de $w_0 = -1$ e $w_1 = 0$ descartaria a constante cosmológica como descrição da energia escura (naturalmente, supondo válidas todas as outras premissas, como homogeneidade em grandes escalas, etc.).

Além de ser uma expressão bastante geral, a equação (2.36) tem sido utilizada como um padrão para comparar a habilidade de diferentes projetos e observáveis em limitar a energia escura. Por exemplo, o *Dark Energy Task Force* [95] utiliza a *figura de mérito* no plano (w_0, w_1) para comparar e definir distintas categorias de projetos para estudar a energia escura.

Outra equação de estado que tem sido muito estudada na literatura é $w = const.$, que pode ser considerada um caso particular da equação 2.36 com $w_1 = 0$.

Nesta tese vamos considerar apenas uma equação de estado do tipo 2.36, em alguns casos nos restringindo ao caso $w = w_0$, modelo ao qual nos referiremos como XCDM.

Quando for dito parâmetros da energia escura, estaremos nos referindo aos parâmetros w_0 e w_1 .

No caso da equação de estado 2.36, o parâmetro de Hubble pode ser escrito como

$$H(z) = H_0 \left[\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_{DE}(1+z)^{3(1+w_0+w_1)} e^{-3w_1 \frac{z}{1+z}} + \Omega_k(1+z)^2 \right]^{1/2}. \quad (2.37)$$

Vale ressaltar que, em todo o trabalho, nós faremos a suposição de que o universo é plano, ou seja, $\Omega_k = 0$, e também escreveremos $H_0 = h \times 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

Na figura 2.8, mostramos um exemplo de como dN/dz varia com o valor do parâmetro

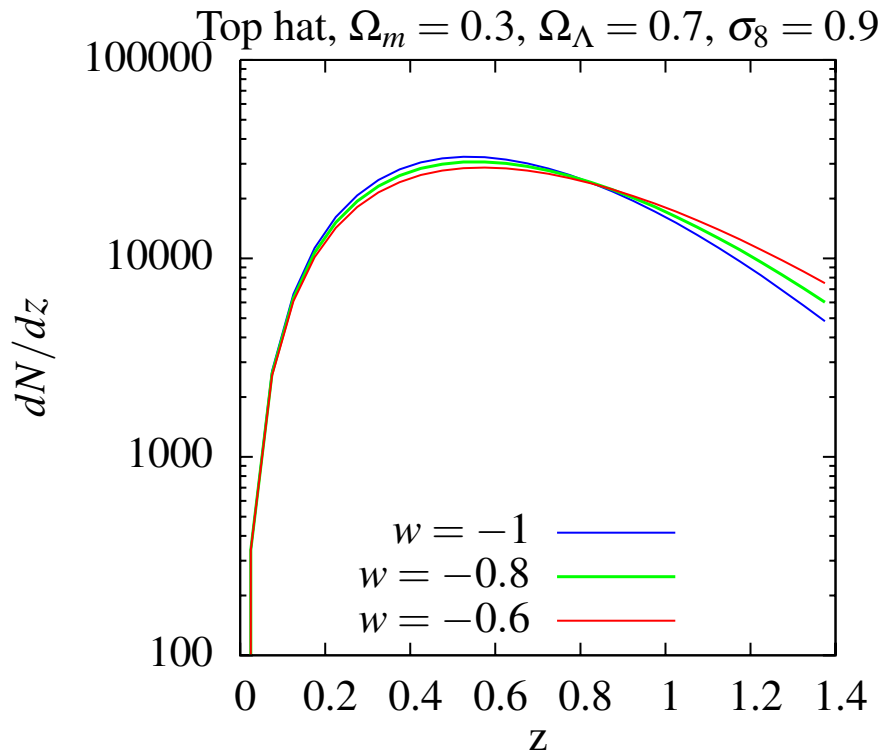


Figura 2: Distribuição da abundância de halos em uma área de 5000deg^2 em função do desvio para o vermelho no modelo Λ CDM com $w = -0,6$ (curva vermelha), $w = -0,8$ (curva verde) e $w = -1$ (curva azul).

da equação de estado da energia escura w_0 do modelo Λ CDM em três diferentes casos: $w_0 = -0,6$, $w_0 = -0,8$ e $w_0 = -1$. Podemos notar que a distribuição de aglomerados é praticamente insensível à equação de estado w_0 até $z \simeq 0,2$. Essa figura também mostra que para distinguirmos os modelos de energia escura a partir da abundância de aglomerados, há a necessidade de medir esses objetos para altos valores de z .

Todas as funções apresentadas neste capítulo, que são necessárias para obter a abundância de aglomerados de galáxias (equações 2.33 e 2.34), foram implementadas na linguagem C e extensivamente testadas, validadas e otimizadas. No apêndice B, apresentamos, entre outras coisas, testes do controle de precisão dos códigos e a sua validação reproduzindo alguns resultados encontrados na literatura.

3 Verossimilhanças para Abundância de Aglomerados de Galáxias

No capítulo 1, mencionamos a grande disponibilidade de dados de aglomerados de galáxias e a sensibilidade de sua abundância em relação aos parâmetros cosmológicos e da energia escura. Portanto, para que possamos restringir esses parâmetros, precisamos implementar uma análise estatística adequada para esses dados.

Neste capítulo, apresentamos verossimilhanças¹ construídas a partir da abundância de aglomerados. Na seção 3.1, revisamos duas verossimilhanças utilizadas na literatura, as quais são obtidas considerando o número de aglomerados em *bins* do desvio para o vermelho. Na seção 3.3, nós propomos uma verossimilhança onde levamos em conta a distribuição dos aglomerados sem binagem em z . Finalizamos o capítulo construindo uma verossimilhança a partir da distribuição de objetos em função do desvio para o vermelho e da massa na seção 3.5.

3.1 Verossimilhanças binadas no desvio para o vermelho

A abundância média de aglomerados no intervalo $z_{\min} \leq z' \leq z_{\max}$ é dada pela equação (2.34). Dividindo esse intervalo em n *bins*, temos que o número médio de aglomerados no k -ésimo *bin* do desvio para o vermelho de largura Δz é

$$N_k = A_{\text{survey}} \int_{z_k - \frac{\Delta z}{2}}^{z_k + \frac{\Delta z}{2}} dz' \left[\frac{dV}{dz' d\Omega} \int_{M_{\min}}^{\infty} dM \frac{dn(M, z')}{dM} \right]. \quad (3.1)$$

Supondo que a contagem de aglomerados em um bin de desvio para o vermelho segue

¹No apêndice A, nós apresentamos uma breve revisão sobre algumas ferramentas estatísticas utilizadas neste trabalho.

a distribuição de Poisson², construímos a verossimilhança

$$L(\{N_k^{\text{teo}}(\vec{\theta})\}, \{N_k\}) = \prod_{k=1}^n \frac{(N_k^{\text{teo}}(\vec{\theta}))^{N_k} e^{-N_k^{\text{teo}}(\vec{\theta})}}{N_k!}, \quad (3.2)$$

onde n é o número de *bins* no intervalo $[z_{\min}, z_{\max}]$, $\{N_k^{\text{teo}}(\vec{\theta})\}$ é o número de aglomerados dado pelo modelo (equação 3.1), $\vec{\theta}$ é o conjunto de parâmetros do modelo³ e $\{N_k\}$ é o número de aglomerados observados em cada *bin* com massa acima do limiar M_{\min} . Dado um conjunto $\{z_i\}$ dos desvios para o vermelho dos aglomerados com $M > M_{\min}$, contamos o número N_k como a quantidade de valores de z que caem no intervalo k . Como discutido no apêndice A, os estimadores $\hat{\theta}_j$ são obtidos a partir da maximização da verossimilhança. Na prática, minimizamos a função χ_P^2 , onde

$$\chi_P^2 := -2 \ln L(\{N_k^{\text{teo}}(\theta)\}, \{N_k\}) = -2 \sum_{k=1}^n [N_k \ln(N_k^{\text{teo}}(\theta)) - N_k^{\text{teo}}(\theta) - \ln(N_k!)]. \quad (3.3)$$

Nos referiremos a esse método como *Método de Poisson*. A aplicação desse método pode ser encontrada, por exemplo, nas referências [96, 97, 98, 99].

Vamos considerar também o *Método dos Mínimos Quadrados* cujo uso, por sua simplicidade e generalidade, é muito comum na literatura (veja a referência [100], por exemplo). Neste caso, os estimadores são obtidos minimizando-se a equação abaixo,

$$\chi_{\text{MQ}}^2 = \sum_{k=1}^n \frac{(N_k - N_k^{\text{teo}}(\theta))^2}{N_k^{\text{teo}}(\theta)}. \quad (3.4)$$

Como mostramos na seção A.1, essa equação também pode ser obtida com o método da máxima verossimilhança se considerarmos que N_k segue uma distribuição Gaussiana⁴ em cada *bin*. Chamaremos esse método de *Método* χ^2 .

²A distribuição de Poisson é dada por

$$P(n; \lambda) = \frac{\lambda^n}{n!} e^{-\lambda},$$

onde n é a medida. Essa distribuição possui média e variância $\langle n \rangle = \lambda$ e $V(n) = \lambda$, respectivamente.

³Tanto os cosmológicos quanto os da equação de estado da energia escura: Ω_m , Ω_b , Ω_{DE} , $w(z)$, σ_8 , n_s e h , onde Ω_b é o parâmetro de densidade de bárions e h é a constante de Hubble adimensional ($H_0 = 100 \times h \text{ km s}^{-1} \text{ Mpc}^{-1}$).

⁴Distribuição Gaussiana com média μ e variância σ^2 : $G(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

3.2 Estudo do viés na equação de estado da energia escura para os métodos binados

A distribuição Gaussiana é uma boa aproximação para a distribuição de Poisson quando o número de eventos, no nosso caso a abundância em *bins* de z , é relativamente alto (em muitos casos $N_k \gtrsim 10$ é o suficiente). Portanto, esperamos que, quando houver muitos objetos em cada *bin*, o uso do método dos mínimos quadrados (equação 3.4) seja equivalente ao de Poisson (equação 3.2 ou 3.3).

No entanto, se o número de objetos é baixo, a utilização do método χ^2_{MQ} pode induzir um viés na obtenção dos parâmetros θ_j . Em particular, se M_{min} for muito elevado, o número de aglomerados em cada *bin* pode ser muito baixo. Para quantificar esse viés, é necessário criar uma amostra fictícia de aglomerados z_i (seguindo a distribuição 2.33) com valores fiduciais de θ_j^0 , contar o número de objetos em cada *bin* N_k e obter os parâmetros cosmológicos com os estimadores obtidos a partir da minimização das expressões (3.3) e (3.4). Como nós geramos a amostra a partir de um modelo fiducial, conhecemos a priori os valores verdadeiros θ_j^0 , os quais iremos comparar com os valores esperados dos estimadores ($\langle \hat{\theta}_j \rangle$) a fim de verificar se existe viés.

Note que só podemos obter o viés a partir de uma simulação⁵, como descrito acima. Se simplesmente calculássemos o número de objetos esperado em cada *bin* (N_k) a partir da equação (3.1), usando essa mesma expressão como modelo teórico, naturalmente recuperaríamos os valores fiduciais de θ_j^0 usando tanto a equação (3.3) quanto a (3.4). Por isso é necessário fazer uma **realização** $\{z_i\}$ a partir da distribuição de probabilidade dada pelo modelo.

3.2.1 Geração de uma realização da distribuição de aglomerados

Aqui vamos supor que a distribuição de aglomerados em cada *bin* de z efetivamente segue uma distribuição de Poisson⁶. Para obter uma realização $\{z_i\}$, adotamos o procedimento descrito a seguir.

Primeiramente, calculamos a abundância de aglomerados $N_{\text{total}}^{\text{teo}}(\theta)$ obtida a partir da equação (2.34) para um modelo fiducial⁷. Usando o valor $N_{\text{total}}^{\text{teo}}(\theta)$ como média de uma

⁵Neste caso, nos referimos ao termo simulação como um procedimento do tipo *Fast Monte Carlo* e não a uma simulação de N-corpos.

⁶Com isso estamos ignorando, por exemplo, efeitos da estrutura em grande escala, o que é uma boa aproximação para levantamentos de grande área.

⁷O modelo fiducial inclui os parâmetros da amostra ($A_{\text{survey}}, z_{\text{min}}, z_{\text{max}}, M_{\text{min}}$) e os parâmetros

distribuição de Poisson, nós geramos aleatoriamente o número total de objetos $N_{\text{total}}^{\text{obs}}$.

O segundo passo consiste em criar uma amostra com $N_{\text{total}}^{\text{obs}}$ valores de desvios para o vermelho z_i a partir da distribuição de probabilidade de encontrar um aglomerado com massa maior que M_{min} e com desvio para o vermelho entre z e $z + dz$, ou seja,

$$\mathcal{P}(z)dz = \frac{1}{N_{\text{total}}^{\text{teo}}} \left(\frac{dN}{dz} \right) dz, \quad (3.5)$$

onde $\frac{1}{N_{\text{total}}^{\text{teo}}}$ é a normalização. Para gerar o conjunto $\{z_i\}$ seguindo essa distribuição, usamos sua cumulativa

$$f(z) = \int_0^z \mathcal{P}(z')dz'. \quad (3.6)$$

Como $f(z)$ é uma função monotonicamente crescente de imagem $[0, 1]$, temos uma relação um para um entre z e $f(z)$. Dado também que a função $f(z)$ é uma variável aleatória de distribuição uniforme⁸ entre $0 \leq f(z) \leq 1$ ⁹, nós geramos $N_{\text{total}}^{\text{obs}}$ números aleatórios $\{u_i\}$ de uma distribuição uniforme e invertemos a equação (3.6) obtendo $z(f)$ e, assim, o conjunto $\{z_i\} = \{z(u_i)\}$. Finalmente, temos um conjunto “simulado” $\{z_i\}$, ao qual estamos nos referindo como uma realização da distribuição de aglomerados. A partir dessa amostra criamos um histograma com n bins no desvio para o vermelho e obtemos o conjunto de “dados” $\{N_k\}$. Podemos agora utilizar as expressões (3.3) e (3.4) para obter os estimadores $\hat{\theta}_j$.

3.2.2 Método de Monte Carlo para a abundância de aglomerados

Através do procedimento descrito acima, podemos obter os valores de $\hat{\theta}_j$ que melhor ajustam uma dada realização. No entanto, esses não corresponderão (geralmente) aos valores fiduciais θ_j^0 utilizados na realização, justamente pela sua natureza probabilística e o número finito de elementos na amostra simulada. Como os estimadores dos parâmetros cosmológicos não possuem uma fórmula analítica, o cálculo direto do valor esperado <
cosmológicos e da energia escura.

⁸A distribuição uniforme é definida por

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & x < a \text{ ou } x > b. \end{cases}$$

⁹A probabilidade da função $f(z)$ ter um valor u é dada por

$$P_f(u) = \int_0^{z_f} \delta(u - f(z))\mathcal{P}(z)dz = \int_0^{z_f} \delta(u - f(z))\frac{df(z)}{dz}dz = 1.$$

$\hat{\theta}_j >$ só pode ser feito numericamente (o que o torna inviável¹⁰). Portanto, usaremos o método Monte Carlo para calcular uma estimativa dos valores esperados de $\hat{\theta}_j$.

Esse método consiste em gerar um grande número de realizações e, para cada realização, calcular o melhor ajuste $\hat{\theta}_j$ e, assim, obter uma aproximação da distribuição de probabilidade de $\hat{\theta}_j$.

A partir das equações [101],

$$\bar{\theta}_j = \sum_{l=1}^m \frac{\hat{\theta}_{jl}}{m} \quad (3.7)$$

e

$$\sigma^2(\hat{\theta}_j) = \frac{1}{(m-1)} \sum_{l=1}^m (\hat{\theta}_{jl} - \bar{\theta}_j)^2, \quad (3.8)$$

obtemos uma estimativa do valor esperado e da variância de $\hat{\theta}_j$, respectivamente, onde m é o número de realizações e $\hat{\theta}_{jl}$ é o melhor ajuste para a l -ésima realização.

Portanto, é o valor $\bar{\theta}_j$ que deve ser comparado com o fiducial para avaliar a capacidade de um dado estimador em recuperar o parâmetro de interesse, ou seja, verificar se ele tem um viés. Como a variância da média $\bar{\theta}_j$ é dada por

$$\sigma^2(\bar{\theta}_j) = \frac{\sigma^2(\hat{\theta}_j)}{m}, \quad (3.9)$$

é necessário fazer um número grande o suficiente de realizações para determinar se o valor esperado de $\hat{\theta}_j$ tem ou não viés.

3.2.3 Estimadores para w_0 nas estatísticas binadas

Se por um lado o método de *Monte Carlo* é bastante poderoso, por outro lado ele é computacionalmente dispendioso, já que é necessário fazer inúmeras realizações, maximizar as verossimilhanças (ou minimizar o χ^2) e analisar as distribuições dos valores obtidos. Para ser aplicado de forma direta sem a necessidade de muitos recursos computacionais, é necessário limitar-se a poucos parâmetros. Neste trabalho faremos a análise com apenas um parâmetro livre, a saber a equação de estado da energia escura $\theta = w_0$ (supondo $p = w_0\rho$). A escolha do modelo fiducial foi baseada nas configurações do levantamento *Dark Energy Survey* e nos parâmetros da simulação Carmen (veja a seção 1.2). Assim, usamos $A_s = 5000 \text{ deg}^2$, $z \in [0; 1, 4]$ e os parâmetros $h = 0,7$, $\Omega_b = 0,04$, $\Omega_{\text{DE}} = 0,75$, $n_s = 1$, $\sigma_8 = 0,8$ e $w_0 = -1$.

¹⁰O cálculo direto é dado por $\langle \hat{\theta}_j \rangle = \int d^n N_k \hat{\theta}_j L(\{N_k^{\text{teo}}(\{\theta_j\})\}, \{N_k\})$, que envolve uma integral multidimensional e cada cálculo do integrando uma minimização.

Utilizamos o método Monte Carlo gerando 1000 realizações de $\{N_k\}$ e então, para cada realização, calculamos o melhor ajuste de \hat{w}_0^P e $\hat{w}_0^{\chi^2}$, minimizando as equações (3.3) e (3.4), respectivamente, em relação a w_0 mantendo os outros parâmetros cosmológicos fixos. Com isso, usando as equações (3.7) e (3.8), obtemos as médias $\overline{\hat{w}_0^P}$ e $\overline{\hat{w}_0^{\chi^2}}$ e as variâncias $\sigma^2(\hat{w}_0^P)$ e $\sigma^2(\hat{w}_0^{\chi^2})$, respectivamente.

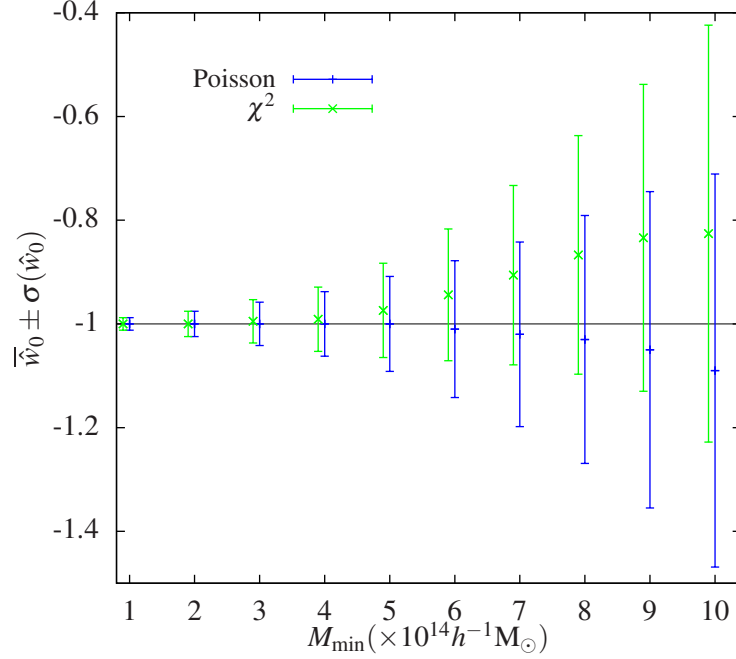


Figura 3: A média $\overline{\hat{w}_0}$ e o desvio padrão $\sigma(\hat{w}_0)$ do estimador \hat{w}_0 , obtidos com método Monte Carlo com 1000 realizações da abundância de aglomerados, para diferentes valores de $M_{\min} = 1 - 10 \times 10^{14} h^{-1} M_{\odot}$, com $\Delta z = 0, 1$ (14 bins). A linha azul corresponde ao método de Poisson e a linha verde ao método dos Mínimos Quadrados (χ^2). Um pequeno desvio foi introduzido no eixo x para facilitar a visualização.

Os resultados desse procedimento são apresentados na figura 3, onde usamos $\Delta z = 0, 1$ (14 bins)¹¹ e 10 diferentes valores da massa limite mínima, $M_{\min} = 1, 2, 3, 4, 5, 6, 7, 8, 9$ e $10 \times 10^{14} h^{-1} M_{\odot}$. Notamos que o valor de $\sigma(\hat{w}_0)$ é muito sensível à massa, o que é justificado pelo fato de que a abundância de aglomerados cai rapidamente com o aumento de M_{\min} . Verificamos que o método χ^2 introduz um viés na estimativa de \hat{w}_0 , que cresce com o aumento da massa mínima. Esse viés é de cerca de 1% para $M_{\min} \simeq 3 \times 10^{14} h^{-1} M_{\odot}$ chegando a 17% para $M_{\min} \simeq 10^{15} h^{-1} M_{\odot}$, como mostrado na tabela 1. Isso era esperado já que, como citamos anteriormente, a Gaussiana não é uma boa aproximação para distribuição de Poisson quando $N_k \lesssim 10$. Por exemplo, no caso em que $M_{\min} = 10^{15} h^{-1} M_{\odot}$, as realizações têm de 5 a 7 bins com $N_k \leq 5$. O método de Poisson também possui um viés, mas

¹¹Realizamos este estudo para bins do desvio para o vermelho com diferentes larguras e verificamos que o resultado é quase insensível a Δz .

este só se manifesta a partir de massas maiores (cerca de 1% para $M_{\min} \simeq 6 \times 10^{14} h^{-1} M_{\odot}$) sendo sempre significativamente menor que no método dos Mínimos Quadrados. As incertezas no viés (equação 3.9) para cada massa são mostradas na tabela 1 e na figura 4 tanto para o método χ^2 quanto para o de Poisson.

$M_{\min}[10^{14}h^{-1}M_{\odot}]$	$ \frac{w_0 - \bar{w}_0^{\chi^2}}{w_0} $	$\sigma(\bar{w}_0^{\chi^2})$	$ \frac{w_0 - \bar{w}_0^P}{w_0} $	$\sigma(\bar{w}_0^P)$
1	0	4×10^{-4}	0	4×10^{-4}
2	0	8×10^{-4}	0	8×10^{-4}
3	1%	10^{-3}	0	10^{-3}
4	1%	2×10^{-3}	0	2×10^{-3}
5	3%	3×10^{-3}	0	3×10^{-3}
6	6%	4×10^{-3}	1%	4×10^{-3}
7	9%	5×10^{-3}	2%	5×10^{-3}
8	13%	7×10^{-3}	3%	7×10^{-3}
9	17%	9×10^{-3}	5%	9×10^{-3}
10	17%	10^{-2}	9%	10^{-2}

Tabela 1: O viés relativo ao valor fiducial, $|(w_0 - \bar{w}_0)/w_0|$, obtido a partir do método χ^2 (segunda coluna) e do método de Poisson (quarta coluna) para diferentes valores de massa M_{\min} . E a incerteza no viés \bar{w}_0 para os métodos χ^2 (terceira coluna) e Poisson (quinta coluna).

É interessante notar que o viés do método χ^2 começa a ser significativo para massas $M \gtrsim 5 \times 10^{14} M_{sun}$, o que justamente é próximo do limiar para a determinação da massa para métodos utilizando raios-x [21, 22], o efeito Sunyaev–Zel’dovich [23] e medidas individuais a partir do efeito fraco de lentes gravitacionais [89, 88, 90]. Desse modo, especialmente catálogos construídos a partir desses métodos deveriam utilizar o método de Poisson (ou uma nova verossimilhança que será introduzida na próxima seção), o que nem sempre tem sido o caso na literatura (veja, por exemplo, a ref. [100]).

Vale ressaltar que o gasto computacional para calcular $\hat{w}_0^{\chi^2}$ é equivalente ao do cálculo de \hat{w}_0^P . Portanto, devido a esse fato e também ao que foi discutido ao longo dessa seção, deve-se escolher o método de Poisson ao utilizar a abundância de aglomerados binada no desvio para o vermelho para restringir os parâmetros cosmológicos.

3.3 Verossimilhança para a abundância sem binagem no desvio para o vermelho

Nas análises que requerem a utilização de *bins* (“análises binadas”) sempre há alguma ambiguidade para a escolha do número/tamanho desses *bins*. Embora haja critérios pro-

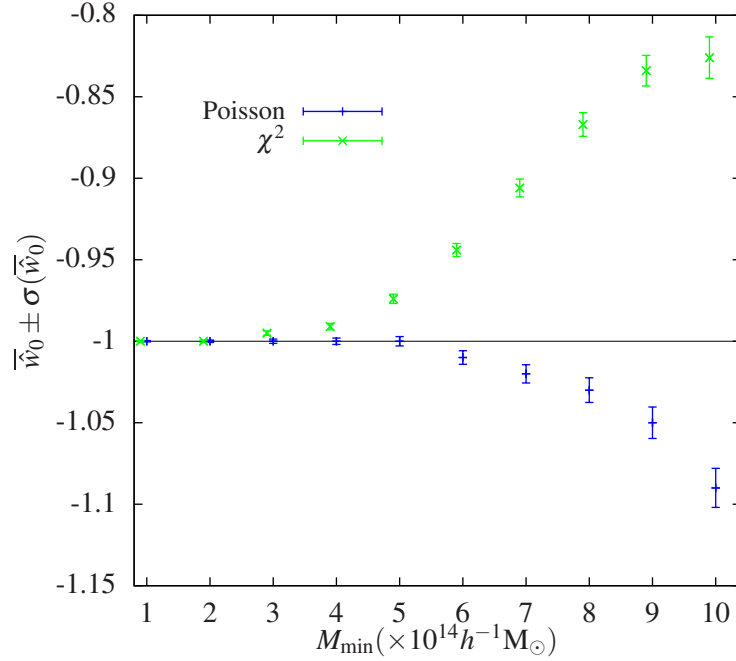


Figura 4: A média \bar{w}_0 e o **seu** desvio padrão $\sigma(\bar{w}_0)$, obtidos com método Monte Carlo com 1000 realizações da abundância de aglomerados, para diferentes valores de $M_{\min} = 1 - 10 \times 10^{14} h^{-1} M_{\odot}$, com $\Delta z = 0, 1$ (14 bins). A linha azul corresponde ao método de Poisson e a linha verde ao método dos Mínimos Quadrados (χ^2). Um pequeno desvio foi introduzido no eixo x para facilitar a visualização.

postos para definir o seu tamanho, não há uma forma única que assegure que não há perda de informação ao se agrupar os dados em *bins*. Em particular, em distribuições que exibem grandes variações, como é o caso da abundância de aglomerados para altos z , a utilização de métodos binados pode acarretar em vieses nos parâmetros. Por isso é conveniente termos um método não-binado para a contagem de aglomerados.

Há ainda uma razão mais fundamental para buscarmos uma abordagem não binada para a abundância de aglomerados. Uma questão fundamental na modelagem da abundância de aglomerados é a inclusão das incertezas no desvio para o vermelho (devido à utilização de desvios para o vermelho fotométricos) e na massa. Essas questões serão abordadas e estudadas no capítulo 6. Nas abordagens da abundância em *bins* é possível incluir essas incertezas, mas apenas de forma média, ou seja, as distribuições de probabilidade de z^p ($p(z^p|z)$) e M^{obs} ($P(M^{obs}|M)$) têm que ser as mesmas para todos os objetos com a mesma massa e desvio para o vermelho. A abordagem binada não permite levar em conta as incertezas individuais em M^{obs} e z^p .

Alguns métodos para obter a distribuição de desvios para o vermelho e para a calibração da massa realmente requerem que os aglomerados sejam agrupados em *bins* de M

e/ou z . Nesse caso, naturalmente, as distribuições são válidas, por construção, para toda a população contida nessas amostras. No entanto, há vários métodos que levam a medidas individuais de z e M . Por exemplo, para aglomerados massivos é possível ter medidas da massa individuais a partir da emissão em raios-X, do efeito Sunyaev-Zel'dovich e do efeito fraco de lente gravitacional. Para essas medidas, cada aglomerado tem uma incerteza na massa estimada de forma também individual, com uma distribuição de probabilidade associada. Por outro lado, vários algoritmos de desvio para o vermelho fotométrico fornecem uma distribuição de probabilidade de z para cada galáxia. Sabendo quais galáxias estão associadas a um determinado aglomerado, é possível combinar essas probabilidades para obter a distribuição de probabilidade de z desse aglomerado.

Um método não binado, que leve em conta cada aglomerado de galáxias, e não o número em *bins*, pode permitir naturalmente a inclusão das incertezas individuais em z e M , como veremos no capítulo 6.

Para construir uma verossimilhança não binada, devemos atribuir uma probabilidade para um dado aglomerado estar em um desvio para o vermelho z_i . Isso é dado naturalmente pela expressão (3.5). A probabilidade de um conjunto de aglomerados ter o conjunto de valores $\{z_i\}$ seria portanto o produto das probabilidades associadas a cada valor de z_i . Além da distribuição dos aglomerados em z , precisamos levar em conta o número total de aglomerados na amostra em comparação com a previsão teórica. Desse modo, construímos a verossimilhança a partir do produto das probabilidades (3.5), multiplicando pela probabilidade do número total de objetos:

$$L(\{\theta_j\}, \{z_i\}) = P(N_{\text{total}}^{\text{obs}}; N_{\text{total}}^{\text{teo}}(\{\theta_j\})) \prod_{i=1}^{N_{\text{total}}^{\text{obs}}} \mathcal{P}(z_i), \quad (3.10)$$

onde $P(N_{\text{total}}^{\text{obs}}; N_{\text{total}}^{\text{teo}}(\{\theta_j\}))$ é a distribuição de Poisson. O termo \mathcal{P} (equação 3.5) leva em conta a abundância relativa em função de z , ou seja, a distribuição dos aglomerados ao longo do intervalo do desvio para o vermelho. Já a distribuição de Poisson leva em consideração o número total de aglomerados. Nos referiremos a esse método como *Método ULC* (*unbinned likelihood for clusters*).

Agora iremos validar a verossimilhança (3.10) construindo estimadores a partir dela, utilizando o método Monte Carlo e comparando os resultados com os do método de Poisson.

Analogamente à análise que fizemos na seção 3.2.3 para os métodos de Poisson e χ^2 , nós aplicamos o método Monte Carlo usando as mesmas 1000 realizações de $\{z_i\}$ e

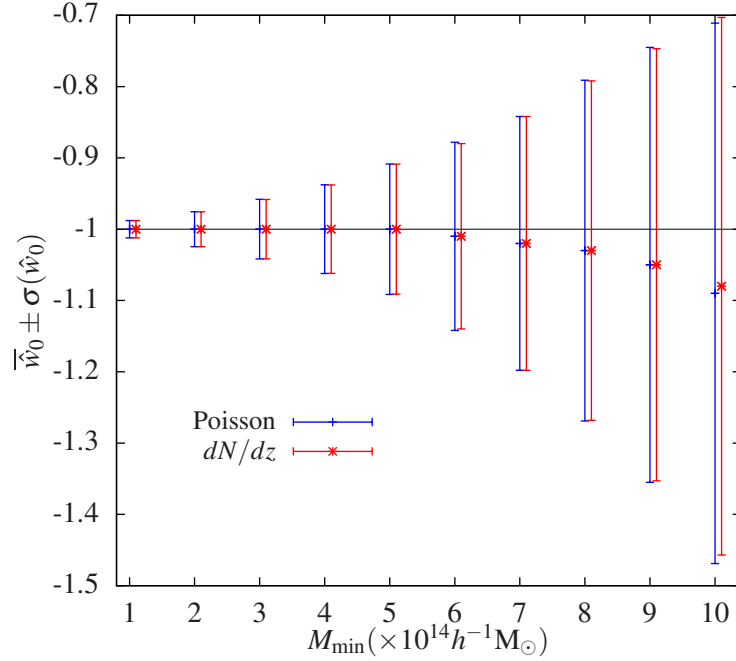


Figura 5: A média \bar{w}_0 e o desvio padrão $\sigma(\hat{w}_0)$, obtidos com método Monte Carlo com 1000 realizações da abundância de aglomerados, para diferentes valores de $M_{\min} = 1 - 10 \times 10^{14} h^{-1} M_{\odot}$ e $\Delta z = 0, 1$ (14 bins). A linha azul corresponde ao método de Poisson e a linha vermelha ao método ULC. A diferença relativa entre as barras de erro desses métodos varia de 0 a 0.5%. Um pequeno desvio foi introduzido no eixo x para melhorar a visualização da figura.

obtemos os melhores ajustes minimizando a expressão $-2 \ln L$, com L dado pela equação (3.10), em relação ao parâmetro w_0 e mantendo os outros parâmetros fixos. Na figura 5, mostramos os resultados da média \bar{w}_0 e do desvio padrão $\sigma(\hat{w}_0)$ do estimador \hat{w}_0 para diferentes massas limites. Nós incluímos também o resultado de Poisson para comparar os dois procedimentos.

Notamos que os resultados das médias \bar{w}_0 são praticamente iguais para ambos os métodos, exceto para $M_{\min} = 10^{15} h^{-1} M_{\odot}$ em que o método ULC parece introduzir um viés ligeiramente menor do que o do método de Poisson. O desvio padrão $\sigma(\hat{w}_0)$ é também essencialmente o mesmo nos dois casos.

3.3.1 Comparação do método de Monte Carlo com a Matriz de Fisher e o Perfil de Verossimilhança

O método de Monte Carlo é bastante genérico e robusto não necessitando de muitas hipóteses sobre a verossimilhança e permitindo abordar questões como o viés em esti-

madores, como foi feito na seção anterior e na seção 3.2.3. Por outro lado, ele é um método bastante dispendioso do ponto de vista numérico. Para obter a estimativa de um dado parâmetro (ou conjunto de parâmetros) e sua incerteza, é preciso fazer inúmeras realizações simulando o conjunto de dados e, para cada realização, é preciso fazer uma minimização.

Nesta tese utilizaremos outros dois métodos, a saber, o *perfil de verossimilhança* (PV) e a *Matriz de Fisher* (MF), que são discutidos nos apêndices A.3 e A.4 respectivamente, para obter $\hat{\theta}$ e sua incerteza. Em ambos os casos, pode-se utilizar apenas um conjunto de dados, sem a necessidade de fazer várias realizações. Em geral, utilizam-se os valores dos observáveis produzidos a partir do modelo fiducial. Ou seja, é suficiente obter os valores médios, sem a necessidade de realizações. No caso das estatísticas binadas, bastaria calcular como observável a contagem esperada em cada bin (equação 3.1).

No caso do perfil da verossimilhança, nós resolvemos, primeiramente, a equação (A.11) para χ_ν^2 com nível de significância $p = 0.3173$ e $\nu = 1$ dado que estamos ajustando somente w_0 e, por isso, temos somente um grau de liberdade. Em seguida, substituímos esse valor de χ_ν^2 na equação (A.12) e calculamos a verossimilhança (equação 3.10) no melhor ajuste \hat{w}_0 ,

$$\chi_\nu^2 = -2 \ln \left(\frac{L(w_0)}{L(\hat{w}_0)} \right), \quad (3.11)$$

e, assim, encontramos os dois valores de w_0 que satisfazem essa equação. Dessa forma, nós obtemos o intervalo de confiança em 1σ de \hat{w}_0 .

No caso da matriz Fisher, calculamos (veja equação A.17)

$$\sigma^2(\hat{w}_0) = - \left(\frac{\partial^2 \ln(L(w_0))}{\partial w_0^2} \Big|_{\hat{w}_0} \right)^{-1}, \quad (3.12)$$

onde L é dado pela equação (3.10).

Note que as barras de erro do método perfil de verossimilhança não são, necessariamente, simétricas enquanto que, por construção, no método de matriz de Fisher elas são sempre simétricas.

Na tabela 2, mostramos os resultados obtidos para as barras de erro de \hat{w}_0 a partir dos métodos de Monte Carlo, PV e MF. Para os dois últimos, nós calculamos as equações (3.11) e (3.12), respectivamente, usando uma realização de $\{z_i\}$, das 1000 que geramos no método Monte Carlo, para cada valor de M_{\min} apresentado na tabela.

Verificamos que os métodos fornecem resultados semelhantes, sendo a barra de erro

$M_{\min}[10^{14}h^{-1}M_{\odot}]$	Monte Carlo	PV	MF
1	0,012	0,013	0,013
2	0,025	0,026	0,026
3	0,042	+0,045 -0,046	0,046
4	0,062	+0,067 -0,069	0,068
5	0,091	+0,095 -0,10	0,097
6	0,13	+0,15 -0,16	0,16
7	0,18	+0,18 -0,19	0,19
8	0,24	+0,22 -0,24	0,23
9	0,30	+0,29 -0,34	0,31
10	0,38	+0,42 -0,54	0,47

Tabela 2: Intervalo de confiança em 1σ do parâmetro w_0 obtido para dez diferentes valores de M_{\min} a partir dos métodos de Monte Carlo (segunda coluna), perfil de verossimilhança (terceira coluna) e matriz de Fisher (quarta coluna).

do método Monte Carlo sistematicamente um pouco menor que as dos outros métodos. Isso é esperado dado que a incerteza do método de Monte Carlo é obtida em torno dos valores fiduciais, enquanto os outros são obtidos em torno do melhor ajuste de uma realização somente. A partir de agora, nós utilizaremos somente os métodos do perfil de verossimilhança e matriz de Fisher para restringir os parâmetros cosmológicos.

Embora esses dois métodos sejam muito mais rápidos que o Monte Carlo, há diferenças importantes entre eles. No caso do método de MF (veja o apêndice A.4) é necessário fazer uma única minimização da verossimilhança (para um dado conjunto de dados) para obter os parâmetros que melhor ajustam os dados. As incertezas (“regiões de confiança”) são obtidas a partir das segundas derivadas da verossimilhança em relação aos parâmetros, calculadas no ponto de melhor ajuste. Portanto, i) no MF é preciso calcular a verossimilhança em poucos pontos, ii) as incertezas dependem da forma local da verossimilhança próximo ao seu mínimo. Portanto, esse é o método mais rápido, mas que utiliza apenas uma informação “local” da verossimilhança. Por construção, as regiões de confiança são elipses (no caso em que ajustamos dois parâmetros simultaneamente) e as barras de erro são simétricas.

Já no método PV (veja o apêndice A.3), a verossimilhança é calculada em muitos pontos no espaço de parâmetros de interesse e ainda é feita uma minimização para cada um desses pontos no restante dos parâmetros “livres”. Esse é portanto um método mais demorado, mas que incorpora informação “não-local” sobre a verossimilhança. Nesse caso, as “curvas de confiança” podem possuir uma forma arbitrária.

O método de MF é muito utilizado para fazer uma estimativa “rápida e grosseira” sobre

o poder de um método em limitar um dado conjunto de parâmetros. Já o PV permite fazer uma estimativa mais acurada e que é especialmente desejável no caso de se combinar as restrições obtidas por vários conjuntos de dados (veja por exemplo a seção 4.2). O método MF será utilizado em todas as análises desta tese, sendo que em alguns casos (como no capítulo 4 e na seção 5.4), utilizaremos também o método PV, comparando os resultados obtidos entre eles.

3.4 Limites para os parâmetros cosmológicos nos métodos de Poisson e ULC

O estudo feito nas seções 3.2.3 e 3.3 consistiu em ajustar o parâmetro w_0 mantendo todos os outros fixos. Para ampliar essa análise e, assim, comparar os métodos de Poisson e ULC levando em conta as incertezas em outros parâmetros cosmológicos, obtivemos regiões de confiança a partir do método de matriz de Fisher para diferentes combinações de parâmetros (planos (θ_1, θ_2)).

Para obter essas regiões, vamos considerar a equação de estado da energia escura dada pela equação (2.36), usar uma realização $\{z_i\}$ gerada com os mesmos parâmetros mencionados na seção 3.2.3 e $M_{\min} = 5 \times 10^{13} h^{-1} M_{\odot}$ e manter os parâmetros $\Omega_b = 0,04$, $\Omega_{DE} = 0,75$, $n_s = 1$ e $h = 0,7$ fixos.

No painel (a) da figura 6, mostramos as regiões de confiança de 1, 2 e 3σ no plano (Ω_m, σ_8) obtidas com os métodos de Poisson (equação 3.2) e ULC (equação 3.10). Em ambos os casos os parâmetros w_0 e w_1 foram marginalizados. Vemos que os resultados obtidos com esses dois métodos são praticamente idênticos. Notamos que a pequena diferença entre os métodos é, essencialmente, um deslocamento das regiões de confiança obtidas com os dois métodos. O mesmo também ocorre nos painéis (b) e (c), onde mostramos as regiões nos planos (Ω_m, w_0) e (w_0, w_1) , respectivamente, onde os dois parâmetros não mostrados em cada painel são marginalizados.

Nos três painéis da figura 6, podemos notar que os melhores ajustes do método ULC estão mais próximos dos valores fiduciais que os de Poisson. Para quantificar a diferença entre as incertezas esperadas a partir desses dois métodos, calculamos a razão das áreas das regiões de 1σ para os três planos, onde a área de cada contorno é dada por $A = \pi \sqrt{C_{ii}C_{jj} - C_{ij}^2}$ e os termos C_{ij} são dados pela equação (A.17) (veja a referência [102]). Dessa forma, temos que o método de Poisson apresenta uma área $\sim 6\%$ maior que a do método ULC no plano (Ω_m, σ_8) e $\sim 7\%$ maior nos planos (Ω_m, w_0) e (w_0, w_1) . Isso será

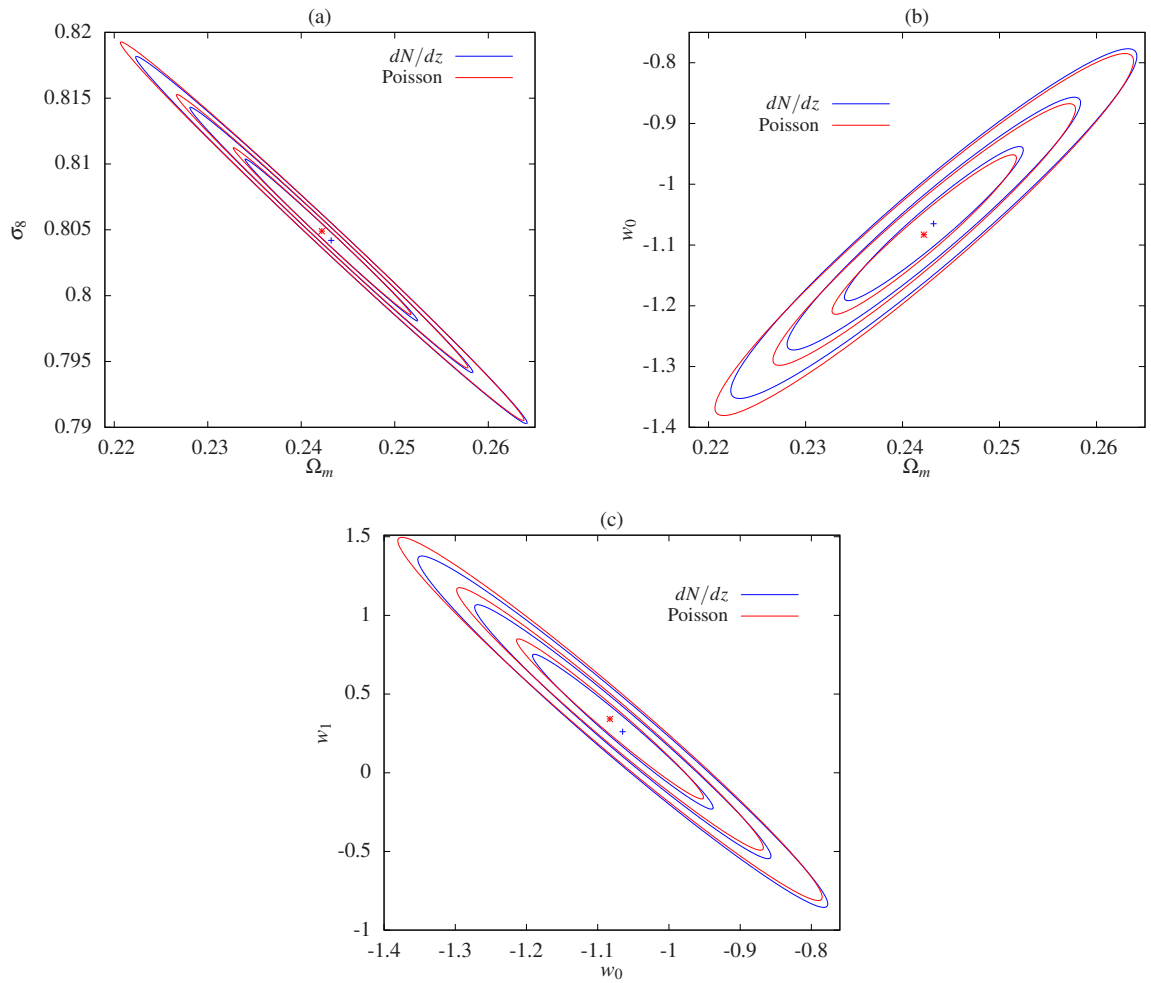


Figura 6: Regiões de confiança de 1, 2 e 3σ obtidas com o método de matriz Fisher nos planos (Ω_m, σ_8) , (Ω_m, w_0) e (w_0, w_1) . Painel (a): contornos obtidos marginalizando os parâmetros w_0 e w_1 . Painel (b): marginalização sobre σ_8 e w_1 . Painel (c): marginalização sobre Ω_m e σ_8 . Os parâmetros cosmológicos são os mesmos utilizados na simulação carmen [34], $M_{\min} = 5 \times 10^{13} h^{-1} M_{\odot}$, $0 \leq z \leq 1, 4$, $A_s = 5000 \text{deg}^2$. As curvas azuis foram obtidas utilizando a verossimilhança sem binagem em z (equação 3.10) e as curvas vermelhas com o método de Poisson com 14 bins em z .

investigado em mais detalhes no futuro, mas esse resultado parece indicar que o estimador obtido a partir do método ULC apresenta não só um viés menor, mas também uma menor variância em relação ao método de Poisson.

3.5 Verossimilhança para a abundância de aglomerados sem binagem na massa

Os aglomerados são caracterizados pelo desvio para o vermelho, assim como pela massa. Até agora consideramos o caso em que conhecemos somente um limite inferior da massa desses objetos, M_{\min} . No caso em que haja a disponibilidade de dados dessas duas grandezas, o cálculo da abundância pode ser feito de forma mais em *bins* menores de M . Seguindo a ideia apresentada na seção 3.3, propomos também uma verossimilhança em que levamos em conta tanto a posição z quanto o valor da massa de cada aglomerado, ou seja, uma verossimilhança sem binagem em z nem em M .

A probabilidade de um aglomerado de galáxias com massa entre M e $M + dM$ ser encontrado entre z e $z + dz$ é

$$P(M, z)dzdM = \frac{1}{N_{\text{total}}^{\text{teo}}} \frac{d^2 N(M, z)}{dzd \ln M} dzd \ln M, \quad (3.13)$$

onde

$$\frac{d^2 N(M, z)}{dzd \ln M} = A_{\text{survey}} \frac{dV}{dzd\Omega} \frac{dn(M, z)}{d \ln M}. \quad (3.14)$$

Com isso, construímos a verossimilhança na massa e no desvio para o vermelho com a expressão:

$$L(\{\theta_j\}, \{M_i, z_i\}) = P(N_{\text{total}}^{\text{obs}}; N_{\text{total}}^{\text{teo}}) \prod_{i=1}^{N_{\text{total}}^{\text{obs}}} \frac{1}{N_{\text{total}}^{\text{teo}}} \frac{d^2 N(M_i, z_i)}{dzd \ln M}, \quad (3.15)$$

onde M_i é a massa de cada aglomerado.

Da mesma maneira que precisamos gerar uma realização $\{z_i\}$ para utilizar a verossimilhança (3.10) na obtenção dos estimadores dos parâmetros cosmológicos, neste caso precisamos gerar uma realização $\{M_i, z_i\}$. O procedimento para obter essa realização é muito semelhante ao que apresentamos na seção 3.2.1, como descrevemos a seguir.

Geramos uma realização $\{z_i\}$ e, para cada valor de z_i , nós obtemos um valor M_i a partir da distribuição condicional de obter um aglomerado com massa entre M e $M + dM$ dado z_i , que é definida como

$$P(M|z_i) = \frac{P(M, z_i)}{\mathcal{P}(z_i)}, \quad (3.16)$$

onde $P(M, z_i)$ é dado pela equação (3.13) e $\mathcal{P}(z_i)$ pela equação (3.5). Com isso, usamos a cumulativa

$$f(M|z_i) = \int_0^M P(M'|z_i)dM', \quad (3.17)$$

para definir $M(f, z_i)$ e procedemos de maneira análoga à descrita na seção 3.2.1 para obter $\{M_i, z_i\}$.

Para estudar a restrição nos parâmetros cosmológicos a partir da verossimilhança (3.15), geramos uma realização $\{M_i, z_i\}$ a partir da realização $\{z_i\}$ utilizada na seção 3.4. Calculamos as regiões de confiança através do método de matriz de Fisher mantendo fixos os parâmetros Ω_b , Ω_{DE} , n_s e h (nos mesmos valores da seção anterior).

Na figura 7, mostramos as regiões de confiança de 1, 2 e 3σ no plano (Ω_m, σ_8) obtidas marginalizando nos parâmetros w_0 e w_1 . Para comparar esse resultado com o método ULC (equação 3.10), incluímos na figura os respectivos contornos mostrados no painel (a) da figura 6. Como esperado, o efeito de usar explicitamente a dependência da distribuição de aglomerados em função da massa e do desvio para o vermelho, ou seja, de considerar a massa de cada aglomerado ao invés de somente selecionar os objetos a partir de uma massa limite M_{\min} , reflete-se em restrições significativamente mais fortes nos parâmetros.

Observamos o mesmo comportamento nos painéis (b) e (c) da figura 7, onde mostramos os contornos de 1, 2 e 3σ nos planos (Ω_m, w_0) e (w_0, w_1) , respectivamente. Os parâmetros σ_8 e w_1 foram marginalizados para obtermos as regiões no painel (b) da figura 7. Analogamente, marginalizamos nos parâmetros Ω_m e σ_8 no caso do painel (c) da figura 7.

Verificamos portanto que utilizar o método sem binagem na massa nem no desvio para o vermelho possibilita recuperar os parâmetros fiduciais com uma precisão maior que nos métodos ULC e de Poisson usando apenas $\frac{dN}{dz}$. Para quantificar essa diferença, calculamos a razão das áreas das regiões de confiança de 1σ . As regiões de confiança obtidas levando em conta a dependência com M $\left(\frac{d^2N}{dzd\ln M}\right)$ são cerca de 50% menores do que os que levam em conta apenas a contagem em z $\left(\frac{dN}{dz}\right)$.

Apesar de se mostrar muito promissor, no restante desta tese não exploraremos o observável de abundância de aglomerados em função de z e da massa M . Vamos nos ater somente à contagem de aglomerados com $M > M_{\min}$ em função de z , que é o observável que tem sido mais utilizado na literatura para estudos da energia escura com aglomerados.

Deixamos para um futuro trabalho a validação da verossimilhança (3.15), utilizando o método de Monte Carlo e comparando com os métodos binados, a exemplo do que foi

feito neste capítulo para o método ULC.

Nesta tese investigaremos os limites nos parâmetros cosmológicos obtidos apenas a partir da abundância de aglomerados. Não utilizaremos nenhuma combinação com outros conjuntos de dados, como supernovas, nem distribuições a priori nos parâmetros obtidas a partir de outros observáveis. No entanto, é direto fazer a combinação das verossimilhanças da abundância de aglomerados com as de outros observáveis e isso será explorado no futuro.

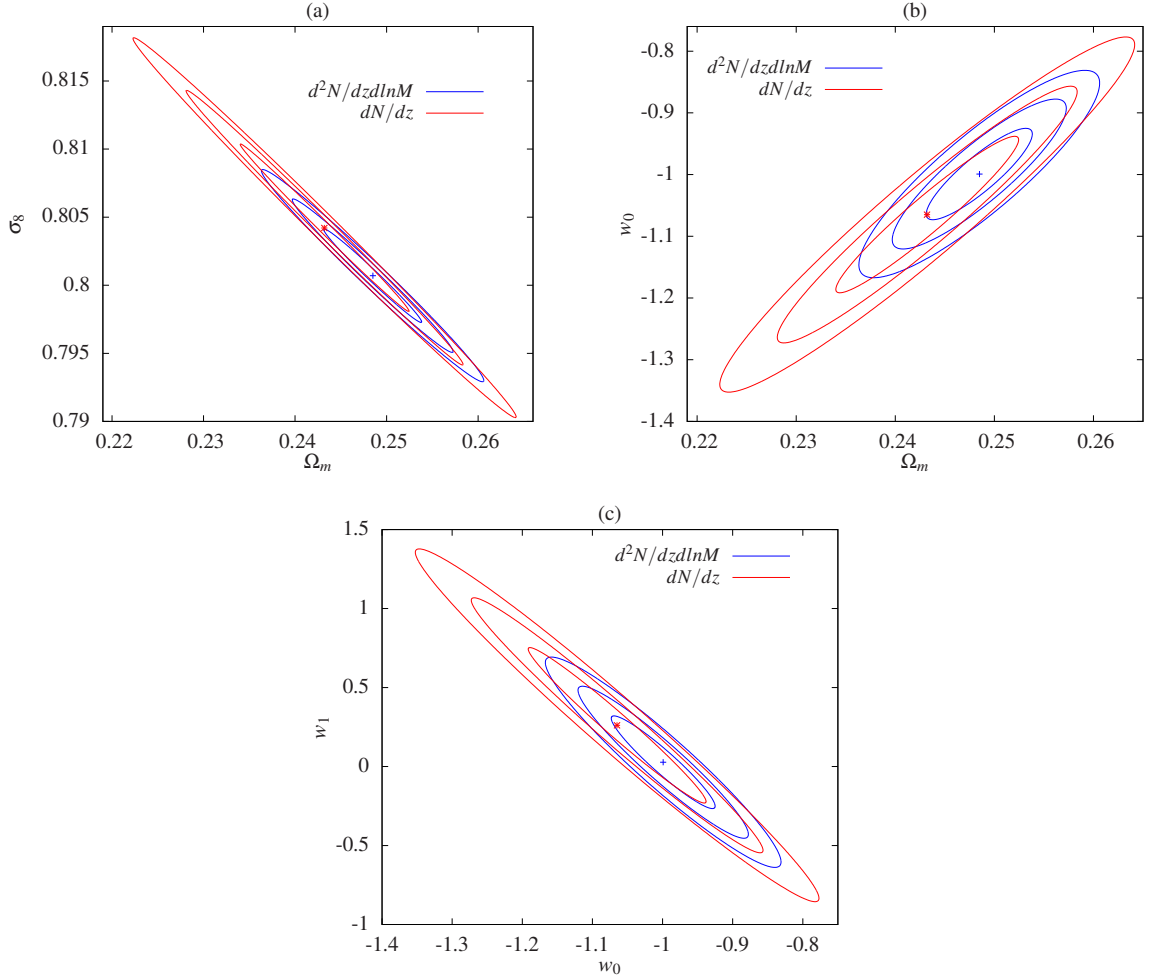


Figura 7: Regiões de confiança de 1, 2 e 3σ obtidas através do método de matriz de Fisher nos planos (Ω_m, σ_8) , (Ω_m, w_0) e (w_0, w_1) . Painel (a): contornos obtidos marginalizando nos parâmetros w_0 e w_1 . Painel (b): marginalização sobre σ_8 e w_1 . Painel (c): marginalização sobre Ω_m e σ_8 . Os parâmetros cosmológicos são os mesmos utilizados na simulação Carmen [34]. Os outros parâmetros utilizados foram $M_{\min} = 5 \times 10^{13} h^{-1} M_{\odot}$, $0 \leq z \leq 1,4$, $A_s = 5000 \text{ deg}^2$. As curvas azuis foram obtidas utilizando a verossimilhança sem binagem na massa nem em z , $\frac{d^2N}{dz d \ln M}$ e as curvas vermelhas com o método ULC (para dN/dz).

4 *Estudo da Combinação de Possíveis Catálogos de Aglomerados no SDSS*

Como mencionamos na seção 1.3, as três etapas do levantamento SDSS possuem medidas correspondentes a diferentes áreas e profundidades do céu que proporcionam catálogos de aglomerados. Na seção 4.1, investigamos as restrições nos parâmetros cosmológicos esperadas a partir desses catálogos. Na seção 4.2, fazemos uma análise combinada usando dois catálogos esperados do SDSS com diferentes áreas e profundidades.

4.1 **Comparação dos limites em Λ CDM para três amostras fictícias de aglomerados no SDSS**

Os catálogos de aglomerados de galáxias no SDSS, temos que:

- Já há catálogos que cobrem uma área de 8000 graus quadrados com uma profundidade de até $z_{\max} \simeq 0,3$ obtidos a partir dos dados das fases anteriores *Data Release 7* [103, 104].
- Há catálogos de aglomerados correspondentes a uma área de 270 graus quadrados referentes à região da faixa-82¹ (“Stripe 82”) e atingindo uma profundidade $z_{\max} \simeq 0,7$ [105].
- Os novos dados do SDSS-III, *Data Release 8*, permitirão a construção de novos catálogos em uma área de 10600 graus quadrados (área total do SDSS) e chegando até $z_{\max} \simeq 0,45$ devido a melhorias nos algoritmos de detecção de aglomerados [105].

¹A faixa-82 corresponde a uma região equatorial que foi imageada dezenas de vezes pelo projeto de supernovas do SDSS-II e para a qual foram construídos catálogos fotométricos mais profundos que no restante da área do projeto.

4.1 Comparação dos limites em Λ CDM para três amostras fictícias de aglomerados no SDSS43

Esses dados já estão disponíveis e espera-se que os novos catálogos de aglomerados estejam disponíveis em 2011.

Para o DR8, além de podermos usar um catálogo que identifica aglomerados até z mais altos, é importante enfatizar que teremos dados de espectroscopia do BOSS que cobrirão toda a área para galáxias vermelhas em mais alto desvio para o vermelho que no SDSS-I e II. Uma boa parte das galáxias centrais de aglomerados terá seus z obtidos espectroscopicamente, o que permitirá calibrar a relação $P(z^p|z)$ até valores mais elevados de z . Isso também contribuirá para a elaboração de catálogos de aglomerados mais profundos na área total do SDSS (DR8) com dados do BOSS. Além disso, várias propostas de observação foram e estão sendo feitas na área do SDSS que ajudarão a aumentar esse limite de z dos catálogos atuais do SDSS.

Geramos uma realização $R = \{z_i\}$ com os mesmos parâmetros cosmológicos da simulação Carmen ($\Omega_b = 0,04$, $\Omega_{DE} = 0,75$, $w_0 = -1$, $\sigma_8 = 0,8$ e $n_s = 1$), com $M_{\min} = 1 \times 10^{14} h^{-1} M_{\odot}$, área $A_s = 10600 \text{ deg}^2$, no intervalo $0 \leq z \leq 0,7$. A partir dela, obtemos os três catálogos com diferentes áreas e profundidades. Chamaremos esses catálogos hipotéticos de DR7, faixa 82 e DR8 baseados no que é discutido na seção 1.3. Portanto, o catálogo DR7 corresponderá a uma área $A_s = 8000 \text{ deg}^2$ e profundidade $z_{\max} = 0,3$, a faixa 82 terá $A_s = 270 \text{ deg}^2$ e $z_{\max} = 0,7$ e o DR8 terá $A_s = 10600 \text{ deg}^2$ e $z_{\max} = 0,45$. Para determinar a quantidade de aglomerados que irá compor cada um desses catálogos, nós calculamos os números médios de aglomerados $N_{\text{total}}^{\text{teo}}(\{\theta_j\})$, a partir da equação (2.34), com as áreas e profundidades desejadas. Em seguida, para levar em conta a flutuação de Poisson, utilizamos $N_{\text{total}}^{\text{teo}}(\{\theta_j\})$ como a média de uma distribuição de Poisson e sorteamos aleatoriamente um número $N_{\text{total}}^{\text{obs}}$ para cada um dos três casos. Seleccionamos, então, $N_{\text{total}}^{\text{obs}}$ valores de z_i do conjunto R no intervalo $0 \leq z \leq z_{\max}$ específico de cada caso. Em particular, obtivemos $N_{\text{total}}^{\text{obs}} = 7186$ para o DR7, $N_{\text{total}}^{\text{obs}} = 1275$ para a faixa 82 e $N_{\text{total}}^{\text{obs}} = 23251$ para o DR8.

Dado que temos três amostras distintas, utilizaremos o método de perfil de verossimilhança para estudar a restrição nos parâmetros cosmológicos, já que não esperamos que as regiões de confiança obtidas com essas realizações tenham o mesmo formato. No painel esquerdo da figura 8, mostramos as regiões de confiança de 1σ no plano (w_0, σ_8) do modelo Λ CDM (veja a seção 2.8) obtidas utilizando a verossimilhança ULC (equação 3.10) mantendo fixos os outros parâmetros cosmológicos, para os três tipos de catálogos considerados. Notamos que, apesar do catálogo DR7 ter ≈ 6 vezes mais objetos que o da faixa 82, a restrição nos parâmetros é comparável. Isso é devido ao catálogo DR7

ter profundidade máxima $z_{\max} = 0,3$ e, como pode ser visto na figura 2, a abundância de aglomerados até esse desvio para o vermelho varia pouco com a equação de estado da energia escura. Em relação a σ_8 , o resultado também é comparável devido à forte correlação entre os estimadores $\hat{\sigma}_8$ e \hat{w}_0 . No caso da configuração do DR8, notamos que a região de confiança fornece a maior restrição no plano (w_0, σ_8) , pois, além de possuir a maior área, a sua profundidade atinge um ponto mais sensível à equação de estado (veja a figura 2). Desse modo, é muito importante investir em algoritmos de busca de aglomerados que permitam aumentar a profundidade dos catálogos de aglomerados.

A mesma análise descrita acima foi feita utilizando o método de matriz de Fisher para comparar esses dois métodos. As regiões em 1σ são mostradas no painel direito da figura 8. Como os contornos do método de perfil da verossimilhança possuem uma forma elíptica, temos que, neste caso, a matriz de Fisher fornece uma boa aproximação para as regiões de confiança. Porém, vale ressaltar que as assimetrias das regiões em torno dos melhores ajustes são perdidas.

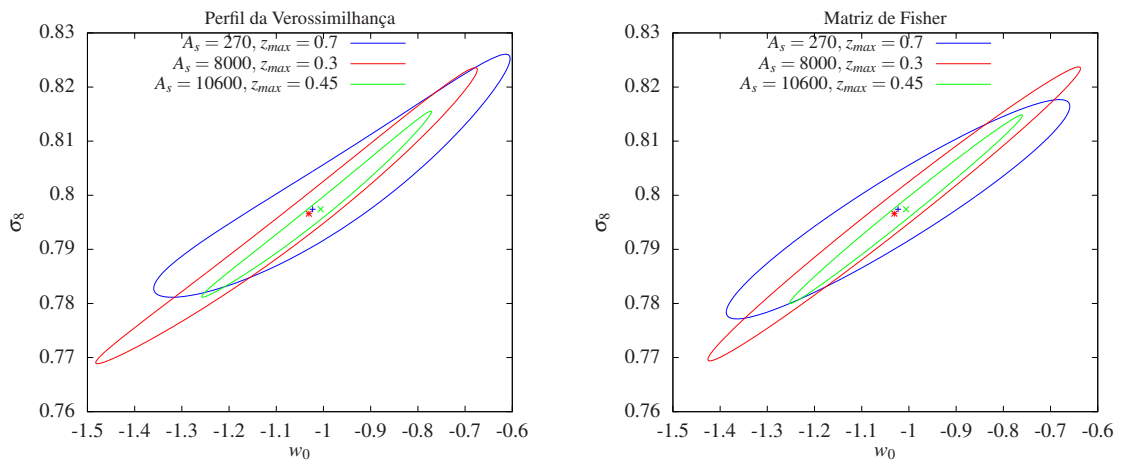


Figura 8: Regiões de confiança em 1σ no plano (w_0, σ_8) do modelo XCDM. Painel esquerdo: contornos obtidos com o método de perfil de verossimilhança. Painel direito: contornos obtidos com o método de matriz de Fisher. Essas regiões foram calculadas para diferentes valores da área do levantamento (A_s) e diferentes profundidades (z_{\max}), conforme a legenda.

4.2 Combinação de duas amostras fictícias de aglomerados no SDSS

Motivados pelo resultado apresentado na figura 8, que mostra a relevância de ter aglomerados identificados na faixa 82 devido a sua profundidade, fizemos um estudo

para prever o efeito de combinar as amostras da faixa 82 e DR8. Para evitar a dupla contagem dos objetos que pertencem aos dois catálogos, primeiramente, nós criamos um novo catálogo R' , a partir da realização original R com $A_s = 10600 \text{ deg}^2$ e $z_{\text{max}} = 0,7$, excluindo os 1275 z_i 's que compõem o catálogo da faixa 82. Seleccionamos, então, desse novo conjunto R' os desvios para o vermelho até $z_{\text{max}} = 0,45$. Com isso, obtemos uma amostra, que chamaremos de DR8', com $N_{\text{total}}^{\text{obs}} = 22679$ correspondente à área $A_s = 10330 \text{ deg}^2$ e profundidade $z_{\text{max}} = 0,45$.

A análise combinada é feita a partir da verossimilhança conjunta (*joint likelihood*) dada por

$$L(\{\theta_j\}, \{z_i\}) = L_1 \times L_2, \quad (4.1)$$

onde $\{z_i\}$ é o conjunto com 23954 (1275+22679) elementos, L_1 é a verossimilhança referente à faixa 82 e L_2 refere-se ao catálogo DR8', sendo ambas obtidas com o método ULC (equação 3.10). Calculamos a região de 1σ no plano (w_0, σ_8) para as verossimilhanças L , L_1 e L_2 tanto com o método de perfil da verossimilhança quanto com matriz de Fisher.

Na figura 9, apresentamos os resultados da análise combinada e, também, das regiões obtidas com a faixa 82 e DR8', onde mantivemos os outros parâmetros cosmológicos fixos. Como esperado, o efeito da combinação dos dados é diminuir a região de confiança na direção mais larga. Apesar dos contornos de confiança esperados da faixa 82 serem bem mais largos que os da amostra do DR8, a combinação das duas amostras ainda permite obter limites mais fortes no espaço (w_0, σ_8) .

Estudamos, também, o efeito de combinar as amostras da faixa 82 e DR8' em um caso mais geral onde o parâmetro Ω_m é deixado livre. Na figura 10, mostramos os contornos de 1σ no mesmo plano (w_0, σ_8) obtidos com as realizações da faixa 82, da amostra DR8' e da análise combinada. Ao deixar o parâmetro Ω_m livre, naturalmente, as regiões de confiança ficam maiores. É interessante notar que, no caso da faixa 82, os estimadores \hat{w}_0 e $\hat{\sigma}_8$ parecem ser anti-correlacionados, ao contrário do caso em que Ω_m é fixo. Note que os resultados obtidos com o DR8' e a faixa 82 são incompatíveis em 1σ , o que não é surpreendente posto que os conjuntos de dados são efetivamente independentes. Como a incerteza na faixa 82 é maior no parâmetro w_0 enquanto que no DR8' a maior degenerescência ocorre em σ_8 , temos que a análise combinada reduz essas duas incertezas, resultando em uma área bem menor no plano (σ_8, w_0) .

No que refere-se aos métodos de perfil da verossimilhança e matriz de Fisher, temos que as regiões de confiança da amostra DR8' evidenciam a diferença entre eles. Podemos dizer que a matriz de Fisher fornece a ordem de grandeza da barra de erro e da correlação

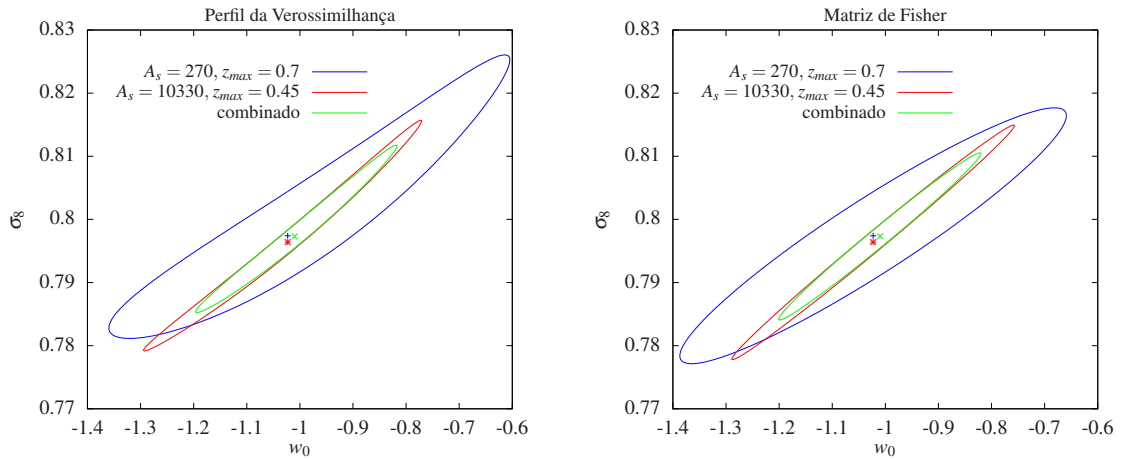


Figura 9: Regiões de confiança em 1σ no plano (w_0, σ_8) do modelo Λ CDM, fixando os outros parâmetros cosmológicos. Painel esquerdo: contornos obtidos com o método de perfil de verossimilhança. Painel direito: contornos obtidos com o método de matriz de Fisher. Essas regiões foram calculadas com as verossimilhanças da faixa 82, DR8' e a combinada.

dos estimadores. Porém, como por definição ela gera regiões elípticas em torno do melhor ajuste, não podemos obter informações mais detalhadas sobre o formato das regiões.

Os resultados da análise combinada mostram a complementariedade de se ter diferentes catálogos de aglomerados de galáxias tanto em grandes áreas quanto com uma profundidade maior.

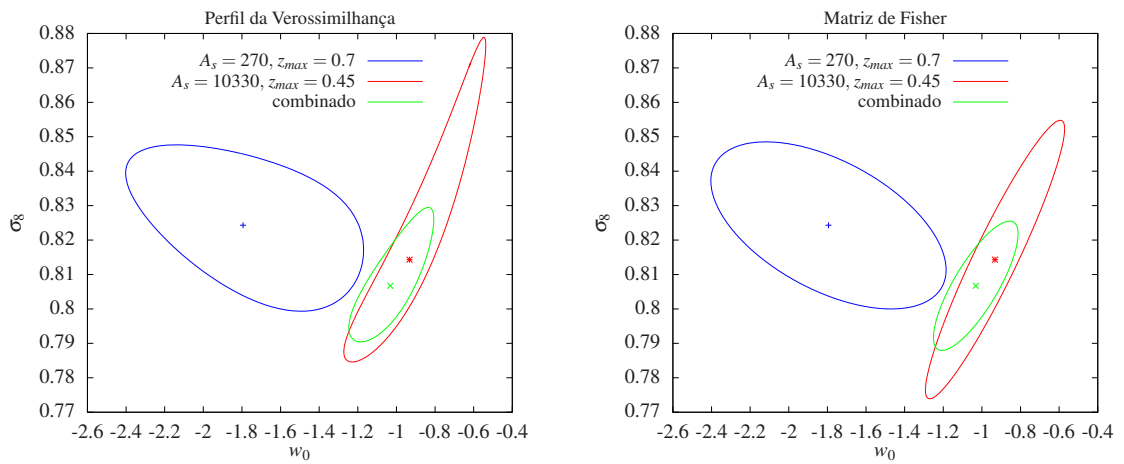


Figura 10: Regiões de confiança em 1σ no plano (w_0, σ_8) do modelo XCDM. Painel esquerdo: contornos obtidos com o método de perfil de verossimilhança, perfilando (veja a seção A.3) o parâmetro Ω_m . Painel direito: contornos obtidos com o método de matriz de Fisher, marginalizando sobre o parâmetro Ω_m . Essas regiões foram calculadas com as verossimilhanças da faixa 82, DR8' e a combinada.

5 *Validação do Catálogo de Halos do Dark Energy Survey*

Neste capítulo apresentamos os resultados que foram obtidos no processo de validação de um catálogo de halos para o DES e as restrições nos parâmetros cosmológicos obtidas a partir desse catálogo. Esse processo foi uma das etapas de um programa do grupo de aglomerados do DES denominado *visible cosmology challenge* que descrevemos sucintamente na seção 5.1. Nas seções 5.2 e 5.3, nós discutimos o procedimento para validar o catálogo de halos. Finalmente, na seção 5.4, apresentamos os primeiros resultados da restrição dos parâmetros cosmológicos para o DES com base em simulações (de N-corpos).

5.1 *Visible Cosmology Challenge*

Para testar e validar os algoritmos de identificação de aglomerados de galáxias, primeiramente, esses algoritmos são aplicados em catálogos de galáxias, gerados a partir de simulações de N-corpos em conjunto com o algoritmo ADDGALS[36], produzindo um catálogo de aglomerados. Esse catálogo é, então, usado para obter restrições nos parâmetros cosmológicos. Caso essas restrições não recuperem os valores fiduciais dos parâmetros das simulações, isso indicará problemas no algoritmo de identificação de aglomerados. Um trabalho da equipe de aglomerados da colaboração do DES, o *blind cosmology challenge* (BCC), consistirá em realizar o procedimento descrito acima, sem que os valores fiduciais dos parâmetros cosmológicos da simulação sejam conhecidos previamente.

Antes de ser feito o BCC, um outro trabalho está sendo realizado, é o chamado *visible cosmology challenge* (VCC). A metodologia é semelhante, no entanto, os parâmetros fiduciais são conhecidos. Como o catálogo de aglomerados de galáxias tem que ser compatível com o catálogo de halos, já que ambos são produzidos a partir da simulação de N-corpos, temos também que validar esses catálogos, ou seja, verificar se eles são consistentes com o modelo “teórico” para a abundância de halos (já que esse modelo também é obtido a

partir de simulações de N-corpos). Para tanto, iremos:

1. Comparar a densidade de halos por unidade de massa em um intervalo do desvio para o vermelho obtida com o catálogo de halos com a densidade obtida a partir da equação (2.6) com a função multiplicidade de Tinker et al. [10].
2. Calcular o *goodness of fit*, ou seja, o nível de concordância entre o catálogo e o modelo.
3. Ajustar parâmetros cosmológicos com estimadores obtidos com o método ULC.

5.2 Catálogo de halos

Os catálogos de halos produzidos pela colaboração do DES foram obtidos a partir da simulação Carmen¹ [34, 35] e dos algoritmos de identificação de halos FoF e SO (veja seção 2.5), sendo a massa dos halos (equação 2.24) definida com $\Delta = 200$ e $\rho_{ref} = \rho_{crit}$. Neste capítulo, apresentaremos o estudo da validação dos catálogos *DES_Mock_v2.11_halos* e *LC_halos_check*.

Como discutimos no seção 2.5, a função multiplicidade de Tinker et al. (equação 2.26) foi obtida com SO e com a massa dos halos definida em relação a $\bar{\rho}_m$. Portanto, para testar a consistência do catálogo de halos com a função de Tinker et al., é necessário adaptar a função multiplicidade à definição de massa com $\rho_{ref} = \rho_{crit}$. Para tanto, temos que redefinir Δ como $\Delta' = \Delta/\Omega_m(z)$, já que

$$\begin{aligned} M_\Delta &= \frac{4\pi}{3} \Delta \rho_{crit} R_\Delta^3 \\ &= \frac{4\pi}{3} \frac{\Delta}{\Omega_m(z)} \bar{\rho}_m R_\Delta^3, \end{aligned} \quad (5.1)$$

onde $\Omega_m(z) = \frac{\Omega_m(0)(1+z)^3}{E(z)^2}$.

Para converter a função multiplicidade de Tinker et al. para qualquer valor Δ' no intervalo $200 \leq \Delta' \leq 3200$, utilizamos os coeficientes, fornecidos na referência [10], para construir splines (veja seção B.1) de $A_0(\Delta')$, $a_0(\Delta')$, $b_0(\Delta')$, $c(\Delta')$ e $\log \alpha(\Delta') = - \left[\frac{0.75}{\log(\Delta'/75)} \right]^{1.2}$ e, dessa forma, obter a forma adequada (equação 2.26) para o catálogo de halos do VCC. Essa função de massa, dada pela função multiplicidade de Tinker et al. com $\Delta'(\Omega_m, z)$, define o que denominamos por *modelo teórico* neste capítulo.

¹Os parâmetros cosmológicos utilizados na simulação Carmen [34] foram $h = 0.7$, $\Omega_b = 0.04$, $\Omega_m = 0.25$, $\Omega_{DE} = 0.75$, $\sigma_8 = 0.8$, $n_s = 1$.

Para comparar o catálogo de halos com o modelo, obtemos a densidade do número de halos $n(M, z) = N/V$ no catálogo, onde V é o volume dado pela integral da equação (2.32), $A_s = 220\text{deg}^2$ e N é o número de halos no catálogo em um dado bin de M e z . Nós utilizamos 100 *bins* no logaritmo da massa (dada em unidades de $h^{-1}M_\odot$) igualmente espaçados entre $11 \leq \log_{10} M \leq 15$ e diferentes intervalos do desvio para o vermelho. Esse resultado foi comparado com a função de massa obtida a partir da equação (2.6), nos mesmos bins de massa e intervalos em z e com os mesmos parâmetros cosmológicos utilizados na simulação de N-corpos.

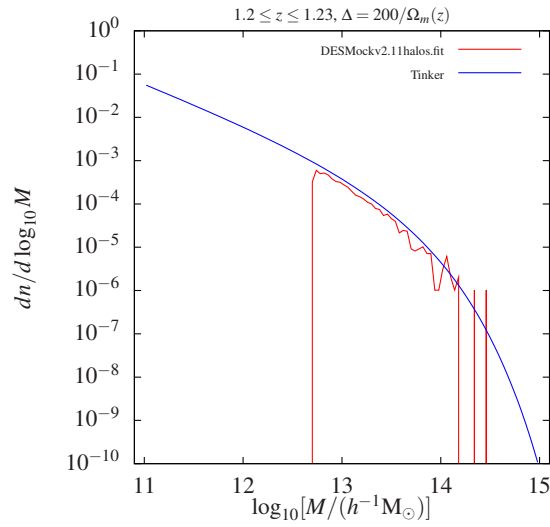


Figura 11: Densidade de halos obtida com o modelo (curva azul) e com o catálogo *DESMock_v2.11_halos* (curva vermelha) em 100 bins da massa e no intervalo $1.2 \leq z \leq 1.23$.

Primeiramente, comparamos a densidade $n(M, z)$ do catálogo de halos *DESMock_v2.11_halos*² com a função de massa ($dn/d \log_{10} M$) em diferentes intervalos do desvio para o vermelho. Na figura 11, apresentamos o resultado obtido para $1.2 \leq z \leq 1.23$. Notamos que a densidade do catálogo é menor que a do modelo para praticamente todo o intervalo de massa. Esse número inferior de halos no catálogo em relação ao modelo apresenta-se no intervalo $0.3 \lesssim z \leq 1.35$ para todas as massas. Verificamos também que a diferença aumenta com o aumento do desvio para o vermelho.

Relatamos esse problema ao grupo de trabalho de aglomerados do DES, especialmente, à pessoa responsável por gerar o catálogo de halos (Michael Busha, Standford).

²O primeiro catálogo de halos utilizado foi o *DESMock_v2.10_halos*. No entanto, como a diferença entre esse catálogo e o *DESMock_v2.11_halos* é que no segundo há mais halos para $z \gtrsim 1.25$, nós restringimos a discussão aos resultados deste último catálogo.

Ela identificou que, ao aplicar o algoritmo FoF³, estava sendo usado o comprimento de correlação comóvel entre as partículas do halo e não o comprimento físico (o que explica o fato da densidade do catálogo diferir mais do modelo para altos z 's). Com isso, foi feita a correção e produzido um novo catálogo, denominado *LC_halos_check*, com a finalidade específica de verificar a consistência com o modelo teórico (ou seja, com a função de massa de Tinker et al.).

Na figura 12, nós mostramos a densidade de halos obtida com o catálogo *LC_halos_check* e a função de massa teórica em quatro intervalos de desvio para o vermelho, $0.3 \leq z \leq 0.33$, $0.6 \leq z \leq 0.63$, $0.9 \leq z \leq 0.93$ e $1.2 \leq z \leq 1.23$. Comparando o painel inferior direito com a figura 11, percebemos que a densidade de halos é maior no novo catálogo, ficando significativamente mais próxima do obtido com o modelo. Em todos os casos há um intervalo de massa para o qual o catálogo parece estar visualmente de acordo com a previsão.

5.3 Completeza do catálogo de halos

Os algoritmos de identificação de halos tornam-se mais eficientes quanto maior a massa dos halos, já que o número de partículas que os compõem cresce. A fim de utilizar um catálogo de halos para estudar a restrição dos parâmetros cosmológicos, por exemplo, precisamos definir o menor valor da massa do halo (M_{\min}) para o qual obtemos uma boa completeza desse catálogo.

Notamos, na figura 12, que a completeza do catálogo ocorre para massas $M \gtrsim 10^{12.2} h^{-1} M_{\odot}$ nos quatro intervalos de z , já que as curvas da densidade do catálogo e do modelo passam a ser da mesma ordem. Mas, para quantificar o nível de concordância entre o catálogo e o modelo, nós calculamos o *goodness of fit* para diferentes massas inferiores nos quatro intervalos de z apresentados na figura. Para tanto, nós usamos dois métodos, o teste de Pearson (usualmente chamado de teste χ^2) e o *teste da razão de verossimilhança* (veja seção A.2).

O teste de Pearson consiste em calcular

$$\chi^2_{\nu} = \sum_{i=1}^n \frac{(N_i - N_i^{\text{teo}})^2}{N_i^{\text{teo}}}, \quad (5.2)$$

³Como mencionamos na seção 2.5, os halos podem ser identificados com o algoritmo FoF e, a partir dos locais de potencial mínimo, aplicar o algoritmo SO. Esse é o procedimento utilizado na produção dos catálogos de halos do DES.

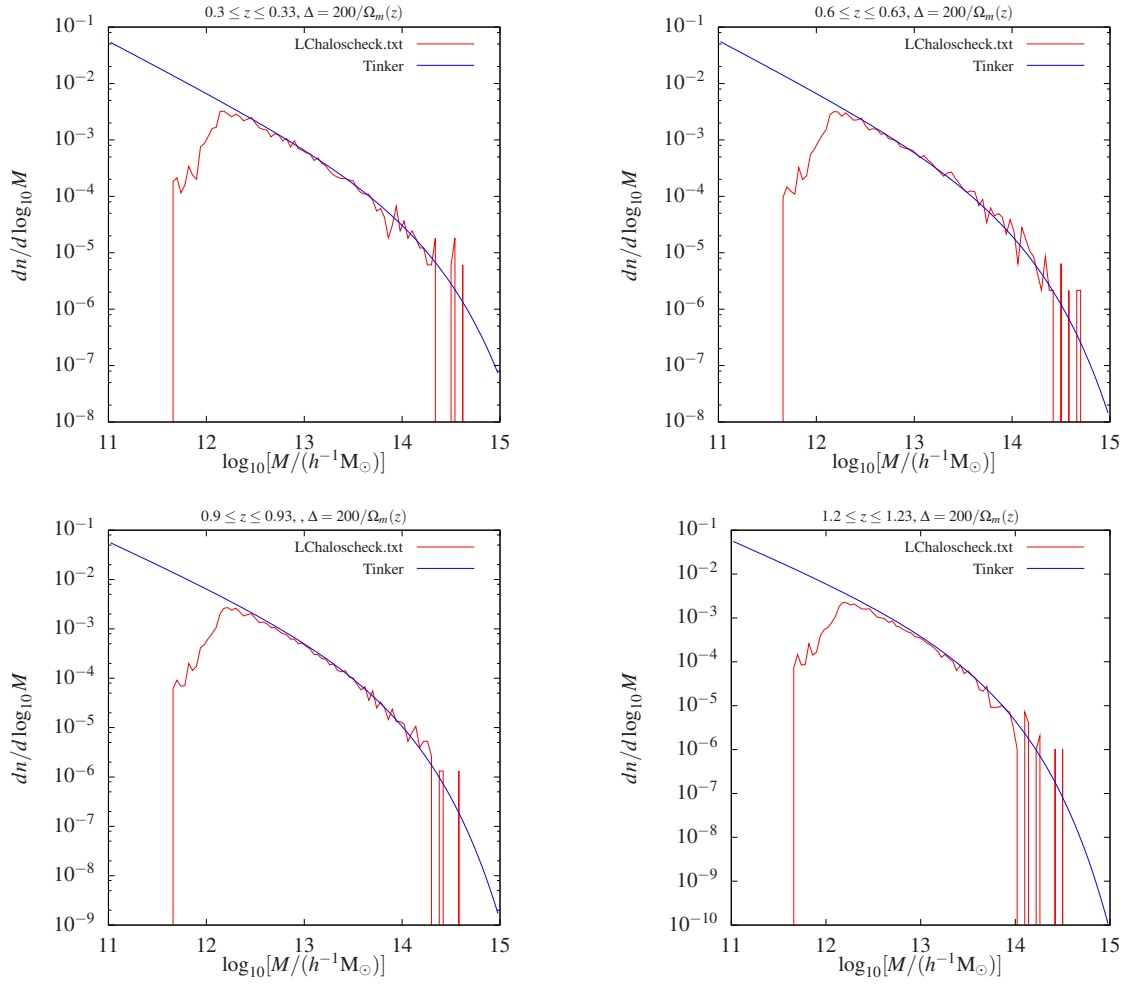


Figura 12: Densidade de halos obtida com o modelo (curva azul) e com o catálogo *LC_halo_check* (curva vermelha) em 100 bins da massa. Cada painel corresponde a um intervalo do desvio para o vermelho, tendo todos $\Delta z = 0.03$.

onde o número de graus de liberdade ν é igual ao número de bins na massa ($\nu = n$), N_i e N_i^{teo} são os números de halos no i -ésimo bin de massa do catálogo e do modelo, respectivamente, calculados com os mesmos parâmetros cosmológicos.

No teste da razão de verossimilhança, nós calculamos

$$\frac{L}{L_0} = \prod_{i=1}^n \left(\frac{(N_i^{\text{teo}})^{N_i} e^{-N_i^{\text{teo}}}}{N_i!} \frac{N_i!}{N_i^{N_i} e^{-N_i}} \right), \quad (5.3)$$

e, dessa forma,

$$\chi_\nu^2 = -2 \ln \left(\frac{L}{L_0} \right) = -2 \sum_{i=1}^n \left[N_i - N_i^{\text{teo}} + N_i \ln \left(\frac{N_i^{\text{teo}}}{N_i} \right) \right]. \quad (5.4)$$

Na tabela 3, nós apresentamos o resultado do *goodness of fit* χ_ν^2/ν , entre o catálogo e o

modelo para $M_{\min} = 5 \times 10^{12}$, 10^{13} , 5×10^{13} e $10^{14}h^{-1}M_{\odot}$, que correspondem aos números de bins na massa $\nu = 58$, 50 , 33 e 25 , respectivamente. O primeiro número corresponde ao teste de Pearson e o segundo refere-se ao teste da razão de verossimilhança. Notamos que a concordância⁴ entre catálogo e modelo torna-se pior com aumento do desvio para o vermelho.

$z/M_{\min}[h^{-1}M_{\odot}]$	$5 \times 10^{12} (\nu = 58)$	$1 \times 10^{13} (\nu = 50)$	$5 \times 10^{13} (\nu = 33)$	$1 \times 10^{14} (\nu = 25)$
$0.3 \leq z \leq 0.33$	0.86/0.91	0.66/0.78	0.73/0.91	0.56/0.80
$0.6 \leq z \leq 0.63$	1.72/1.89	1.75/1.70	1.48/1.49	1.13/1.20
$0.9 \leq z \leq 0.93$	1.94/2.06	1.07/1.15	0.68/0.75	0.55/0.92
$1.2 \leq z \leq 1.23$	3.68/4.05	1.68/1.94	1.00/1.26	0.91/1.18

Tabela 3: *Goodness of fit*, χ_{ν}^2/ν , entre o catálogo e o modelo, calculado para quatro limites de massa e quatro diferentes intervalos no desvio para o vermelho. Os valores x/y correspondem ao *goodness of fit* calculado com o teste de Pearson e o teste da razão de verossimilhança, respectivamente.

No entanto, o intervalo em z que requer o maior valor da massa limite para alcançar uma boa concordância entre catálogo e modelo é $0.6 \leq z \leq 0.63$. Isso ocorre porque o número de halos neste intervalo é maior e, portanto, os valores de χ_{ν}^2 obtidos com as equações (5.2) e (5.4) são também maiores.

Nós estimamos a completeza do catálogo para cada intervalo em z escolhendo os valores de massa mais baixo possíveis para os quais $\chi_{\nu}^2/\nu \simeq 1$. Na tabela 4, mostramos esses valores de M_{\min} e, também, o nível de significância (veja equação A.11) de obter o catálogo dado o modelo.

Apesar dos dois métodos fornecerem, essencialmente, os mesmos resultados para o *goodness of fit*, notamos que no método de Pearson os valores são sistematicamente maiores. Por exemplo, no intervalo $1.2 \leq z \leq 1.23$ e $\nu = 33$, o *goodness of fit* do teste de Pearson é $\chi_{\nu}^2/\nu = 1$ com um nível de significância de 46%, enquanto o do teste da razão da verossimilhança, $\chi_{\nu}^2/\nu = 1.26$, está fora de 1σ dado que $\sigma_{\chi_{\nu}^2} = \sqrt{\frac{2}{\nu}} \simeq 0.246$, tendo um nível de significância de 14%. Essa diferença pode ser entendida pelo fato que quando $N_i \lesssim 10$ a aplicação do teste de Pearson não é adequada por razões análogas às discutidas na seção 3.2.

⁴Para interpretar o resultado do *goodness of fit* χ^2/ν , ou seja, entender o que este valor diz sobre a concordância entre o catálogo e o modelo, vejamos algumas propriedades da distribuição $\chi^2(x, \nu)$. Essa distribuição tem média $\langle \chi^2(x, \nu) \rangle = \nu$ e variância $V = 2\nu$. Com isso, temos que um *goodness of fit* que indica uma concordância dentro de 1σ terá um valor

$$\frac{\chi^2}{\nu} \simeq 1 \pm \sqrt{\frac{2}{\nu}}.$$

z	$M_{\min}[h^{-1}M_{\odot}]$	Nível de Significância
$0.3 \leq z \leq 0.33$	5×10^{12}	86% / 67%
$0.6 \leq z \leq 0.63$	1×10^{14}	30% / 23%
$0.9 \leq z \leq 0.93$	1×10^{13}	37% / 24%
$1.2 \leq z \leq 1.23$	5×10^{13}	46% / 14%

Tabela 4: Nível de significância de obter o catálogo, dado o modelo, para a completeza a partir de M_{\min} em quatro intervalos do desvio para o vermelho. Os números $x\%/y\%$ correspondem aos níveis de significância calculados a partir de χ^2_{ν}/ν obtidos com teste de Pearson e o teste da razão de verossimilhança, respectivamente.

A partir desses resultados vemos que a massa mínima que fornece uma boa completeza em qualquer valor de z é $M_{\min} \simeq 10^{14}h^{-1}M_{\odot}$.

5.4 Cosmologia com o catálogo de halos

Tendo como principal motivação usar o catálogo de halos para fazer previsões para o DES baseadas no resultado de simulações, e não na amostra obtida a partir da distribuição dN/dz , nós estudamos as restrições nos parâmetros Ω_m , σ_8 e w_0 , obtidas com o catálogo *LC_halos_check* e supondo o modelo Λ CDM para a equação de estado da energia escura.

Estudamos os planos (Ω_m, σ_8) , (Ω_m, w_0) e (w_0, σ_8) usando a verossimilhança sem binagem no desvio para o vermelho (equação 3.10) em dois casos. No primeiro, calculamos as regiões de confiança nesses planos a partir do método de matriz de Fisher (veja a seção A.4) mantendo fixos, respectivamente, os parâmetros w_0 , σ_8 e Ω_m . No segundo caso, as regiões de confiança são obtidas marginalizando nos parâmetros não mostrados (ou seja, w_0 , σ_8 e Ω_m , respectivamente).

Na figura 13, mostramos os contornos de 1, 2 e 3σ no plano (Ω_m, σ_8) em ambos os casos (isto é, fixando w_0 e, também, mantendo-o livre) para as massas limites $M_{\min} = 5 \times 10^{12}$, 5×10^{13} , 10^{14} e $2 \times 10^{14}h^{-1}M_{\odot}$. Já nas figuras 14 e 15, nós mostramos os resultados da matriz de Fisher referentes aos planos (Ω_m, w_0) e (w_0, σ_8) , respectivamente, para os mesmos valores de M_{\min} .

Notamos nas três figuras que, para $M_{\min} = 5 \times 10^{12}h^{-1}M_{\odot}$, os parâmetros fiduciais do catálogo não são recuperados com mais de 3σ de confiança tanto quando os parâmetros estão fixos quanto quando são marginalizados.

Esse fato também ocorre para $M_{\min} = 5 \times 10^{13}h^{-1}M_{\odot}$ quando fixamos todos os outros parâmetros cosmológicos. Já quando o terceiro parâmetro é deixado livre, isto é, margi-

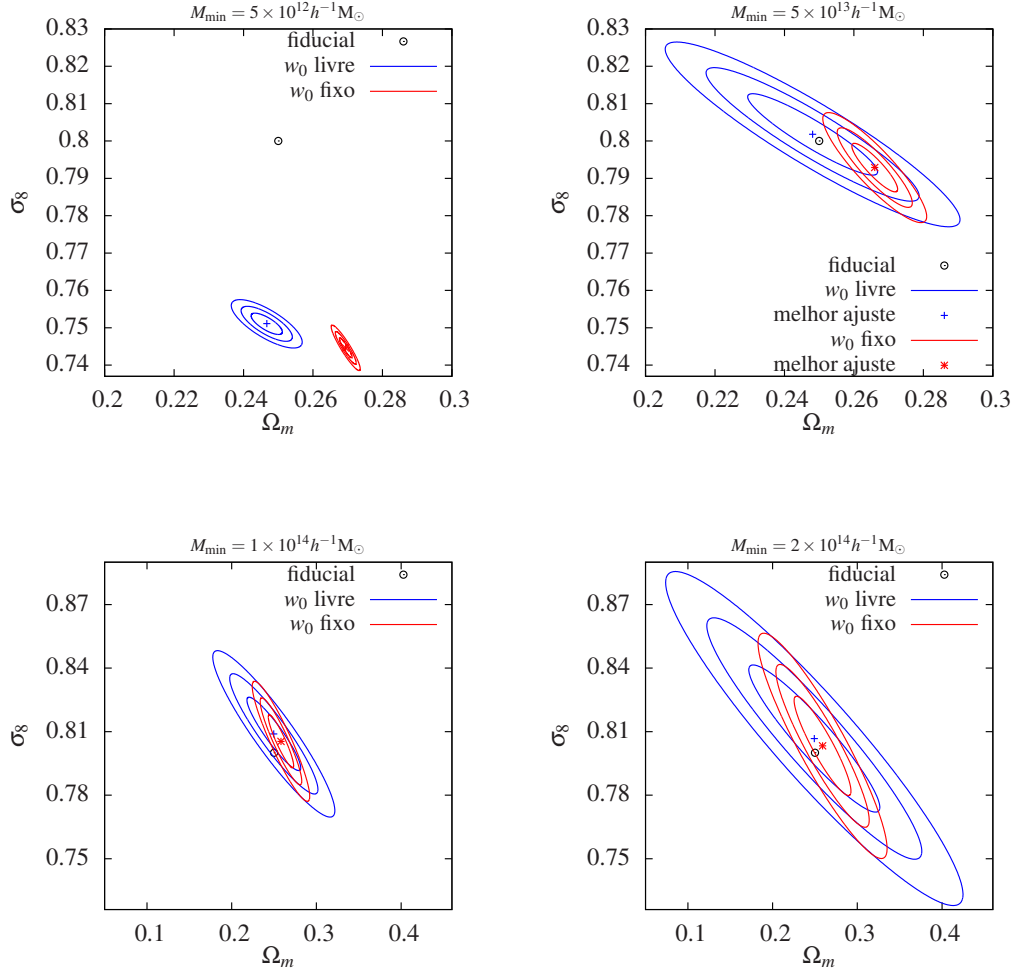


Figura 13: Contornos de 1, 2 e 3σ no plano (Ω_m, σ_8) obtidos com o método de matriz de Fisher fixando os outros parâmetros cosmológicos (curvas vermelhas) e deixando w_0 livre (curvas azuis). Cada painel corresponde a uma massa limite mínima utilizada.

nalizando sobre w_0 (figura 13), σ_8 (figura 14) e Ω_m (figura 15), naturalmente, ocorre um aumento nas áreas das regiões de confiança e os resultados, nos três planos, são consistentes com o modelo fiducial em 3σ , ou seja, os valores fiduciais encontram-se dentro dos contornos até 3σ para massas a partir de $M_{\min} = 5 \times 10^{13} h^{-1} M_{\odot}$.

Mantendo os parâmetros cosmológicos fixos, só recuperamos os parâmetros fiduciais dentro da região $\leq 2\sigma$ quando $M_{\min} = 2 \times 10^{14} h^{-1} M_{\odot}$, onde obtemos os maiores intervalos nos parâmetros Ω_m , σ_8 e w_0 , dado que um número menor de objetos fornece uma restrição mais fraca nesses parâmetros.

Como só obtemos uma boa concordância entre catálogo e modelo para massas a partir de $M_{\min} \simeq 10^{14} h^{-1} M_{\odot}$, entendemos o porquê de não recuperarmos os parâmetros fiduciais

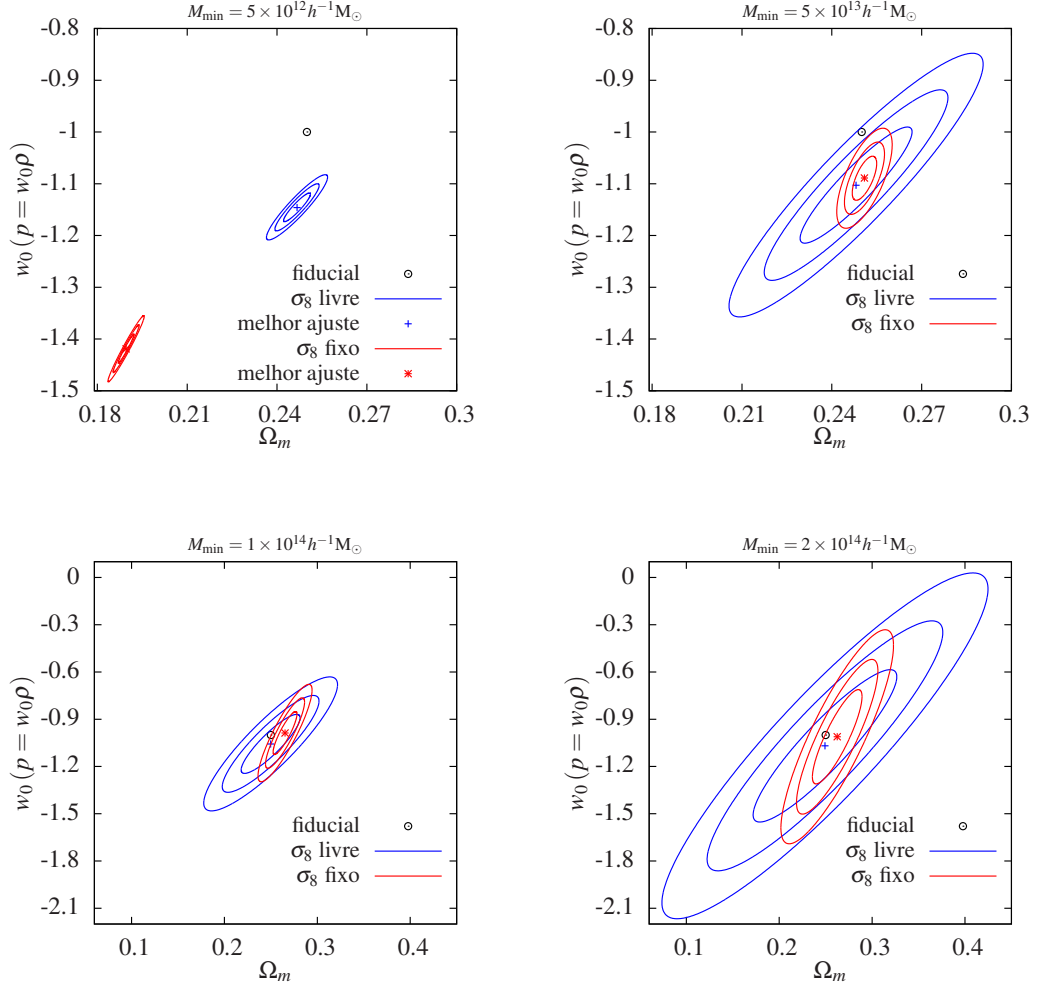


Figura 14: Contornos de 1, 2 e 3σ no plano (Ω_m, w_0) obtidos com o método de matriz de Fisher fixando os outros parâmetros cosmológicos (curvas vermelhas) e deixando σ_8 livre (curvas azuis). Cada painel corresponde a uma massa limite mínima utilizada, conforme indicado na parte superior de cada um deles.

com confiança $> 3\sigma$ quando usamos $M_{\min} \simeq 5 \times 10^{12} h^{-1} M_{\odot}$. Esses resultados nos levaram a estudar com mais detalhes o catálogo *LC_halos_check* e verificamos que a massa dos halos é subestimada nos desvios para o vermelho $z \simeq 1.18$ e $z \simeq 1.25$, como mostramos na figura 16. Isso reforça o fato de que o estudo da completeza do catálogo de halos é essencial para que possamos utilizá-los como referência para checar os resultados da restrição dos parâmetros cosmológicos obtidos com o catálogo de aglomerados.

A matriz de Fisher é um método excelente para obter uma primeira estimativa das regiões de confiança, já que ela fornece somente o menor limite da variância dos estimadores, como dado pela equação (A.16). Se desejamos fazer um estudo mais detalhado

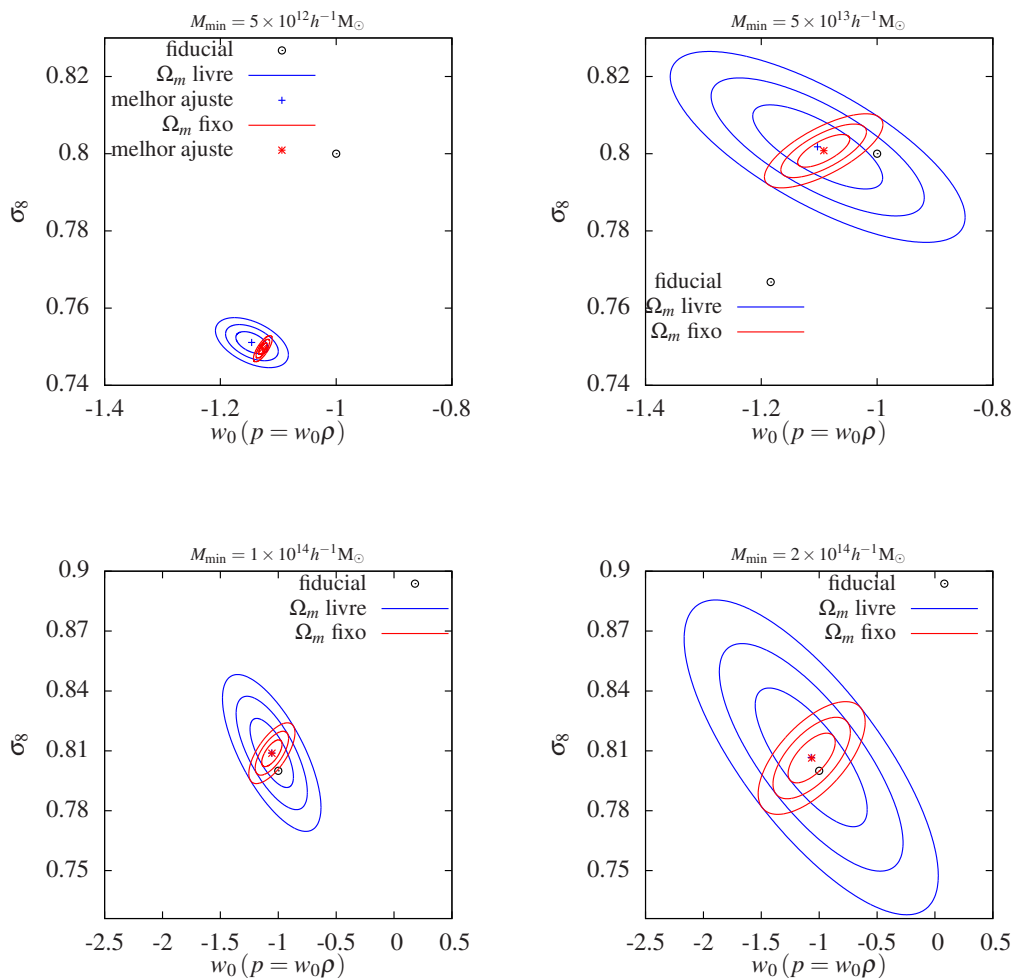


Figura 15: Contornos de 1–3 σ no plano (w_0, σ_8) obtidos com o método de matriz de Fisher fixando os outros parâmetros cosmológicos (curvas vermelhas) e deixando Ω_m livre (curvas azuis). Cada painel corresponde a uma massa limite mínima utilizada.

da restrição dos parâmetros cosmológicos devemos usar um outro método. Para finalizar esta etapa do VCC para o catálogo de halos, nós obtivemos algumas regiões de confiança com o método de perfil de verossimilhança.

Na figura 17, mostramos os contornos de 1 e 2 σ nos planos (Ω_m, σ_8) , (Ω_m, w_0) e (w_0, σ_8) , perfilando (deixando livres) os parâmetros w_0 , σ_8 e Ω_m , respectivamente. Comparamos esses resultados com as regiões obtidas anteriormente com a matriz de Fisher. Neste caso, diferentemente do resultado do capítulo 4, fica evidente como podemos perder informações sobre a correlação entre dois parâmetros e suas barras de erro ao se usar somente o método da matriz de Fisher.

Vale ressaltar que os resultados apresentados nesta seção são as primeiras previsões

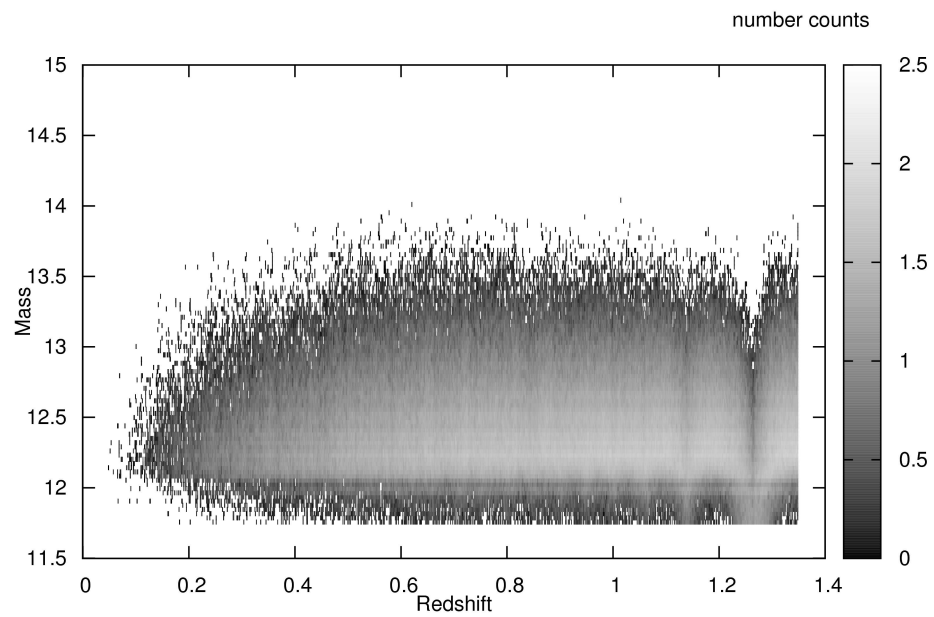


Figura 16: Abundância de halos do catálogo *LC_halos_check* divididos em 100 *bins* na massa e 1350 *bins* no desvio para o vermelho.

utilizando simulações para o DES.

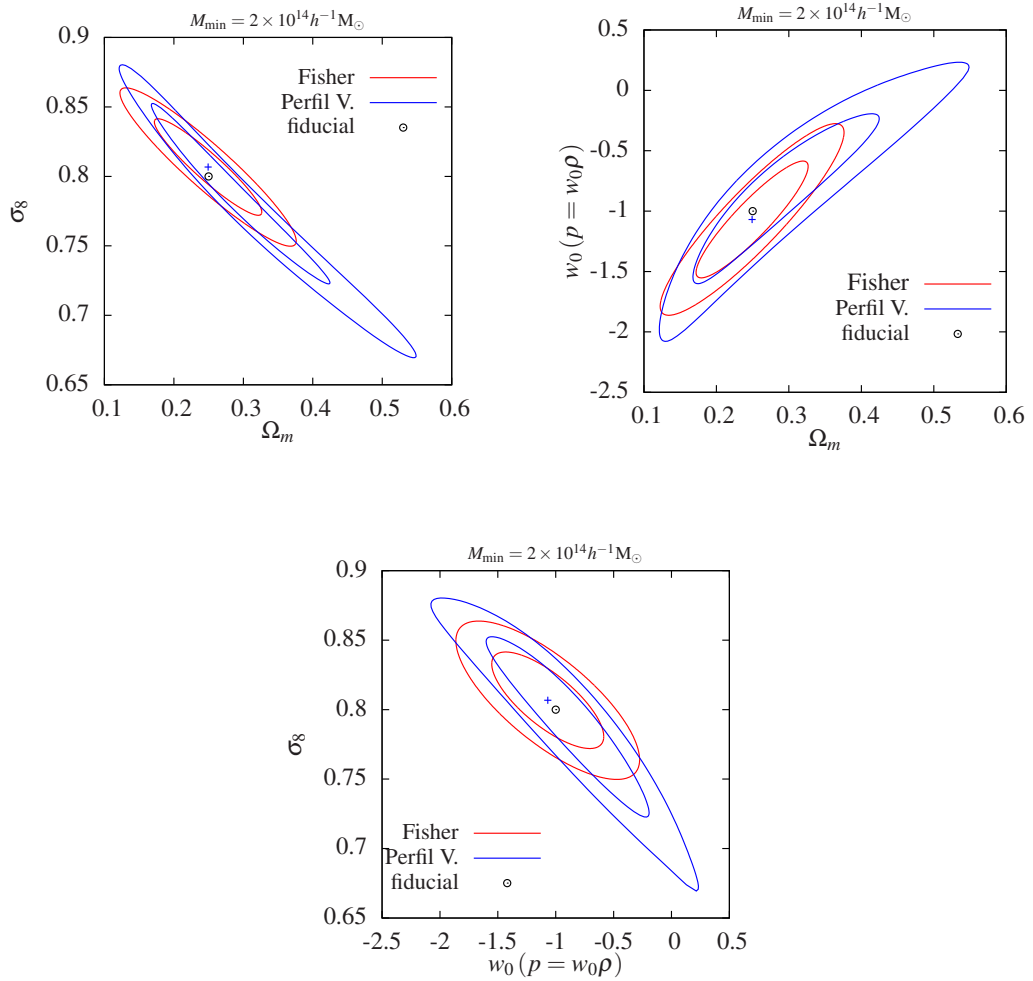


Figura 17: Contornos de 1-2 σ nos planos (Ω_m, σ_8) , (Ω_m, w_0) , (w_0, σ_8) obtidos com o método perfil de verossimilhança (curvas azuis) e matriz de Fisher (curvas vermelhas) para $M_{\min} = 2 \times 10^{14} h^{-1} M_{\odot}$ e deixando livres os parâmetros w_0 , σ_8 e Ω_m , respectivamente.

6 *Inclusão de Efeitos Observacionais nos Estimadores da Contagem de Aglomerados de Galáxias*

O cálculo do número de aglomerados a partir da Eq. (3.1) pressupõe que os desvios para o vermelho e as massas dos aglomerados sejam conhecidos perfeitamente. No entanto, quando utilizamos dados observacionais para estimar a abundância de aglomerados, temos que levar em conta as incertezas das medidas desses dados na modelagem teórica.

Nas seções 6.1 e 6.2, discutimos, respectivamente, a inclusão das incertezas nas medidas do desvio para o vermelho fotométrico e da massa e no cálculo da abundância de aglomerados. Na seção 6.3, construímos a verossimilhança (3.10) considerando essas incertezas e discutimos os efeitos na restrição dos parâmetros cosmológicos.

6.1 *Inclusão da incerteza do desvio para o vermelho fotométrico*

Ao considerar as medidas dos desvios para o vermelho fotométricos $\{z_i^p\}$, a abundância de aglomerados em função de z^p é dada por[106]

$$\frac{dN^p}{dz^p} = \int_0^\infty dz \frac{dN}{dz} p(z^p|z), \quad (6.1)$$

onde $p(z^p|z)$ é a densidade de probabilidade de medir um aglomerado entre z^p e $z^p + dz^p$ cujo desvio para o vermelho verdadeiro é z . Portanto, a abundância no k -ésimo *bin* em z^p é

$$N_k = \int_{z_k^p}^{z_{k+1}^p} dz^p \frac{dN^p}{dz^p}, \quad (6.2)$$

onde $z_k^p \leq z^p \leq z_{k+1}^p$ é o intervalo do k -ésimo *bin*.

No caso de uma distribuição de probabilidade Gaussiana, temos

$$p(z^p|z) dz^p = \sqrt{\frac{2}{\pi}} \frac{\exp[-y^2(z^p)]}{\sigma_z(1 - \operatorname{erf}(y_{z^p=0}))} dz^p, \quad z^p > 0, \quad (6.3)$$

onde

$$y(z^p) = \frac{z^p - z - z^{bias}}{\sqrt{2\sigma_z^2}}, \quad (6.4)$$

$\sigma_z = [\sum(z^p - z^{true})^2]^{1/2}$ e $z^{bias} = \sum(z^p - z^{true})$. Dessa forma, fazendo a transformação de variável $z^p \rightarrow y$,

$$dy = \frac{1}{\sqrt{2\sigma_z^2}} dz^p, \quad (6.5)$$

temos que,

$$\begin{aligned} N_i &= \int_0^\infty dz \frac{dN}{dz} \frac{2}{\sqrt{\pi}} \frac{1}{(1 - \operatorname{erf}(y_{z^p=0}))} \int_{y_i}^{y_{i+1}} dy e^{-y^2} \\ &= \int_0^\infty dz \frac{dN}{dz} \frac{2}{\sqrt{\pi}} \frac{1}{(1 - \operatorname{erf}(y_{z^p=0}))} \left[\int_0^{y_{i+1}} dy e^{-y^2} - \int_0^{y_i} dy e^{-y^2} \right] \\ &= \int_0^\infty dz \frac{dN}{dz} \left(\frac{\operatorname{erf}(y_{i+1}) - \operatorname{erf}(y_i)}{1 - \operatorname{erf}(y_{z^p=0})} \right). \end{aligned} \quad (6.6)$$

Na figura 18, mostramos a abundância em *bins* no desvio para o vermelho fotométrico, onde $\Delta z^p = 0.04$ e $z^{bias} = 0$. Notamos que, com a incerteza em z^p , existe a tendência de tornar a distribuição dos desvios para o vermelho dos aglomerados constante, já que $p(z^p|z) \rightarrow 1$ quando $\sigma_z \rightarrow \infty$.

6.2 Inclusão da incerteza na massa no cálculo da contagem de aglomerados

Neste trabalho, vamos supor que a densidade de probabilidade de medir a massa de um aglomerado, a partir de um observável, com valor entre M^{obs} e $M^{obs} + dM^{obs}$, dado que a massa verdadeira é M , é dada por uma distribuição log-normal [107, 108],

$$p(M^{obs}|M) dM^{obs} = \frac{1}{\sqrt{2\pi}\sigma_{\ln M}} \exp^{-x^2(M^{obs})} d \ln M^{obs}, \quad -\infty < \ln M^{obs} < \infty, \quad (6.7)$$

onde

$$x(M^{obs}) = \frac{\ln M^{obs} - \ln M - \ln M_{bias}(M^{obs}, z)}{\sqrt{2}\sigma_{\ln M}(M^{obs}, z)}, \quad (6.8)$$

e, de forma geral, o viés $\ln M_{bias}$ e a variância $\sigma_{\ln M}$ são funções de M^{obs} e z [107].

Com isso, temos que a densidade de aglomerados no desvio para o vermelho z e com

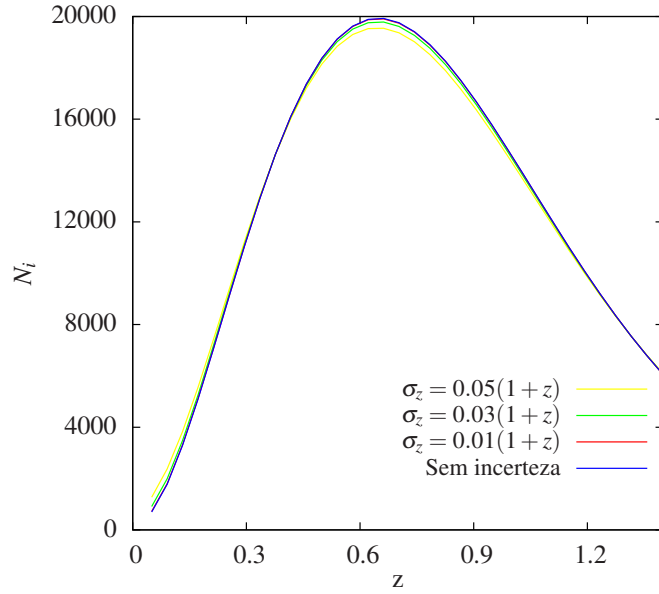


Figura 18: Abundância de aglomerados em *bins* do desvio para o vermelho fotométrico, $\Delta z^p = 0.04$. As curvas amarela, verde e vermelha levam em conta a incerteza em z^p com $\sigma_z = 0.01, 0.03$ e 0.05 , respectivamente. A curva azul é obtida no caso em que conhecemos z (equação 3.1). Utilizamos os mesmos parâmetros cosmológicos da simulação Carmen e $M_{\min}^{\text{obs}} = 5 \times 10^{13} h^{-1} M_{\odot}$.

massa observada no intervalo $M_{\alpha}^{\text{obs}} \leq M^{\text{obs}} \leq M_{\alpha+1}^{\text{obs}}$ é

$$\frac{dN^M}{dV} = \int_{M_{\alpha}^{\text{obs}}}^{M_{\alpha+1}^{\text{obs}}} d \ln M^{\text{obs}} \int_{-\infty}^{\infty} d \ln M \frac{dn(M, z)}{d \ln M} p(M^{\text{obs}} | M). \quad (6.9)$$

No caso em que a variância e o viés independem da medida da massa, $\sigma_{\ln M} = \sigma_{\ln M}(z)$ e $\ln M_{\text{bias}} = \ln M_{\text{bias}}(z)$, a integração na variável $\ln M^{\text{obs}}$ da equação (6.9) é analítica. Usando a equação (6.8) e fazendo a mudança de variável,

$$dx = \frac{1}{\sqrt{2}\sigma_{\ln M}} d \ln M^{\text{obs}}, \quad (6.10)$$

obtemos,

$$\begin{aligned} \frac{dN^M}{dV} &= \int_{-\infty}^{\infty} d \ln M \frac{dn(M, z)}{d \ln M} \frac{\sqrt{2}\sigma_{\ln M}}{\sqrt{2\pi}\sigma_{\ln M}} \int_{x_{\alpha}}^{x_{\alpha+1}} dx e^{-x^2} \\ &= \int_{-\infty}^{\infty} d \ln M \frac{dn(M, z)}{d \ln M} \frac{1}{\sqrt{\pi}} \left[\int_0^{x_{\alpha+1}} dx e^{-x^2} - \int_0^{x_{\alpha}} dx e^{-x^2} \right] \\ &= \int_{-\infty}^{\infty} d \ln M \frac{dn(M, z)}{d \ln M} \frac{1}{2} [\text{erf}(x_{\alpha+1}) - \text{erf}(x_{\alpha})], \end{aligned} \quad (6.11)$$

onde $\text{erf}(x)$ é a função erro cuja definição pode ser vista na equação (2.4). Com isso,

temos que a abundância no k -ésimo *bin* do desvio para o vermelho é

$$N_k = A_{\text{survey}} \int_{z_k - \frac{\Delta z}{2}}^{z_k + \frac{\Delta z}{2}} dz' \frac{dN^M}{dz}, \quad (6.12)$$

onde

$$\frac{dN^M}{dz} = \frac{dV}{dz d\Omega} \int_{-\infty}^{\infty} d \ln M \frac{dn(M, z)}{d \ln M} \frac{1}{2} [\text{erf}(x_{\alpha+1}) - \text{erf}(x_{\alpha})] \quad (6.13)$$

Na figura 19, mostramos a abundância de aglomerados em *bins* no desvio para o vermelho, onde $\Delta z = 0.04$, com e sem incerteza na massa. Consideramos $x_{\alpha} = 5 \times 10^{13} h^{-1} M_{\odot}$ e $x_{\alpha+1} = \infty$ e $\sigma_{\ln M} = 0.1, 0.3$ e 0.5 [109, 43]. Notamos que, quanto maior a largura da distribuição da massa observável $\sigma_{\ln M}$, maior é o número médio de aglomerados nos *bins* de z . Isso ocorre pois a densidade de objetos cresce com a diminuição da massa, como podemos ver na figura 12, conseqüentemente quanto maior $\sigma_{\ln M}$ massas menores serão consideradas na integração em $\ln M$ (equação 6.12). Em outras palavras, há uma migração dos objetos de baixa massa para os bins de massa maior, aumentando artificialmente a abundância.

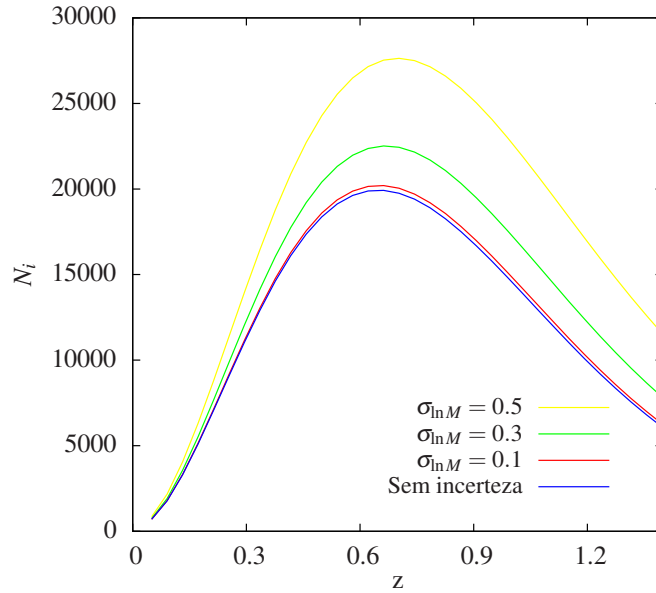


Figura 19: Abundância de aglomerados em *bins* do desvio para o vermelho, onde $\Delta z = 0.04$. As curvas amarela, verde e vermelha foram obtidas a partir da equação (6.12) com $\sigma_{\ln M} = 0.1, 0.3$ e 0.5 , respectivamente. A curva azul corresponde à abundância sem incerteza na massa (equação 3.1). Utilizamos os mesmos parâmetros cosmológicos da simulação Carmen e $M_{\min}^{\text{obs}} = 5 \times 10^{13} h^{-1} M_{\odot}$.

6.3 Inclusão de incertezas na verossimilhança

Como vimos na seção 3.3, a verossimilhança sem binagem no desvio para o vermelho é construída a partir da distribuição dos aglomerados $\frac{dN(z)}{dz}$. Para incluir as incertezas da massa e do desvio para o vermelho fotométrico na análise estatística, nós obtemos, primeiramente, as distribuições que levam em conta essas incertezas. Considerando somente o desvio para o vermelho fotométrico, temos que

$$\mathcal{P}^p(z_i^p) = \frac{1}{N_{\text{total}}^p} \frac{dN^p}{dz^p}, \quad (6.14)$$

onde a normalização N_{total}^p é dada pela equação (6.2) integrada no intervalo $z_{\text{min}}^p \leq z^p \leq z_{\text{max}}^p$.

De maneira análoga, a distribuição que leva em conta a incerteza na massa é definida a partir da equação (6.13) como

$$\mathcal{P}^M(z_i) = \frac{1}{N_{\text{total}}^M} \frac{dN^M}{dz}, \quad (6.15)$$

onde N_{total}^M é obtida integrando a equação (2.34) no intervalo $z_{\text{min}} \leq z \leq z_{\text{max}}$. Finalmente, combinando as duas incertezas, a distribuição da abundância de aglomerados em função do desvio para o vermelho fotométrico é dada por

$$\mathcal{P}^{pM}(z_i^{pM}) = \frac{1}{N_{\text{total}}^{pM}} \frac{dN^{pM}}{dz^{pM}} = \frac{1}{N_{\text{total}}^{pM}} \int_0^\infty dz \frac{dN^M}{dz} p(z^{pM}|z), \quad (6.16)$$

onde N_{total}^{pM} é o número de objetos otidos no intervalo $z_{\text{min}}^{pM} \leq z^{pM} \leq z_{\text{max}}^{pM}$.

Definidas as distribuições dadas pelas equações (6.14) – (6.16), nós construímos a verossimilhança como

$$L(\{\theta_j\}, \{z_i^\beta\}) = P(N_{\text{total}}^{\text{obs}}; N_{\text{total}}^\beta) \prod_{i=1}^{N_{\text{total}}^{\text{obs}}} \mathcal{P}^\beta(z_i^\beta), \quad (6.17)$$

onde o índice β pode ser p , M ou pM .

Nas seções que seguem, nós utilizaremos uma realização obtida levando em conta a incerteza no desvio para o vermelho fotométrico e/ou na massa para estudar os efeitos nas estimativas dos parâmetros cosmológicos em dois casos diferentes, quando essas incertezas são consideradas na análise estatística e, também, quando elas são ignoradas.

6.3.1 Efeito da incerteza no desvio para o vermelho

Neste caso, geramos uma realização $\{z_i^p\}$ seguindo o procedimento descrito na seção 3.2.1, porém utilizando a distribuição $\mathcal{P}^p(z^p)$ (equação 3.6). Como modelo fiducial, utilizamos os mesmos parâmetros cosmológicos da simulação Carmen, escolhemos os parâmetros da distribuição $p(z^p|z)$ como $z^{bias} = 0$ e $\sigma_z = 0.01(1+z)$. Esse último corresponde à incerteza esperada nas medidas dos desvios para o vermelho fotométricos dos aglomerados no DES [20]. O limiar da massa é fixo em $M_{\min} = 5 \times 10^{13} h^{-1} M_{\odot}$, $A_s = 5000 \text{deg}^2$ e $z_{\max} = 1.4$.

Supondo que a equação de estado da energia escura é dada pela equação (2.36), nós calculamos as regiões de confiança de 1, 2 e 3σ com o método de matriz de Fisher no plano (Ω_m, σ_8) , marginalizando nos parâmetros w_0 e w_1 , utilizando a verossimilhança (6.17) com $\beta = p$. Esses contornos são mostrados no primeiro painel da figura 20. Ainda nesse painel, mostramos as regiões de confiança obtidas com a verossimilhança (3.10), ou seja, ajustamos os parâmetros sem levar em conta a incerteza nas medidas de z^p .

No segundo painel da figura 20, mostramos as regiões de confiança de 1, 2 e 3σ no plano (Ω_m, w_0) , obtidas marginalizando os parâmetros σ_8 e w_1 , em relação tanto à verossimilhança (6.17), com $\beta = p$, quanto à equação (3.10). E, no terceiro painel, temos os contornos obtidos com o mesmo procedimento, porém no plano (w_0, w_1) , onde Ω_m e σ_8 são deixados livres.

Notamos nos três planos que ignorar a incerteza no desvio para o vermelho fotométrico quando $\sigma_z = 0.01$ acarreta em obter as estimativas dos parâmetros com um viés de 0.005 em Ω_m , 0.003 em σ_8 , 0.06 em w_0 e 0.2 em w_1 . No entanto, esse deslocamento ainda está dentro da região de 1σ .

Repetimos a análise descrita acima, porém considerando $\sigma_z = 0.03(1+z)$, que seria a pior estimativa para o erro do desvio para o vermelho fotométrico de um aglomerado (já que essa é a incerteza esperada nos z^p da população geral de galáxias). Na figura 21, mostramos os resultados das regiões de confiança de 1, 2 e 3σ . Podemos notar que, para $\sigma_z = 0.03(1+z)$, não só o viés está fora da região de 3σ como as próprias regiões estão disjuntas. Portanto, ignorar a distribuição do desvio para o vermelho fotométrico na análise acarretaria em um resultado significativamente enviesado, em vários σ . Neste caso temos um viés de 0.5 em Ω_m , 0.3 em σ_8 , 0.51 em w_0 e 1.68 em w_1 .

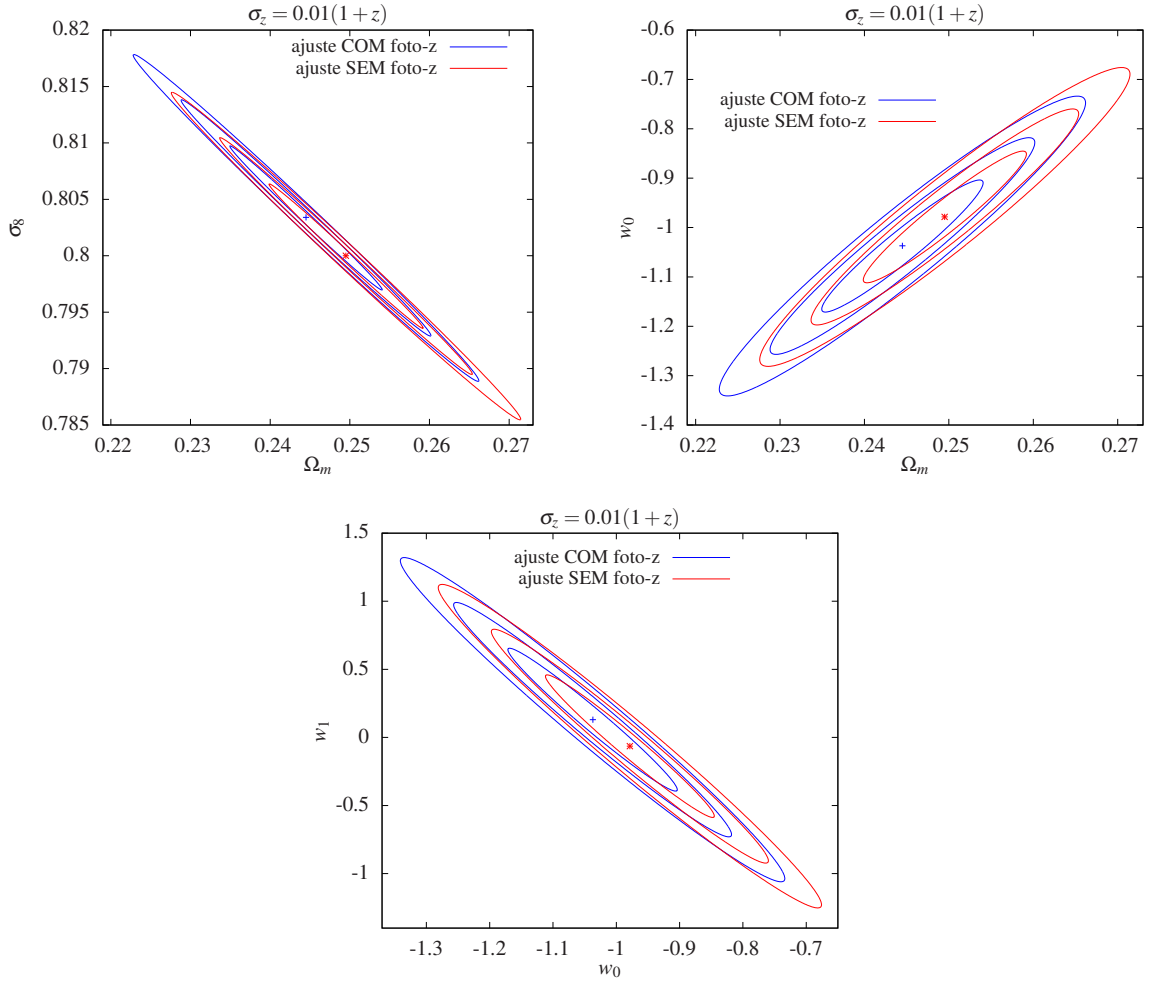


Figura 20: A realização $\{z_i\}$ foi gerada incluindo a distribuição do desvio para o vermelho fotométrico com $\sigma_z = 0.01(1+z)$. As curvas azuis são as matrizes Fisher obtidas levando-se em conta o efeito do desvio para o vermelho fotométrico, enquanto as curvas vermelhas foram obtidas sem considerar a incerteza no conjunto $\{z_i\}$.

6.3.2 Efeito da incerteza na massa nos parâmetros cosmológicos

Seguindo o mesmo procedimento das seção 6.3.1, nós agora vamos estudar o efeito da incerteza na massa. Utilizando os mesmos parâmetros como modelo fiducial e escolhendo $\ln M_{bias} = 0$ e $\sigma_{\ln M} = 0.5$ [107] como os parâmetros da distribuição de probabilidade $p(M^{obs}|M)$, geramos uma realização $\{M_i^{obs}, z_i\}$ (veja seções 3.2.1 e 3.5). Com isso, calculamos as regiões de confiança através do método de matriz de Fisher com a verossimilhança dada por (6.17) e $\beta = M$ e, também com (3.10), onde, neste caso, não consideramos a incerteza na massa.

Na figura 22, mostramos os contornos de 1, 2 e 3σ nos três diferentes planos, como feito na seção 6.3.1. Notamos, no primeiro painel, que σ_8 é o parâmetro mais afetado pela

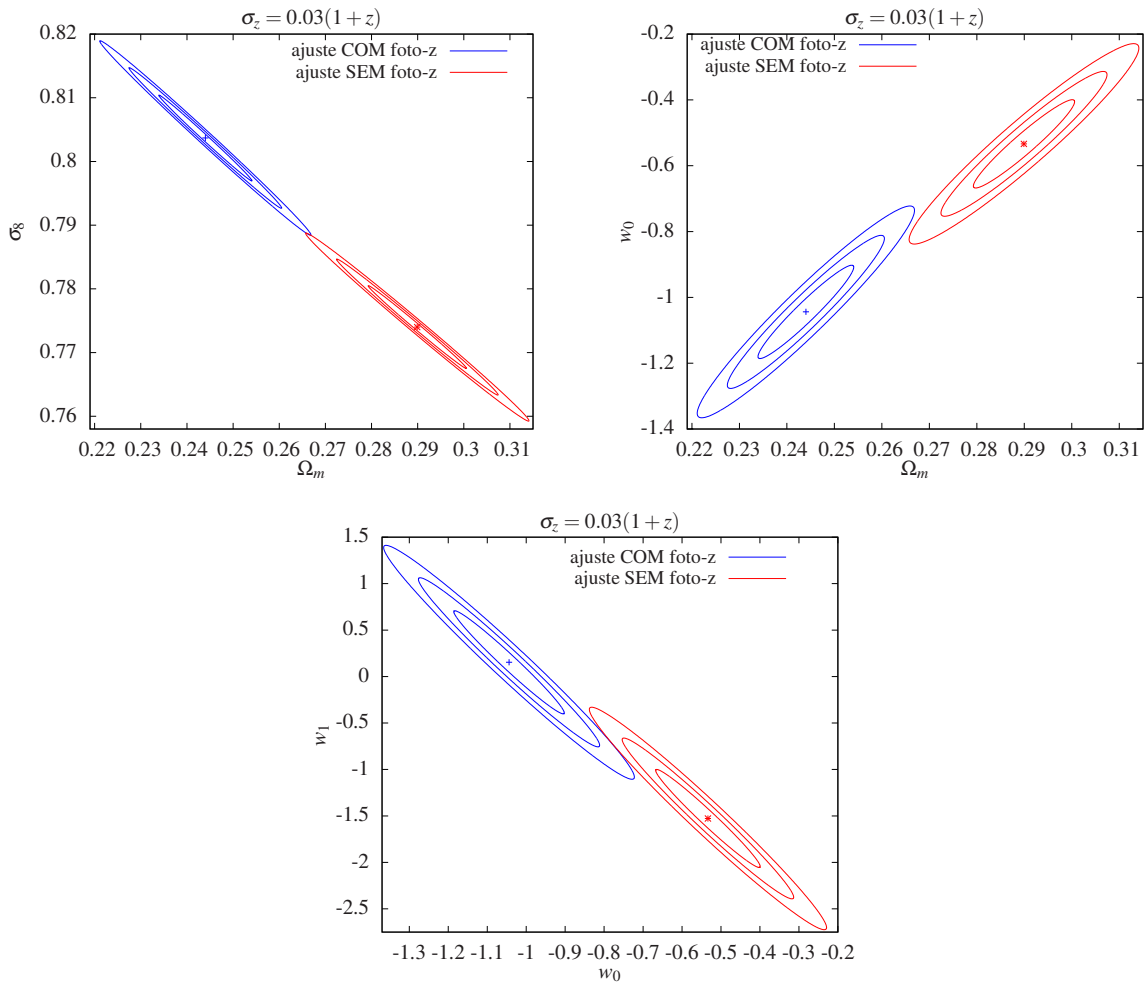


Figura 21: A realização $\{z_i\}$ foi gerada incluindo a distribuição do desvio para o vermelho fotométrico com $\sigma_z = 0.03(1+z)$. As curvas azuis são obtidas levando-se em conta o efeito do desvio para o vermelho fotométrico, enquanto as curvas vermelhas foram obtidas sem considerar a incerteza no conjunto $\{z_i\}$.

incerteza na massa, sendo a diferença entre os dois casos maior que 3σ . Os outros painéis mostram que o deslocamento ocorre em $\simeq 3\sigma$. O viés é de 0.005 em Ω_m , 0.07 em σ_8 , 0.03 em w_0 e 0.32 em w_1 . Então a incerteza na massa acarreta em uma diferença grande e é fundamental modelar corretamente a relação massa–observável.

6.3.3 Efeito da combinação de erros nos parâmetros cosmológicos

Para incluir os efeitos tanto da incerteza na massa quanto no desvio para o vermelho fotométrico, geramos uma realização $\{M_i^{obs}, z_i^p\}$ para o mesmo modelo fiducial das seções anteriores e com $\sigma_z = 0.01(1+z)$ e $\sigma_{\ln M} = 0.5$.

No painel esquerdo da figura 23, apresentamos as regiões de confiança de 1, 2 e 3σ

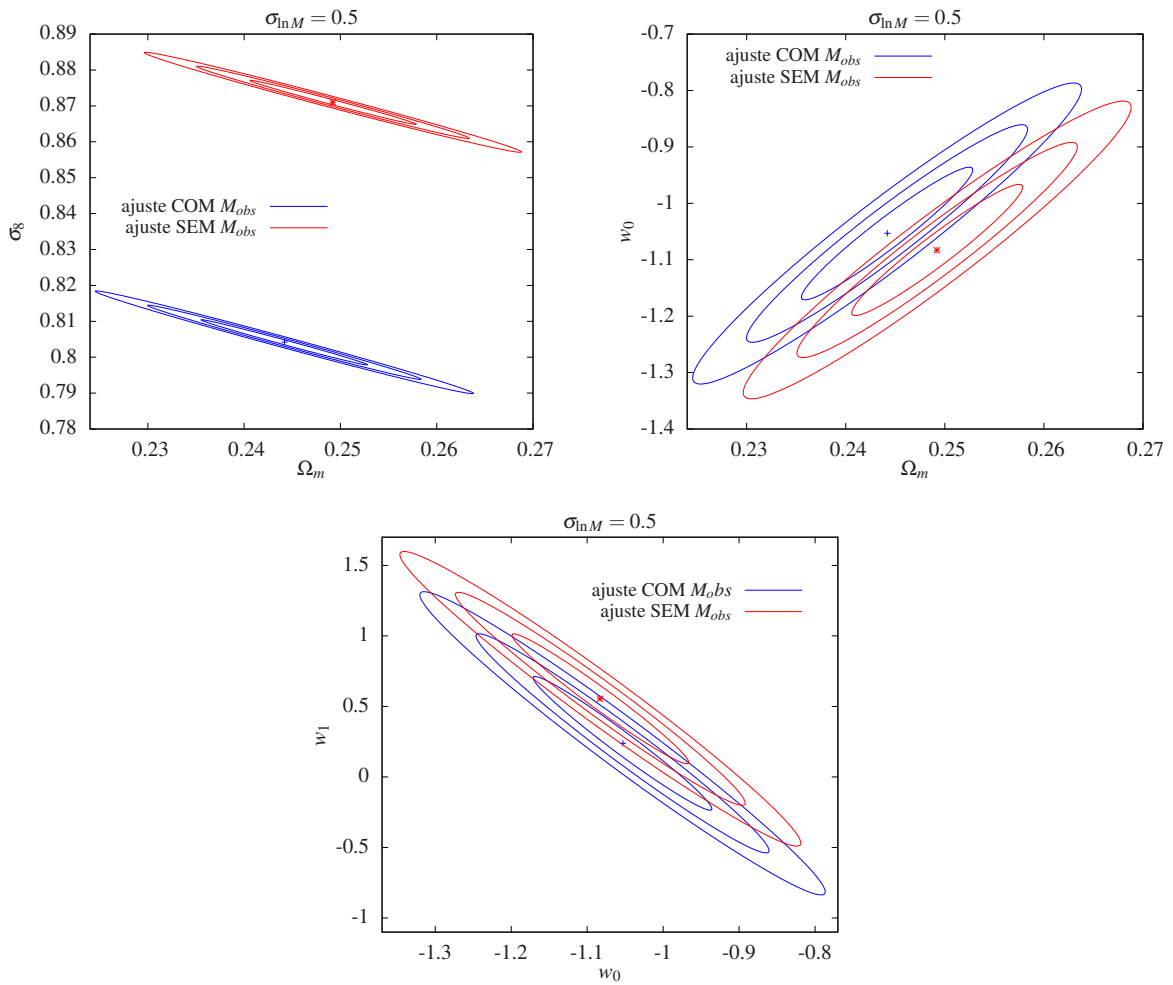


Figura 22: A realização $\{z_i\}$ foi gerada incluindo a distribuição $p(M^{obs}|M)$ com $\sigma_{\ln M} = 0.5$. As curvas azuis são obtidas levando-se em conta o efeito da incerteza na medida da massa observada, enquanto as curvas vermelhas foram obtidas sem considerar a incerteza em M_{\min}^{obs} .

obtidas com matriz de Fisher no plano (w_0, w_1) , marginalizando sobre Ω_m e σ_8 , e com as verossimilhanças (6.17), onde $\beta = pM$, e (3.10). Comparando esta figura com os terceiros painéis das figuras 20 e 22, vemos que o efeito causado pela incerteza da massa domina.

Repetindo esse procedimento, mas agora com $\sigma_z = 0.03(1+z)$, verificamos que, ainda assim, o efeito da massa domina, como mostrado no painel direito da figura 23. Surpreendentemente, a variação das curvas é mínima, ao contrário do que pode ser visto na figura 21. Não compreendemos porque a incerteza no desvio para o vermelho fotométrico parece ter um papel tão menor neste caso. Investigaremos essa questão em um futuro próximo.

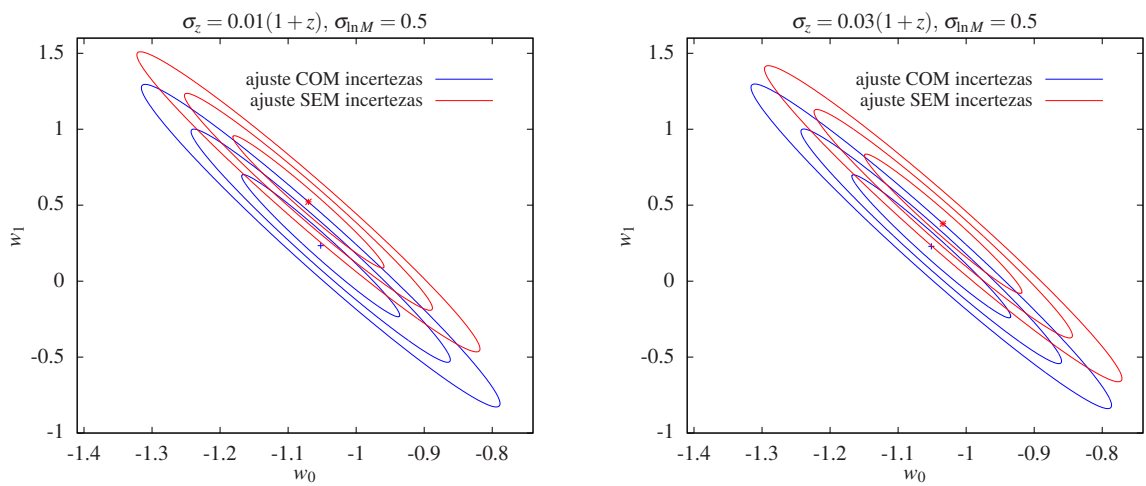


Figura 23: A realização $\{z_i\}$ foi gerada incluindo as distribuições $p(M^{obs}|M)$, com $\sigma_{\ln M} = 0.5$, e $p(z^p|z)$, com $\sigma_z = 0.01(1+z)$ (painel esquerdo) e $\sigma_z = 0.03(1+z)$ (painel direito). As curvas azuis são as regiões de confiança de 1, 2 e 3σ obtidas levando-se em conta o efeito das incertezas nas medidas da massa observada e do desvio para o vermelho fotométrico, enquanto as curvas vermelhas foram obtidas sem considerar essas incertezas.

7 *Conclusões e Perspectivas Futuras*

Neste trabalho, introduzimos uma nova verossimilhança, o método ULC (equação 3.10), para a abundância de aglomerados. Ela difere dos métodos usuais de Poisson e χ^2 por não requerer a utilização de *bins* em z e por permitir incluir a incerteza em z para cada aglomerado de forma individual. Propusemos também a verossimilhança baseada em $\frac{d^2N}{dzd\ln M}$ (equação (3.15)) que, além de possuir as mesmas características de ULC, explora a informação sobre a massa de cada objeto e permite tratar individualmente a incerteza na relação massa-observável.

Utilizamos o método de Monte Carlo para obter o viés dos estimadores do parâmetro da equação de estado da energia escura w_0 e a sua variância e verificamos que a verossimilhança ULC fornece essencialmente os mesmos resultados que a verossimilhança de Poisson (e claramente superior ao método χ^2). Quando estimamos simultaneamente diferentes parâmetros cosmológicos, a diferença entre os métodos ULC e Poisson torna-se maior. No exemplo tratado na seção 3.3 a área das regiões de confiança é cerca de 7% menor no método ULC que no método de Poisson. Ainda no capítulo 3, mostramos um resultado preliminar de alguns estimadores obtidos com a verossimilhança dada pela equação (3.13), encontrando um resultado promissor que será explorado em um trabalho futuro.

No capítulo 4 aplicamos o método ULC para obter previsões para os limites que podem ser impostos nos parâmetros cosmológicos a partir de várias amostras de aglomerados no SDSS. Verificamos que a faixa 82 fornece uma restrição comparável à obtida com o DR7. Vimos também que há um ganho considerável na restrição dada pelo DR8 em relação ao DR7 devido ao aumento na profundidade de $z_{\max} = 0.3$ para $z_{\max} = 0.45$. É possível obter uma restrição ainda mais forte quando combinamos o DR8 com a faixa 82.

Também aplicamos a metodologia ULC para validar os catálogos de halos do DES, como mostramos no capítulo 5. Através da comparação entre a função de massa, $\frac{dn}{d\log m}$,

e o catálogo de halos em diferentes intervalos no desvio para o vermelho, verificamos que o número de objetos identificados no catálogo era sistematicamente menor do que o esperado para qualquer valor de M e $z \gtrsim 0.3$. Produzido um novo catálogo, refizemos o mesmo estudo e ainda verificamos os valores de M_{\min} a partir dos quais existe uma boa concordância entre catálogo e modelo. Essa completeza varia com o desvio para o vermelho, sendo que para considerar todo o intervalo $0 \leq z \leq 1.35$, para restringir os parâmetros cosmológicos, a M_{\min} adequada deve ser $\gtrsim 10^{14} h^{-1} M_{\odot}$.

Finalmente, no capítulo 6 modelamos as incertezas no desvio para o vermelho fotométrico e na relação massa–observável supondo que essas seguem uma distribuição Gaussiana em z^p e em $\ln M^{obs}$, respectivamente. Utilizando o método ULC, estudamos o efeito na restrição dos parâmetros cosmológicos ao ignorar essas incertezas na análise. Notamos que para a incerteza esperada no desvio para o vermelho fotométrico dos aglomerados no DES, $\sigma_z = 0.01(1 + z)$, não levar em conta esse efeito implica em um viés pequeno (dentro de 1σ) nos parâmetros. Já no caso $\sigma_z = 0.03(1 + z)$, o viés desloca o melhor ajuste de forma que as regiões de confiança ficam disjuntas. Já o efeito da incerteza na massa introduz um viés que desloca o melhor ajuste em torno de 3σ . Como é conhecido na literatura, a inclusão das incertezas dos observáveis z^p e M^{obs} é necessária para se restringir com acurácia modelos cosmológicos nos levantamentos de grande área voltados para a cosmologia de precisão. Por isso, é necessário aprofundar o estudo que iniciamos neste capítulo, em particular utilizando o método Monte Carlo para adicionar as distribuições de z^p e M^{obs} , permitindo comparar os limites sem essas distribuições, levando-as em conta na modelagem e ignorando-as quando estão presentes.

Os estudos realizados neste trabalho abrem as perspectivas de uma série de análises complementares visando o desenvolvimento de métodos eficientes para determinar os parâmetros cosmológicos (levando em conta aspectos observacionais) e a sua aplicação a dados reais.

Na seção 3.5, apresentamos uma verossimilhança, construída a partir da função $\frac{d^2 N}{dz d \ln M}$, que permite considerar a massa de cada aglomerado individualmente. Como esperado, o fato de utilizar mais informações faz com que esse método seja mais eficiente na restrição dos parâmetros cosmológicos do que o método ULC. Portanto, aprofundaremos o estudo dessa verossimilhança de forma análoga à desenvolvida para o método ULC (seção 3.3) e, em seguida, verificaremos se o método se mantém mais eficiente ao incluir as modelagens

das incertezas do desvio para o vermelho fotométrico e da relação massa-observável.

Outro aspecto fundamental a ser incluído na modelagem é a função de seleção dos catálogos, o que permitirá aplicar o método a catálogos reais em que essa função é bem estimada.

A partir das duas metodologias desenvolvidas neste trabalho iremos:

- validar os catálogos de aglomerados de galáxias produzidos pela colaboração do DES, assim como validamos o catálogo de halos (capítulo 5);
- fazer previsões que poderão otimizar as estratégias de observação do DES, a partir dos resultados do item acima. Estudar as incertezas nos parâmetros da energia escura e verificar o poder discriminatório de modelos, inclusive gravidade modificada.
- combinar com as verossimilhanças de outros observáveis;
- utilizar os dados reais do SDSS, utilizando a função seleção adequada a cada catálogo, para restringir os parâmetros cosmológicos;

Neste trabalho, tratamos apenas da abundância de aglomerados. No entanto, é conhecido que a “covariância da amostra” (*sample covariance*) precisa ser considerada em conjunto com a estatística da contagem desses objetos para restringirmos modelos cosmológicos [98, 99, 110]. Estudaremos essa questão para incluí-la na nossa análise.

Pretendemos também calcular a função de correlação de dois pontos teórica e combiná-la com a obtida a partir dos catálogos para obter o viés entre a aglomeração de matéria escura e a dos aglomerados [52], permitindo assim ajustar parâmetros da relação massa-observável em conjunto com os parâmetros cosmológicos (auto-calibração).

APÊNDICE A – Métodos Estatísticos Utilizados

Neste apêndice nós apresentaremos algumas ferramentas de estatística que foram utilizadas ao longo deste trabalho para fazer análises como, por exemplo, restringir os parâmetros cosmológicos a partir da abundância de aglomerados de galáxias.

Na seção A.1, nós apresentamos o método da Máxima Verossimilhança a partir do qual obtemos os estimadores dos parâmetros cosmológicos. Já na seção A.2, tratamos do Teste da Razão de Verossimilhança que possibilita avaliar o nível de concordância entre uma modelagem teórica e dados observacionais. Nas seções A.3 e A.4, nós apresentamos os métodos Perfil da Verossimilhança e Matriz de Fisher que são usados para calcular as regiões de confiança.

A.1 Método da Máxima Verossimilhança

Seja x uma variável aleatória com função densidade de probabilidade, à qual nos referiremos por *pdf*, $f(x, \vec{\theta})$ dependente de um vetor de parâmetros $\vec{\theta}$. Dada a amostra de dados independentes $\{x_i\}$, temos que a densidade de probabilidade conjunta de se obter $\{x_i\}$ dado $\vec{\theta}$ é

$$L(\vec{\theta}; x_i) = \prod_{i=1}^N f(x_i; \vec{\theta}), \quad (\text{A.1})$$

onde $L(\vec{\theta}; x_i)$, vista como uma função dos parâmetros não conhecidos $\vec{\theta}$ dadas as medidas $\{x_i\}$, é chamada de verossimilhança.

O método da Máxima Verossimilhança é uma técnica para estimar os valores dos parâmetros $\vec{\theta}$ dada o conjunto de dados finito x_i . Os estimadores $\hat{\theta}_j$ são aqueles que maximizam o valor da verossimilhança,

$$\hat{\theta}(x_i) = \arg \max[L(\vec{\theta}; x_i)]. \quad (\text{A.2})$$

Supondo que $L(\vec{\theta}; x_i)$ é uma função diferenciável dos parâmetros $\{\theta_j\}$ e que o seu máximo não esteja nos limites dos domínios desses parâmetros, temos que $\hat{\theta}_j$'s são dados pelas soluções das equações

$$\frac{\partial L}{\partial \theta_j} = 0, \quad j = 1, \dots, m. \quad (\text{A.3})$$

Um estimador é dito *sem viés* se o seu valor esperado for igual ao valor verdadeiro θ^0 ($\langle \hat{\theta}_j \rangle = \theta_j^0$) e, ele é *eficiente* se a sua variância é a menor possível¹. Uma vantagem de se usar os estimadores $\hat{\theta}_j$'s é dada pelo teorema que diz que se existe um estimador eficiente e sem viés para um dado problema, então esse estimador coincidirá com o estimador de máxima verossimilhança (veja a referência [111]). Outra vantagem é que, sob condições muito gerais, ele é consistente, isto é, ele tende ao seu valor verdadeiro quando o número de dados é suficientemente grande.

No caso em que a pdf da variável aleatória y é a distribuição Gaussiana e dado o conjunto de medidas $\{y_i^{\text{obs}}\}$, temos que

$$L(\vec{\theta}) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(y_i^{\text{obs}} - y_i(\vec{\theta}))^2}{2\sigma_i^2}\right). \quad (\text{A.4})$$

Ao invés de utilizar $L(\vec{\theta}; x_i)$, muitas vezes é mais conveniente utilizar a função log-verossimilhança. Como o logaritmo é uma função monotonicamente crescente, os valores dos parâmetros que maximizam a verossimilhança também maximizam $\log L(\vec{\theta}; x_i)$. Dessa forma, a equação (A.4) é dada por

$$\ln L(\vec{\theta}) = \sum_{i=1}^N \frac{-(y_i^{\text{obs}} - y_i(\vec{\theta}))^2}{2\sigma_i^2} - \sum_{i=1}^N \ln(\sqrt{2\pi}\sigma_i). \quad (\text{A.5})$$

Considerando que σ_i não depende dos parâmetros $\{\theta_j\}$, a equação (A.5) é maximizada encontrando os valores de $\{\theta_j\}$ que minimizam a função

$$-2 \ln L(\vec{\theta}) = \chi^2 = \sum_{i=1}^N \frac{(y_i^{\text{obs}} - y_i(\vec{\theta}))^2}{\sigma_i^2}. \quad (\text{A.6})$$

Neste caso, o método da máxima verossimilhança coincide com o *método dos mínimos quadrados*.

¹Essa característica será discutida na seção A.4.

A.2 Teste da Razão de Verossimilhança

Considere os modelos H_0 e H_1 no qual o último está *aninhado* em relação ao primeiro, isto é, H_1 é um subconjunto de H_0 ou pode ser considerado como um caso especial de H_0 . Portanto, o número de parâmetros θ' de H_1 , m' , tem que ser menor que o número de parâmetros θ de H_0 , m . Dado os estimadores da máxima verossimilhança $\hat{\theta}'$ e $\hat{\theta}$ e um conjunto de dados $\{y_i\}$, o modelo H_1 é testado frente a H_0 , calculando-se a razão das verossimilhanças,

$$\Lambda = \frac{L_1(\hat{\theta}', \{y_i\})}{L_0(\hat{\theta}, \{y_i\})}, \quad (\text{A.7})$$

onde $0 < \Lambda \leq 1$. Como a densidade de probabilidade $L_0(\hat{\theta}, \{y_i\})$ é máxima, se o valor dessa razão for próximo de 0, isso indica que H_1 não é aceito comparado a H_0 . Caso Λ seja próximo de 1, significa que o modelo H_1 é quase tão provável quanto H_0 [112, 113]. Esse é o chamado *teste da razão de verossimilhança* (*likelihood ratio test*).

Para avaliar quão bem ajustado aos dados observacionais é um dado modelo, neste trabalho, nós usamos o *teste da razão de verossimilhança* em conjunto com um teorema que diz que, sob certas condições regulares², a variável aleatória $-2 \ln(L_1(\theta')/L_0(\theta))$ tem uma distribuição que se aproxima da distribuição χ^2 quando $N \rightarrow \infty$, com seus graus de liberdade iguais à diferença no número de parâmetros dos dois modelos [112, 113]. Vejamos o exemplo abaixo.

Considere que foram medidos n aglomerados de galáxias em um intervalo de desvio para o vermelho. Dividindo esse intervalo em i bins, nós diremos que n_i é uma variável de Poisson com média λ_i , ou seja,

$$P(n_i; \lambda_i) = \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}. \quad (\text{A.8})$$

Com isso, para quantificar se um certo modelo para $\lambda_i(\theta)$ está de acordo com os dados n_i , nós calculamos

$$\chi_\nu^2 = -2 \ln \left(\frac{L_1(\{\lambda_i(\theta)\}, \{n_i\})}{L_0(\{n_i\}, \{n_i\})} \right), \quad (\text{A.9})$$

onde $L_0(\{n_i\}, \{n_i\})$ é a verossimilhança construída considerando que os valores das médias λ_i são dadas pelas medidas n_i , em $L_1(\{\lambda_i\}, \{n_i\})$ consideramos $\lambda_i = \lambda_i(\theta)$ com valores de θ fixos e o número de graus de liberdade ν igual ao número de bins, $\nu = N$. Usando a

²A diferenciabilidade da função $L(\vec{\theta})$ e o fato de não possuir o seu máximo nos limites do domínio dos parâmetros $\{\theta_j\}$ são duas condições das chamadas *condições regulares*.

equação (A.8), obtemos

$$\chi_\nu^2 = 2 \sum_{i=1}^N \left[\lambda_i - n_i - n_i \ln \left(\frac{\lambda_i}{n_i} \right) \right]. \quad (\text{A.10})$$

Vale ressaltar que a hipótese de n_i ser variável de Poisson não está sendo testada.

Com o valor χ_ν^2 obtido com a Eq. (A.9), nós podemos calcular a probabilidade de obter os dados observacionais dado o modelo. Para isso calculamos o nível de significância (*P-value*) [114],

$$\int_{\chi_\nu^2}^{\infty} dx \frac{1}{2^{\nu/2} \Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2} = p, \quad (\text{A.11})$$

onde o integrando é a distribuição $\chi^2(x, \nu)$.

A.3 Perfil da Verossimilhança

Dado um modelo para uma quantidade $y(\{\theta_j\})$, onde $j = 1, \dots, m$, e um conjunto de dados $\{y_i\}$, onde $i = 1, \dots, N$, temos que os valores dos parâmetros θ_j , nos quais obtemos a maior probabilidade de, dado o modelo, obter os dados $\{y_i\}$, são aqueles que maximizam a verossimilhança $L_0(\{\hat{\theta}_j\}, \{y_i\})$, onde $\{\hat{\theta}_j\}$ é o ponto de melhor ajuste.

Para obter a região no espaço paramétrico onde a probabilidade de obter os dados $\{y_i\}$ é maior ou igual a $1 - p$, ou seja, para obter a região de confiança nós utilizamos o método *Perfil da Verossimilhança* (*Profile Likelihood*). Esse método, muito utilizado em física de partículas [115], consiste em fazer o teste da razão de verossimilhança no contexto descrito abaixo.

No caso em que queremos obter a região de confiança no subespaço (θ_1, θ_2) , por exemplo, a hipótese H_1 consistirá em manter fixos esses parâmetros. Com isso, temos que a função $-2 \ln \Lambda$ (equação A.7) seguirá uma distribuição χ_ν^2 de 2 graus de liberdade,

$$\chi_{\nu=2}^2 = -2 \ln \left(\frac{L_1(\theta_1, \theta_2, \{\hat{\theta}'_i\}, \{y_i\})}{L_0(\{\hat{\theta}_j\}, \{y_i\})} \right), \quad (\text{A.12})$$

onde $i = 3, \dots, m$ e $\{\hat{\theta}'_i\}$ será o melhor ajuste para o ponto (θ_1, θ_2) . Dado o nível de confiança $1 - p$ da região, calculamos o valor de $\chi_{\nu=2}^2$ a partir da equação (A.11) e, dessa forma, podemos obter os valores de θ_1 e θ_2 que satisfazem a equação (A.12). Esses pontos definem o contorno da região de confiança. Usualmente, são calculadas as regiões em que $1 - p = 68.27\%$, 95.45% e 99.73% (esses valores correspondem às regiões de 1 , 2 e 3σ para a distribuição Gaussiana).

Como não é feita nenhuma suposição sobre os parâmetros θ'_i , dizemos que eles são *livres*. O procedimento descrito acima, chamaremos de *perfilização*, pode ser relacionado (mas vale ressaltar que não é equivalente) ao processo de marginalização na análise Bayesiana.

A região de confiança também pode ser obtida mantendo fixos alguns parâmetros θ_i , onde $i \neq 1, 2$. Essa análise é equivalente a supor que esses parâmetros são conhecidos com nenhuma incerteza.

A.4 Matriz de Fisher

O gasto computacional para calcular as regiões de confiança com o método de perfil da verossimilhança pode ser muito alto, principalmente quando estima-se muitos parâmetros simultaneamente. Um método muito utilizado na literatura, por ser de baixo custo computacional, é o da *Matriz de Fisher*. No entanto, a matriz de Fisher fornece apenas uma aproximação da matriz covariância dos parâmetros $\{\theta_j\}$, como veremos nesta seção, e por isso, o seu uso é limitado.

A matriz de Fisher é definida como o valor esperado da matriz Hessiana,

$$F_{ij} = \left\langle \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_i \partial \theta_j} \Big|_{\vec{\theta}^0} \right\rangle, \quad (\text{A.13})$$

onde $\vec{\theta}^0$ são os valores verdadeiros dos estimadores, e está relacionada com a variância dos estimadores $\hat{\theta}_i$ como dado pelos teoremas (veja a referência [116]) que dizem que

$$\left\langle \frac{\partial \ln(L(\vec{\theta}))}{\partial \theta_i} \Big|_{\vec{\theta}^0} \right\rangle = 0, \quad (\text{A.14})$$

$$\left\langle \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_i \partial \theta_j} \Big|_{\vec{\theta}^0} \right\rangle = -C \left(\frac{\partial \ln(L(\vec{\theta}))}{\partial \theta_i} \Big|_{\vec{\theta}^0}, \frac{\partial \ln(L(\vec{\theta}))}{\partial \theta_j} \Big|_{\vec{\theta}^0} \right), \quad (\text{A.15})$$

e que, se $\hat{\theta}$ não tiver viés, então

$$C(\hat{\theta}_i, \hat{\theta}_j) \geq \left[- \left\langle \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_i \partial \theta_j} \Big|_{\vec{\theta}^0} \right\rangle \right]^{-1}, \quad (\text{A.16})$$

onde $C(\hat{\theta}_i, \hat{\theta}_j)$ são os elementos da matriz covariância, no sentido que a diferença dessas duas matrizes é positiva semi-definida (o lado direito da inequação é chamado de *limite inferior de Cramer-Rao*). Caso o estimador $\hat{\theta}$ seja eficiente, a inequação (A.16) reduz-se

à igualdade.

Na prática, a equação (A.13) não é calculada, já que o cálculo do valor esperado corresponde a uma integração (a não ser que haja solução analítica). Dessa forma, supõe-se que o conjunto de dados é suficientemente grande e calcula-se a derivada segunda de $\ln L(\vec{\theta})$ no melhor ajuste $\hat{\theta}$,

$$C(\hat{\theta}_i, \hat{\theta}_j) = - \left(\frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_i \partial \theta_j} \Big|_{\hat{\theta}} \right)^{-1}. \quad (\text{A.17})$$

Essa é uma aplicação de um teorema que diz que, no limite assintótico, o inverso da equação (A.17) é um estimador da matriz Fisher [116],

$$\lim_{n \rightarrow \infty} \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_i \partial \theta_j} \Big|_{\hat{\theta}} = F_{ij}. \quad (\text{A.18})$$

Portanto, para obter a região de confiança³ dos parâmetros θ_1 e θ_2 , por exemplo, mantendo todos os outros fixos, basta calcular a equação paramétrica da elipse (veja a referência [102])

$$\theta_1(t) = \hat{\theta}_1 + q[a \cos(\alpha) \cos(t) - b \sin(\alpha) \sin(t)] \quad (\text{A.19})$$

e

$$\theta_2(t) = \hat{\theta}_2 + q[a \sin(\alpha) \cos(t) + b \cos(\alpha) \sin(t)], \quad (\text{A.20})$$

onde,

$$a = \frac{C_{11} + C_{22}}{2} + \sqrt{\frac{(C_{11} - C_{22})^2}{4} + C_{12}^2}, \quad (\text{A.21})$$

$$b = \frac{C_{11} + C_{22}}{2} - \sqrt{\frac{(C_{11} - C_{22})^2}{4} + C_{12}^2}, \quad (\text{A.22})$$

$$\tan(2\alpha) = \frac{2C_{12}^2}{C_{11} - C_{22}}, \quad (\text{A.23})$$

e

$$C(\hat{\theta}_i, \hat{\theta}_j) \equiv C_{ij} \doteq - \left(\begin{array}{cc} \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_1^2} \Big|_{\hat{\theta}_1, \hat{\theta}_2} & \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_1 \partial \theta_2} \Big|_{\hat{\theta}_1, \hat{\theta}_2} \\ \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_1 \partial \theta_2} \Big|_{\hat{\theta}_1, \hat{\theta}_2} & \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_2^2} \Big|_{\hat{\theta}_1, \hat{\theta}_2} \end{array} \right)^{-1}.$$

O parâmetro q indica o nível de confiança da região. No caso de 1σ , temos que $q = 1.52$ e para 2 e 3σ temos $q = 2.48$ e $q = 3.44$, respectivamente.

No caso em que os parâmetros θ_i , onde $i \neq 1, 2$, estão livres, a região de confiança é

³A região de confiança obtida com matriz de Fisher se aproxima assintoticamente da região de confiança do perfil da verossimilhança, da mesma forma que a distribuição dos estimadores é assintoticamente Gaussiana.

obtida a partir da matriz inversa

$$C(\hat{\theta}_i, \hat{\theta}_j) \equiv C_{ij} \doteq - \left(\begin{array}{cccc} \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_1^2} \Big|_{\hat{\theta}} & \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_1 \partial \theta_2} \Big|_{\hat{\theta}} & \cdots & \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_1 \partial \theta_m} \Big|_{\hat{\theta}} \\ \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_1 \partial \theta_2} \Big|_{\hat{\theta}} & \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_2^2} \Big|_{\hat{\theta}} & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_1 \partial \theta_m} \Big|_{\hat{\theta}} & \cdots & \cdots & \frac{\partial^2 \ln(L(\vec{\theta}))}{\partial \theta_m^2} \Big|_{\hat{\theta}} \end{array} \right)^{-1},$$

onde os coeficientes das equações (A.19) e (A.20) são calculados a partir da submatriz 2×2 da matriz C dada pelos elementos C_{11} , C_{12} e C_{22} . Com esse procedimento, os parâmetros θ_i são marginalizados.

APÊNDICE B – Otimização, Controle e Validação dos Códigos

As funções necessárias para o cálculo da abundância de aglomerados assim como para a análise estatística foram implementadas na linguagem C e constituem as bibliotecas *CosmoClusterLib* e *CosmoLib*. Essas bibliotecas usam o controle de versão de software *git* [32] e o Doxygen [117] é usado para documentá-las. Os códigos dessas bibliotecas são organizados em estruturas independentes de forma a possibilitar a troca dos componentes de forma simples. Esses componentes referem-se desde modelos e parâmetros cosmológicos aos algoritmos de análise estatística. Foi desenvolvido um sistema de construção (*build system*) para cada biblioteca e, com isso, elas podem ser compiladas em qualquer plataforma. As bibliotecas estão disponíveis para toda a colaboração DES e, futuramente, estarão abertas a toda comunidade científica.

Neste apêndice, nós descrevemos alguns detalhes da implementação numérica de integrais como, por exemplo, da equação (2.14). Restringimos essa descrição aos processos de otimização do tempo gasto no cálculo dessas integrais dada uma precisão escolhida. A motivação para desenvolvermos códigos menos dispendiosos é devida ao fato de que ao fazermos uma análise estatística para restringir parâmetros cosmológicos, por exemplo, aquelas integrais têm que ser calculadas inúmeras vezes. Antes de implementar a otimização, o tempo gasto para calcular o melhor ajuste do parâmetro da equação de estado da energia escura w mantendo os outros parâmetros cosmológicos fixos, por exemplo, levava ~ 7 minutos. Com a otimização, esse tempo caiu para ~ 3 segundos.

Na seção B.1, apresentamos brevemente um método de interpolação, a interpolação por spline, que utilizamos diversas vezes em nossos códigos. Em particular, o uso desse método no cálculo da variância e da contagem de aglomerados em intervalos no desvio para o vermelho são discutidos nas seções B.2 e B.3, respectivamente. Esses dois casos envolvem somente a interpolação das funções de uma variável. Na seção B.4, estendemos esse procedimento para construir a interpolação por spline de funções de duas variáveis.

Finalmente, na seção B.5 apresentamos alguns testes de validação

B.1 Interpolação por Spline

Spline é uma função polinomial definida em trechos (*piecewise*) adjacentes. Em um intervalo $[x_0, x_n]$, n trechos são determinados por $n + 1$ nós tais que

$$x_0 < x_1 < \dots < x_n.$$

Interpolando os nós $(x_i, f(x_i))$ e $(x_j, f(x_j))$ obtemos o polinômio $P_i(x)$, onde $x \in [x_i, x_j]$. Nos restringiremos ao uso do algoritmo de spline cúbica e, com isso, teremos ao todo n polinômios de terceiro grau,

$$P(x) = \begin{cases} P_0(x) = a_0(x - x_0)^3 + b_0(x - x_0)^2 + c_0(x - x_0) + d_0 & x \in [x_0, x_1] \\ P_1(x) = a_1(x - x_1)^3 + b_1(x - x_1)^2 + c_1(x - x_1) + d_1 & x \in [x_1, x_2] \\ \vdots & \vdots \\ P_{n-1}(x) = a_{n-1}(x - x_{n-1})^3 + b_{n-1}(x - x_{n-1})^2 + c_{n-1}(x - x_{n-1}) + d_{n-1} & x \in [x_{n-1}, x_n]. \end{cases}$$

Como cada polinômio possui quatro coeficientes, precisamos de $4n$ condições independentes para determiná-los. Primeiramente, temos que

$$P_i(x_i) = f(x_i) \quad \text{e} \quad P_{i+1}(x_{i+1}) = P_i(x_{i+1}), \quad (\text{B.1})$$

que são as condições de interpolação e as que garantem a continuidade da função, respectivamente. Outras $2n - 2$ condições surgem com o requerimento de que a primeira e a segunda derivadas também sejam contínuas,

$$P'_{i+1}(x_{i+1}) = P'_i(x_{i+1}), \quad P''_{i+1}(x_{i+1}) = P''_i(x_{i+1}) \quad (\text{B.2})$$

e, desta forma, evitamos que a função tenha picos ou troque abruptamente de curvatura nos nós. As Eqs. (B.1) e (B.2) somam $4n - 2$ condições. Usando a spline cúbica natural [118], que é definida com

$$P''_0(x_0) = P''_{n-1}(x_{n-1}) = 0, \quad (\text{B.3})$$

determinamos as duas condições de contorno que faltavam. Dessa forma, obtemos uma aproximação da função $f(x)$ via método de interpolação por spline cúbica, ou seja, $P(x, \{f(x_i)\}) \approx f(x)$.

B.2 Cálculo Numérico da Variância

Como vimos na seção 2.2, a variância do campo de densidade linear filtrada em uma escala R no desvio para o vermelho z é dada por

$$\sigma^2(R, z) = A \left(\frac{D(z)}{D(0)} \right)^2 \int_0^\infty dk \frac{k^{n+2} T(k)^2}{2\pi^2} W^2(k, R), \quad (\text{B.4})$$

quando usamos o espectro de potências primordial de Harrison-Zeldovich. O primeiro passo realizado para otimizar o cálculo da integral da equação (B.4) foi fazer uma spline cúbica do termo $f(k) = (2\pi^2)^{-1} k^{n+2} T(k)^2$. Para tanto, era necessário ter um limite superior de k finito. Esse limite foi obtido calculando a integral da equação (B.4) com regras de integração de Gauss-Legendre implementadas na biblioteca GSL [119] para as funções janela Gaussiana e cartola para diferentes conjuntos de valores dos parâmetros cosmológicos e de R . Em todos os casos, a convergência da integral ocorreu para $k < 1000$ o que nos levou a escolher o intervalo $[0, 1000]$ para calcularmos a spline $P(k)$. Usando o critério $\frac{|P(k)-f(k)|}{f(k)} \leq 10^{-5}$, nós determinamos o número de nós da spline, $n = 230$, e o espaçamento entre eles, conforme descrito na tabela 5. Na figura 24, nós mostramos a diferença fracionária entre $f(k)$ e a spline $P(k)$ que calculamos para obter

$$\int_0^\infty dk f(k) W^2(k, R) \approx \int_0^{1000} dk P(k) W^2(k, R), \quad (\text{B.5})$$

onde,

$$\int_0^{1000} dk P(k) W^2(k, R) = \int_0^{10^{-20}} dk P_0(k) W^2(k, R) + \dots + \int_{990}^{1000} dk P_{229}(k) W^2(k, R). \quad (\text{B.6})$$

Intervalo em k	Número de nós
$0 \leq k \leq 10^{-5}$	17
$2 \times 10^{-5} \leq k \leq 4 \times 10^{-2}$	66
$5 \times 10^{-2} \leq k \leq 0.75$	76
$0.85 \leq k \leq 400$	51
$450 \leq k \leq 850$	10
$900 \leq k \leq 1000$	10

Tabela 5: Distribuição dos nós da spline $P(k)$ que aproxima a função $f(k) = (2\pi^2)^{-1} k^{n+2} T(k)^2$. Em cada intervalo de k , o espaçamento entre os nós é linear, exceto em $0 \leq k \leq 10^{-5}$ e em $0.85 \leq k \leq 400$, no qual utilizamos uma separação logarítmica entre os nós.

O algoritmo da GSL que utilizamos para calcular a spline $P(k)$, gera polinômios

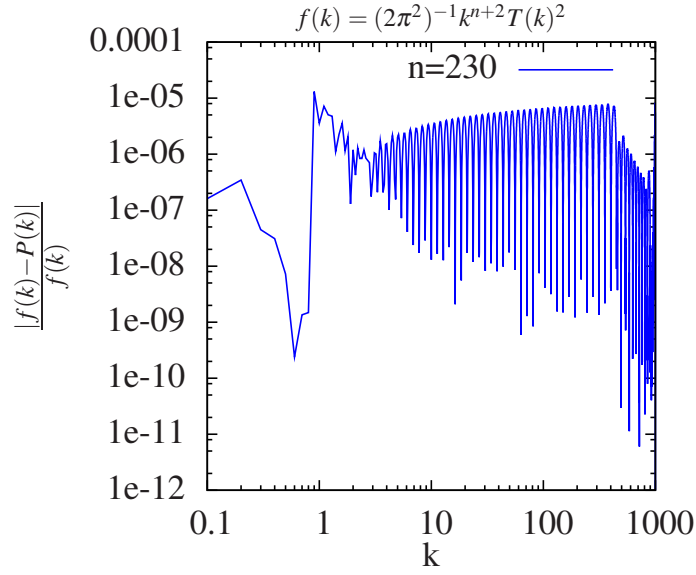


Figura 24: Diferença fracionária entre a função $f(k)$ e a spline que aproxima $P(k)$ em função do modo k . Nesta figura, foram utilizados os parâmetros cosmológicos do *Hubble Volume* [4].

centrados no ponto inicial do seu respectivo trecho, isto é,

$$P_i(k) = a_i(k - k_i)^3 + b_i(k - k_i)^2 + c_i(k - k_i) + d_i. \quad (\text{B.7})$$

Com isso, temos que

$$\begin{aligned} \int_{k_i}^{k_{i+1}} dk P_i(k) W^2(k, R) &= (d_i - c_i k_i + b_i k_i^2 - a_i k_i^3) I_0^W \\ &+ (c_i - 2b_i k_i + 3a_i k_i^2) I_1^W \\ &+ (b_i - 3a_i k_i) I_2^W \\ &+ a_i I_3^W, \end{aligned} \quad (\text{B.8})$$

onde

$$I_0^W = \int_{k_i}^{k_{i+1}} dk W^2(k, R), \quad (\text{B.9})$$

$$I_1^W = \int_{k_i}^{k_{i+1}} dk k W^2(k, R), \quad (\text{B.10})$$

$$I_2^W = \int_{k_i}^{k_{i+1}} dk k^2 W^2(k, R), \quad (\text{B.11})$$

$$I_3^W = \int_{k_i}^{k_{i+1}} dk k^3 W^2(k, R). \quad (\text{B.12})$$

As soluções das integrais da Eq. (B.6) e, conseqüentemente, da Eq. (B.8) são analíticas tanto para a função janela do tipo cartola quanto Gaussiana (veja a subseção 2.2). A princípio, teríamos que calcular essas soluções para cada valor do raio, R , do aglomerado de galáxias. Portanto, como mais um passo da otimização, nós fazemos a mudança de variável $k' = kR$ para que os resultados das integrais sejam independentes de R e, assim, possamos calculá-los somente uma vez dado um conjunto de parâmetros cosmológicos. E, com isso, a variância é obtida calculando-se

$$\begin{aligned} \sigma^2(z, R) &= A \left(\frac{D(z)}{D(0)} \right)^2 \frac{C^W}{(2\pi^2)} \sum_{i=0}^{n-1} \frac{d_i}{R} I_0^{W'} + \frac{c_i}{R^2} (I_1^{W'} - k_i' I_0^{W'}) \\ &+ \frac{b_i}{R^3} (I_2^{W'} - 2k_i' I_1^{W'} + k_i'^2 I_0^{W'}) \\ &+ \frac{a_i}{R^4} (I_3^{W'} - 3k_i' I_2^{W'} + 3k_i'^2 I_1^{W'} - k_i'^3 I_0^{W'}), \end{aligned} \quad (\text{B.13})$$

onde o superscrito $'$ significa que as integrações são feitas na variável k' , W indica qual é a função janela que está sendo utilizada e C^W é uma combinação de constantes da função janela que não incluímos em $I_n^{W'}$ colocando-o em evidência na equação acima. C^W é igual a $C^{TH} = 9$ para a função janela cartola e $C^G = 1$ para a função Gaussiana. Naturalmente, dada a expressão (B.13), temos que o erro fracional de 10^{-5} , imposto por $\frac{|P(k)-f(k)|}{f(k)} \leq 10^{-5}$, é mantido no resultado da variância.

Além da dependência em $\sigma^2(z, R)$, a função de massa (equação 2.6) também depende de $d\sigma/dR$. Como essa derivada é com respeito a R , não há nenhuma mudança na função $f(k)$ e, logo, a mesma spline $P(k)$ é usada para calcular as quantidades σ^2 e $d\sigma^2/dR$. Esse último é obtido a partir do mesmo procedimento descrito acima mas fazendo a substituição

$W^2(k, R) \rightarrow 2WdW/dR$ nas equações (B.5), (B.6) e (B.8). A partir disso, temos que

$$\begin{aligned} \frac{d\sigma^2(z, R)}{dR} &= A \left(\frac{D(z)}{D(0)} \right)^2 \frac{C^W}{(2\pi^2)} \sum_{i=0}^{n-1} \frac{d_i}{R} I_4^{iW} + \frac{c_i}{R^2} (I_5^{iW} - k_i' I_4^{iW}) \\ &+ \frac{b_i}{R^3} (I_6^{iW} - 2k_i' I_5^{iW} + k_i'^2 I_4^{iW}) \\ &+ \frac{a_i}{R^4} (I_7^{iW} - 3k_i' I_6^{iW} + 3k_i'^2 I_5^{iW} - k_i'^3 I_4^{iW}), \end{aligned} \quad (\text{B.14})$$

onde $C^{TH} = -18$ e $C^G = -2$.

As soluções analíticas das equações (B.9) – (B.12) com as funções janela cartola e Gaussiana e das equações equivalentes com $2WdW/dR$, denotadas por $I_4^{iW} - I_7^{iW}$, são apresentadas na subseção B.2.1.

Além do tempo de execução das integrais com solução analítica ser menor que o tempo do cálculo das integrações numéricas “diretas”, nós também ganhamos eficiência com o fato de que precisamos calcular somente uma spline para a variância e a sua derivada com relação à R .

B.2.1 Soluções Analíticas

Seguem abaixo as expressões analíticas.

- **Cartola**

$$\begin{aligned} I_0^{th} &= \int_{k_i'}^{k_{i+1}'} dk' \frac{1}{k'^2} [j_1(k')]^2 = \frac{1}{30k_i'^5} [4k_i'^5 \text{Si}(2k_i') + k_i'(6 + k_i'^2) \sin(2k_i') \\ &+ (3 - k_i'^2 + 2k_i'^4) \cos(2k_i') - 5k_i'^2 - 3]_{k_i}^{k_{i+1}} \end{aligned} \quad (\text{B.15})$$

$$I_1^{th} = \int_{k_i'}^{k_{i+1}'} dk' \frac{1}{k'} [j_1(k')]^2 = \frac{1}{8k_i'^4} \left[2k_i' \sin(2k_i') + \cos(2k_i') - 2k_i'^2 - 1 \right]_{k_i}^{k_{i+1}} \quad (\text{B.16})$$

$$\begin{aligned} I_2^{th} &= \int_{k_i'}^{k_{i+1}'} dk' [j_1(k')]^2 = \frac{1}{6k_i'^3} [2k_i'^3 \text{Si}(2k_i') + 2k_i' \sin(2k_i') \\ &+ (1 + k_i'^2) \cos(2k_i') - 3k_i'^2 - 1]_{k_i}^{k_{i+1}} \end{aligned} \quad (\text{B.17})$$

$$I_3^{th} = \int_{k'_i}^{k'_{i+1}} dk' k' [j_1(k')]^2 = \frac{1}{2k'^2} \{k'[\sin(2k') + k' \log(k') - k' \text{Ci}(2k')] - \sin^2(k')\}_{k_i}^{k_{i+1}} \quad (\text{B.18})$$

$$I_4^{th} = \int_{k'_i}^{k'_{i+1}} dk' \frac{1}{k'} j_1(k') j_2(k') = \frac{1}{60k'^5} [4k'^5 \text{Si}(2k') + k'(36 + k'^2) \sin(2k') + 2(9 - 8k'^2 + k'^4) \cos(2k') - 20k'^2 - 18]_{k_i}^{k_{i+1}} \quad (\text{B.19})$$

$$I_5^{th} = \int_{k'_i}^{k'_{i+1}} dk' j_1(k') j_2(k') = \frac{1}{8k'^4} [6k' \sin(2k') + (3 - 2k'^2) - 4k'^2 - 3]_{k_i}^{k_{i+1}} \quad (\text{B.20})$$

$$I_6^{th} = \int_{k'_i}^{k'_{i+1}} dk' k' j_1(k') j_2(k') = \frac{1}{2k'^3} [k'^3 \text{Si}(2k') + 2k' \sin(2k') + \cos(2k') - 2k'^2 - 1]_{k_i}^{k_{i+1}} \quad (\text{B.21})$$

$$I_7^{th} = \int_{k'_i}^{k'_{i+1}} dk' k'^2 j_1(k') j_2(k') = \ln(k') - \text{Ci}(2k') + \frac{1}{2k'^2} [-3 \sin^2(k') + k' \cos(k')(6 \sin(k') - k' \cos(k'))]_{k_i}^{k_{i+1}}, \quad (\text{B.22})$$

onde $\text{Si}(2k')$ e $\text{Ci}(2k')$ são as funções seno e o cosseno integral, respectivamente.

- **Gaussiana**

$$I_0^g = \int_{k'_i}^{k'_{i+1}} dk' \exp(-k'^2) = \left[\frac{1}{2} \sqrt{\pi} \text{erf}(k') \right]_{k_i}^{k_{i+1}} \quad (\text{B.23})$$

$$I_1^g = \int_{k'_i}^{k'_{i+1}} dk' k' \exp(-k'^2) = \left[-\frac{e^{-k'^2}}{2} \right]_{k_i}^{k_{i+1}} \quad (\text{B.24})$$

$$I_2^g = I_4^g = \int_{k'_i}^{k'_{i+1}} dk' k'^2 \exp(-k'^2) = \left[\frac{1}{4} (-2e^{-k'^2} k' + \sqrt{\pi} \text{erf}(k')) \right]_{k_i}^{k_{i+1}} \quad (\text{B.25})$$

$$I_3'^g = I_5'^g = \int_{k_i'}^{k_{i+1}'} dk' k'^3 \exp(-k'^2) = \left[-\frac{1}{2} e^{-k'^2} (1 + k'^2) \right]_{k_i}^{k_{i+1}} \quad (\text{B.26})$$

$$I_6'^g = \int_{k_i'}^{k_{i+1}'} dk' k'^4 \exp(-k'^2) = \left[-\frac{1}{4} e^{-k'^2} k' (3 + 2k'^2) + \frac{3}{8} \sqrt{\pi} \operatorname{erf}(k') \right]_{k_i}^{k_{i+1}} \quad (\text{B.27})$$

$$I_7'^g = \int_{k_i'}^{k_{i+1}'} dk' k'^5 \exp(-k'^2) = \left[-\frac{1}{2} e^{-k'^2} (2 + 2k'^2 + k'^4) \right]_{k_i}^{k_{i+1}}, \quad (\text{B.28})$$

onde $\operatorname{erf}(k')$ é a função erro cuja definição pode ser vista na equação (2.4).

B.3 Cálculo Numérico da Abundância de Aglomerados

O cálculo do número de aglomerados de galáxias no i -ésimo bin do desvio para o vermelho z envolve 3 integrações, como dado pela equação (3.1). Além da otimização que descrevemos na seção B.2 para melhorarmos o tempo do cálculo da variância e, conseqüentemente, da função de massa (equação 2.1), nós também utilizamos o método de interpolação por spline para efetuar a integração na variável z' da equação (3.1), onde, neste caso, $f(z')$ é o integrando dessa variável.

Usamos o algoritmo *gsl-integration-qag* para calcular as integrais da Eq. (2.34) em z'' e em M e, assim, obter $f(z')$. Já a especificação dos nós da spline $P(z')$ foi obtida requerendo que $\frac{|f(z')-P(z')|}{f(z')} \leq 10^{-3}$. Neste caso, diferentemente da seção B.2 onde só precisamos de um conjunto de parâmetros cosmológicos para determinar o número de nós, é necessário fixar também o valor de M_{\min} e o intervalo do desvio para o vermelho, que escolhemos como $z \in [0, 1.4]$. Para estudar o efeito da massa mínima do aglomerado, nós calculamos $f(z')$ e $P(z')$ para diferentes valores de M_{\min} e notamos que a condição, $\frac{|f(z')-P(z')|}{f(z')} \leq 10^{-3}$, é satisfeita com um número menor de nós quanto menor for a massa limite. Portanto, para termos uma única spline para calcular a abundância de aglomerados em bins de z , nós determinamos $P(z')$ a partir de $M_{\min} = 10^{15} h^{-1} M_{\odot}$. A tabela B.3 mostra a distribuição dos 30 nós e, na figura 25, vemos o resultado de $\frac{|f(z')-P(z')|}{f(z')} \leq 10^{-3}$ para $M_{\min} = 5 \times 10^{13}$, 2×10^{14} e $10^{15} h^{-1} M_{\odot}$.

Com isso, ao invés de calcular as três integrais da Eq. (2.34), reduzimos o gasto computacional a integrais de polinômios de terceira ordem,

$$N_i = \int_{z_i - \frac{\Delta z}{2}}^{z_i + \frac{\Delta z}{2}} dz' P(z'), \quad (\text{B.29})$$

Intervalo em z	Número de nós
$0 \leq z \leq 10^{-2}$	6
$0.03 \leq z \leq 0.09$	3
$0.15 \leq z \leq 0.25$	3
$0.33 \leq z \leq 1.29$	13
$1.33 \leq z \leq 1.36$	2
$1.38 \leq z \leq 1.4$	3

Tabela 6: Distribuição dos nós da spline $P(z')$ que aproxima a função $f(z') = A_{survey} \frac{c}{H(z')}$ $\left(\int_0^{z'} dz'' \frac{c}{H(z'')} \right)^2 \int_{M_{min}}^{\infty} dM \frac{dn}{dM}$. Cada linha representa um intervalo uniforme, exceto a primeira em que $z'_0 = 0$, $z'_1 = 10^{-6}$ e os outros 4 nós seguem uma progressão geométrica com razão igual a 10.

onde

$$P(z') = \sum_{j=0}^{29} P_j(z').$$

Por exemplo, se quisermos saber quantos aglomerados o modelo prevê entre $0.5 \leq z \leq 0.6$, precisamos calcular

$$N_i = \int_{0.5}^{0.57} dz' P_{14}(z') + \int_{0.57}^{0.6} dz' P_{15}(z'). \quad (\text{B.30})$$

B.4 Spline Bidimensional

Na seção 3.5, nós apresentamos a verossimilhança (3.15) que depende da função $\frac{d^2 N(M, z)}{dz d \ln M}$, dada pela equação (3.14). Como essa é uma função de duas variáveis, o processo de otimização por spline tem que ser estendido.

Nós implementamos o cálculo da spline bidimensional de uma função $f(x, y)$ decompondo o problema em $n + 1$ splines unidimensionais da função $f(x, y_i)$, dados n valores de y_i fixos. Chamaremos de splines horizontais $H_i(x)$ as n funções que interpolam $f(x, y_i)$, como descrito na seção B.1. Uma vez preparadas as splines H_i , podemos calcular o valor interpolado de f em um ponto qualquer (x, y) da seguinte maneira. Primeiramente, utilizamos as splines horizontais para calcular $H_i(x)$. Com esses n números, preparamos uma spline vertical $V(y, \{H_i(x)\})$ e a usamos para calcular o valor interpolado de $f(x, y) \approx V(y, \{H_i(x)\})$.

Aplicamos este procedimento para calcular a spline da função $f(M, z) = \frac{d^2 N(M, z)}{dz d \ln M}$. O limite inferior do intervalo na massa, $M_i \leq M \leq M_f$, onde a interpolação será preparada, é dado por M_{min} (lembrando que as massas dos aglomerados são $\gtrsim 10^{12} h^{-1} M_{\odot}$). Já o limite superior M_f é determinado pelo valor em que ocorre a convergência da integração

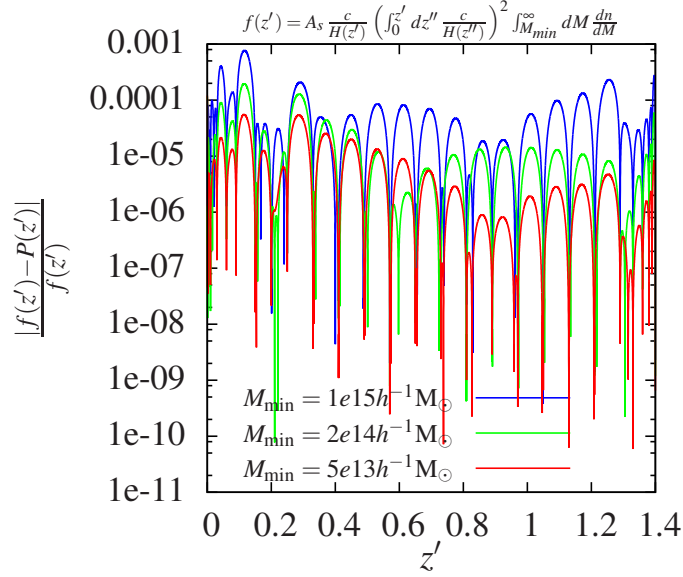


Figura 25: Erro relativo da spline entre a função $f(z')$ e $P(z')$ para os valores de $M_{\min} = 5 \times 10^{13}, 2 \times 10^{14}$ e $10^{15} h^{-1} M_{\odot}$. Utilizamos os parâmetros cosmológicos do *Hubble Volume* [4].

numérica, isto é, $\int_{M_i}^{\infty} f(M, z) dM \approx \int_{M_i}^{M_f} f(M, z) dM$ dentro de uma precisão determinada. Neste caso, escolhemos uma precisão relativa de 10^{-8} e, fazendo a análise para diferentes conjuntos de parâmetros cosmológicos, obtemos $M_f = 10^{16.3} h^{-1} M_{\odot}$. Em relação ao desvio para o vermelho, a spline pode ser preparada no intervalo $z_{\min} \leq z \leq z_{\max}$, onde os valores extremos possíveis são $z_{\min} = 10^{-6}$ e $z_{\max} = 2$.

A distribuição dos nós é não uniforme nem em M e nem em z , sendo a escolha dos mesmos determinada de forma que o erro entre a spline $V(z, \{H_i(M)\})$ e a função $f(M, z)$ seja $|f(M, z) - V(z, \{H_i(M)\})| \leq 5 \times 10^{-4} f(M, z) + 10^{-9}$. Na figura 26, nós mostramos o resultado da diferença fracionária $\frac{|f(M, z) - V(z, \{H_i(M)\})|}{f(M, z)}$ para diferentes conjuntos de parâmetros cosmológicos. No primeiro painel, nós utilizamos os mesmos parâmetros da simulação Carmen ($\Omega_b = 0.04$, $\Omega_{DE} = 0.75$, $w_0 = -1$, $w_1 = 0$, $\sigma_8 = 0.8$, $n_s = 1$ e $h = 0.7$). Os parâmetros utilizados nos painéis 2 e 3 foram escolhidos como exemplos baseados em alguns resultados do capítulo 5 (região de 3σ das figuras 13, 14 e 15 para $M_{\min} = 2 \times 10^{14} h^{-1} M_{\odot}$). Com isso, no segundo painel, modificamos $\sigma_8 = 0.9$, $w_0 = -2.1$, $\Omega_{DE} = 0.91$ e, no terceiro, $\sigma_8 = 0.73$, $w_0 = 0$, $\Omega_{DE} = 0.55$. O terceiro painel apresenta uma região vermelha maior que nos outros casos. Isso ocorre porque apresentamos na

figura o erro relativo. Neste caso, é o erro absoluto que domina já que a densidade de aglomerados diminui para valores mais baixos de σ_8 .

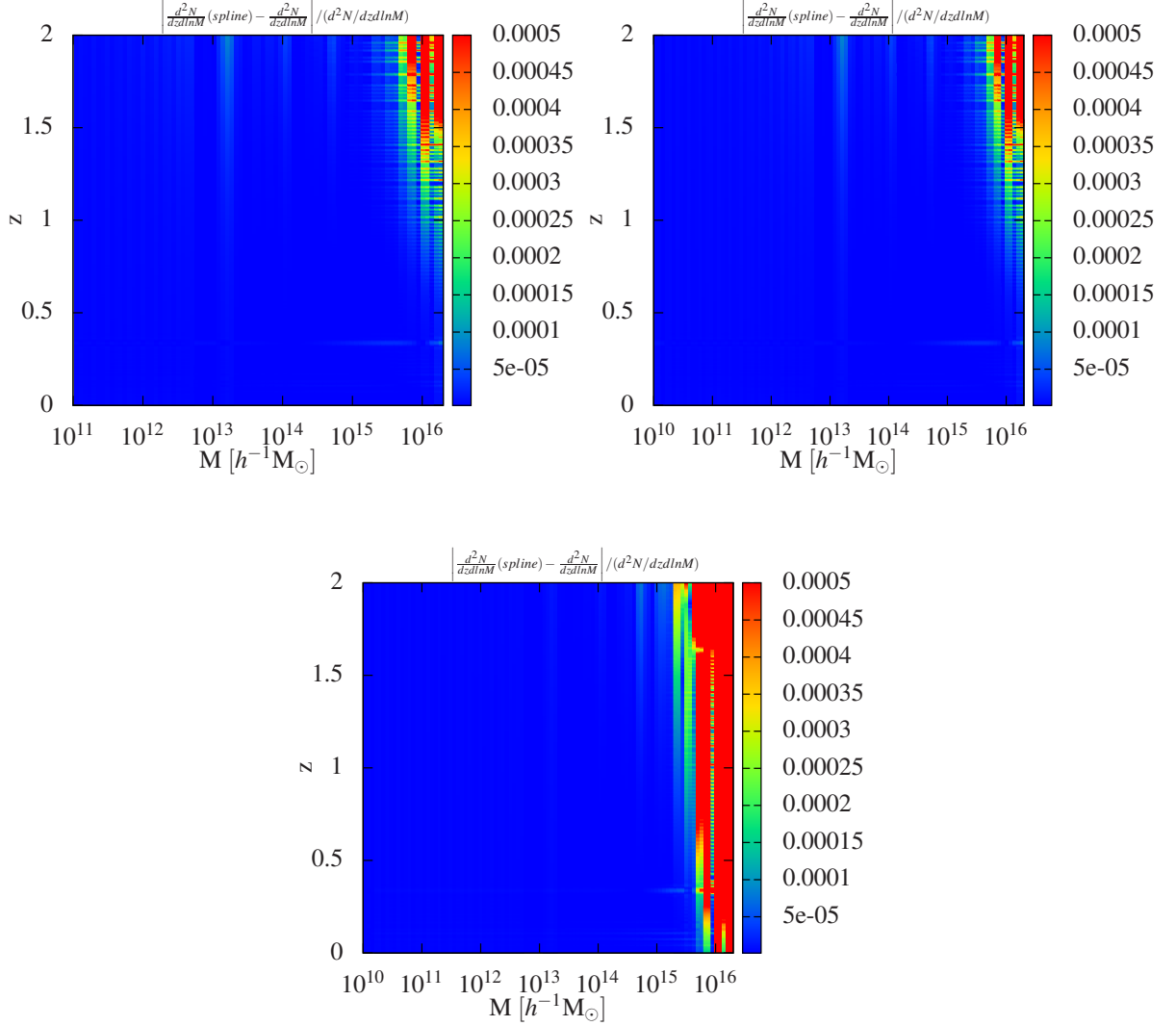


Figura 26: Diferença fracionária entre a função $f(M, z)$ e a spline que a aproxima $V(z, \{H_i(M)\})$. No primeiro painel, utilizamos os parâmetros cosmológicos da simulação Carmen. No segundo, mudamos os valores de $\sigma_8 = 0.9$, $w_0 = -2.1$, $\Omega_{DE} = 0.91$ e, no terceiro, $\sigma_8 = 0.73$, $w_0 = 0$, $\Omega_{DE} = 0.55$.

Essa mesma análise é feita nos casos em que incluímos a incerteza no desvio para o vermelho fotométrico, a incerteza na massa e quando combinamos as duas como mostra as figuras 27, 28 e 29, respectivamente. Nesses casos, os limites de integração das integrais na massa e no desvio para o vermelho das equações (6.9) e (6.17), respectivamente, são obtidas exigindo a convergência das mesmas com precisão de 5×10^{-4} . Obtemos, então, os intervalos de integração $M_{\text{obs}} - 7\sigma_{\ln M} \leq M \leq M_{\text{obs}} + 7\sigma_{\ln M}$ e $z^{\text{phot}} - 10\sigma_z \leq z \leq$

$$z^{phot} + 10\sigma_z.$$

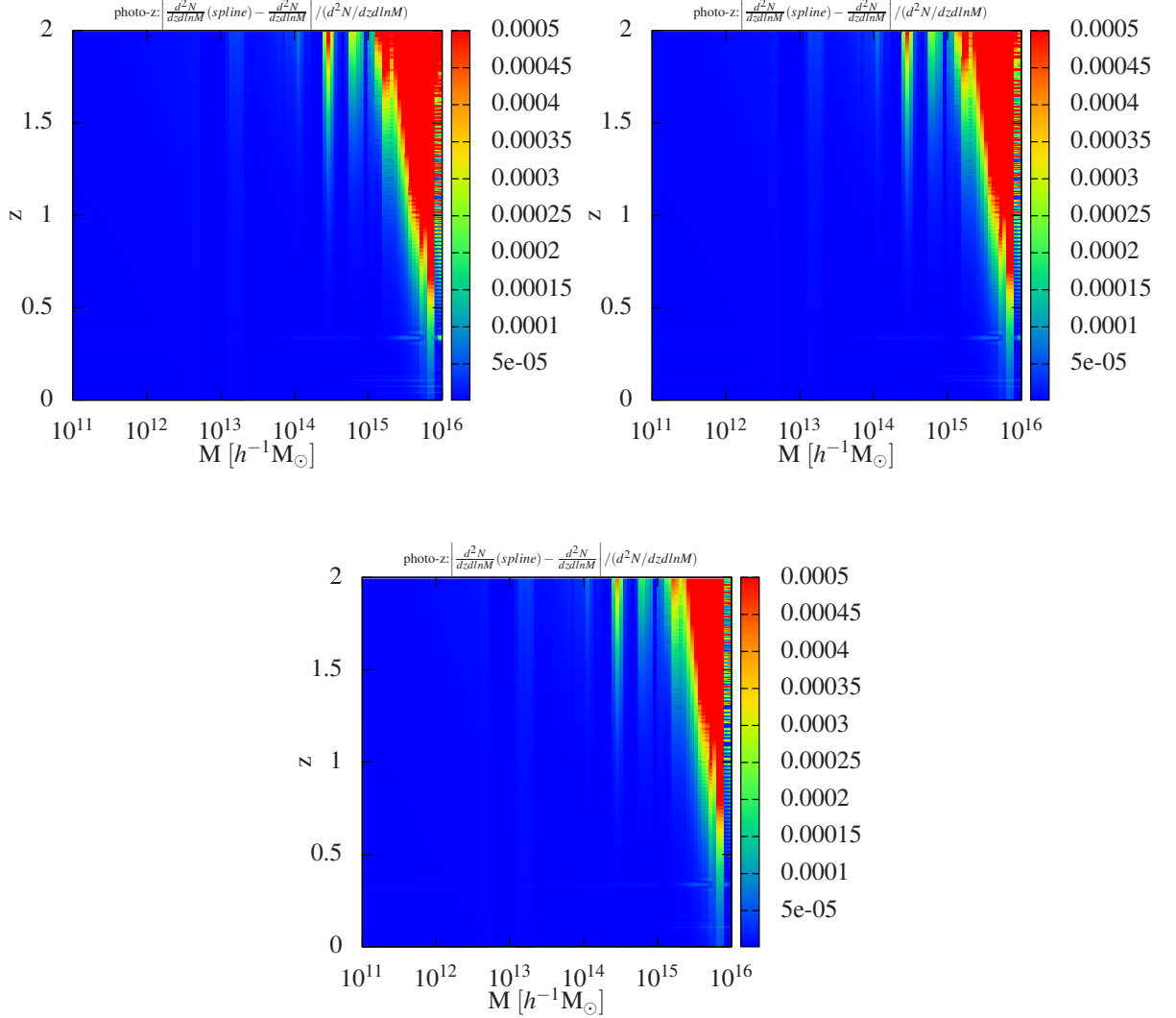


Figura 27: Diferença fracionária entre a função $f(M, z)$ e a spline que a aproxima $V(z, \{H_i(M)\})$ incluindo a incerteza no desvio para o vermelho fotométrico. No primeiro painel, usamos $\sigma_z = 0.01(1+z)$. No segundo, $\sigma_z = 0.03(1+z)$ e, no terceiro painel, fizemos $\sigma_z = 0.05(1+z)$.

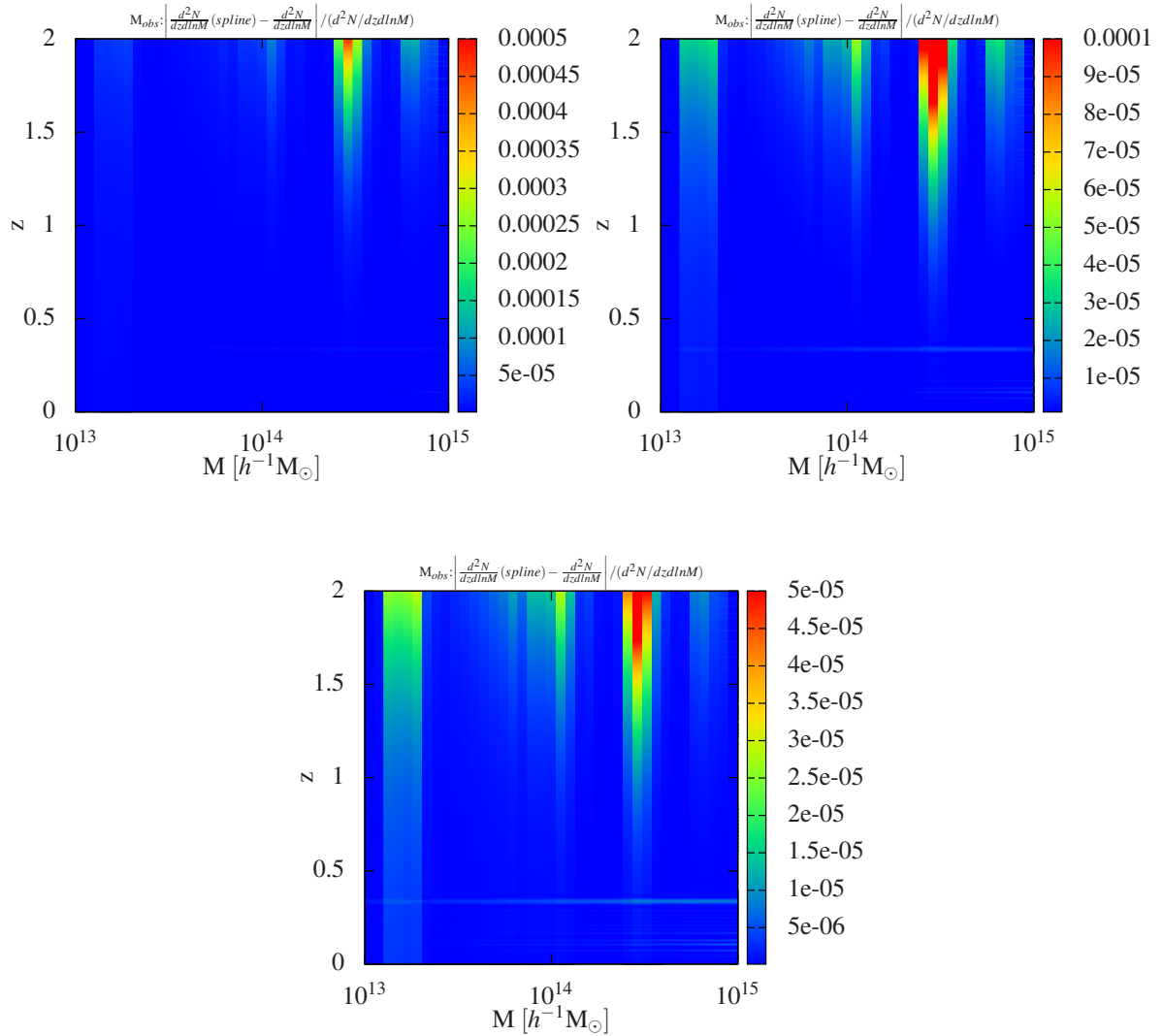


Figura 28: Diferença fracionária entre a função $f(M, z)$ e a spline que a aproxima $V(z, \{H_i(M)\})$ incluindo a incerteza na massa. No primeiro painel, usamos $\sigma_{\ln M} = 0.1$. No segundo, $\sigma_{\ln M} = 0.3$ e, no terceiro painel, fizemos $\sigma_{\ln M} = 0.5$.

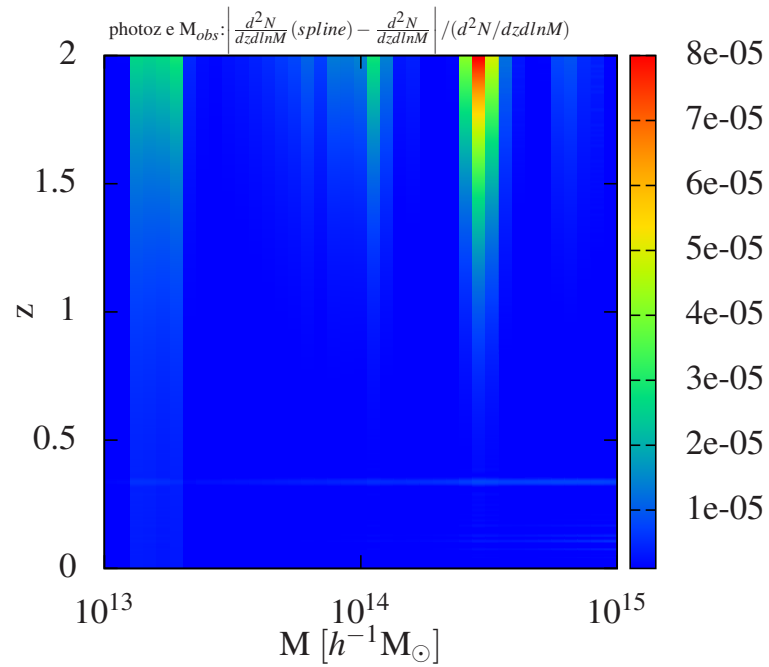


Figura 29: Diferença fracionária entre a função $f(M, z)$ e a spline que a aproxima $V(z, \{H_i(M)\})$ incluindo tanto a incerteza na massa quanto no desvio para o vermelho fotométrico. Utilizamos $\sigma_{\ln M} = 0.1$ e $\sigma_z = 0.01(1 + z)$.

B.5 Validação e Controle do Código

Nesta seção, nós mostramos alguns testes feitos para validar os códigos da biblioteca *CosmoClusterLib*, apresentando comparações dos resultados obtidos desses códigos com resultados da literatura. Os gráficos mostrados abaixo foram produzidos com nossos códigos, reproduzindo a literatura citada.

Função Crescimento

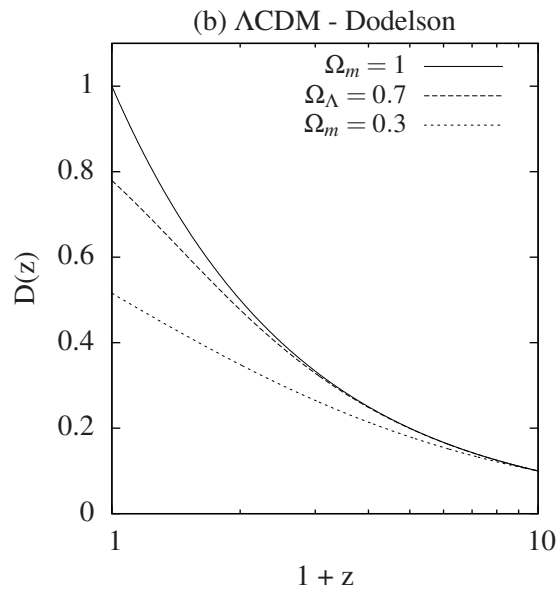


Figura 30: Função crescimento $D(z)$ para o modelo Λ CDM em três diferentes casos: 1) universo plano, $\Omega_m = 1$; 2) universo plano, $\Omega_\Lambda = 0.7$; 3) universo aberto, $\Omega_m = 0.3$. Esta figura é análoga à figura 7.12 da referência [67], que usa como normalização $D(z \simeq 10) = 0.1$.

Função Transferência

Na figura 31, nós mostramos o resultado das funções transferência BBKS [63], BBKS Modificada [71] e EH [72].

Função de Massa

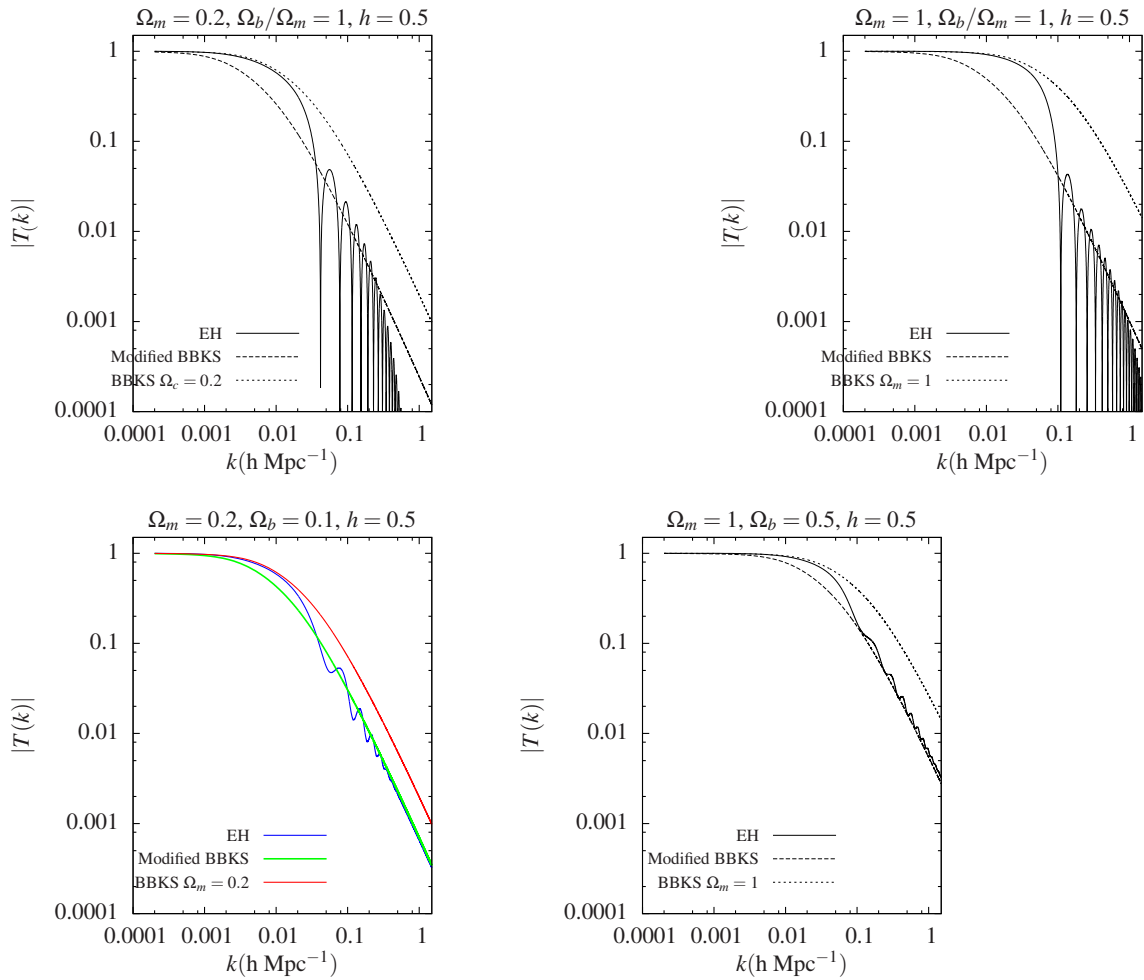


Figura 31: Funções transferências EH [72], BBKS [63] e BBKS Modificada [71] para quatro diferentes cosmologias. Esta figura é similar à figura 3 da referência [72]. Nós estendemos a análise incluindo a função de transferência BBKS modificada.

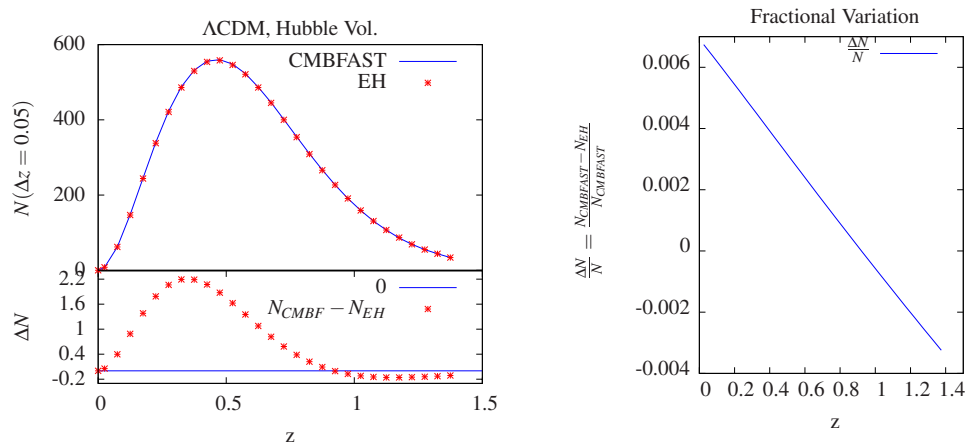


Figura 32: Comparação entre a função transferência EH e CMBFAST. Painel direito: número de aglomerados calculado com os parâmetros da simulação do Hubble Volume [4], função multiplicidade de Jenkins [8] e $M_{min} = 2 \times 10^{14} h^{-1} M_{\odot}$. Painel direito: diferença relativa entre as abundâncias obtidas com EH e CMBFAST.

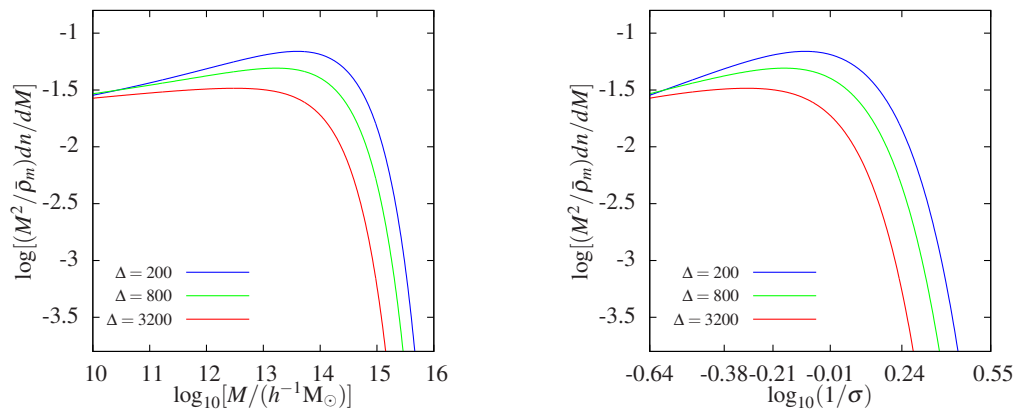


Figura 33: Função de massa de Tinker et al. para $\Delta = 200, 800$ e 3200 em função do logaritmo da massa (painel esquerdo) e em função do logaritmo do inverso do desvio padrão σ (painel direito). Equivalente à figura 5 da referência [10].

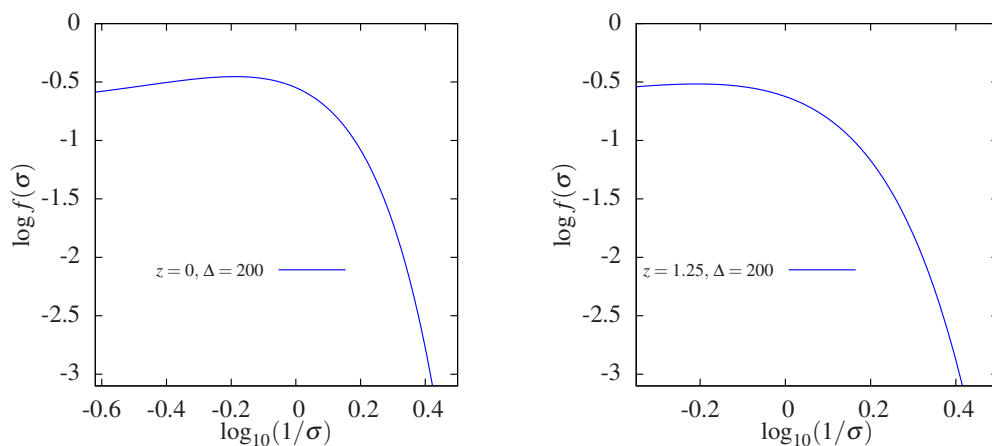


Figura 34: Função multiplicidade de Tinker et al. para $\Delta = 200$ em função do logaritmo do inverso do desvio padrão σ em $z = 0$ (painel esquerdo) e em $z = 1.25$ (painel direito). Equivalente à figura 6 da referência [10].

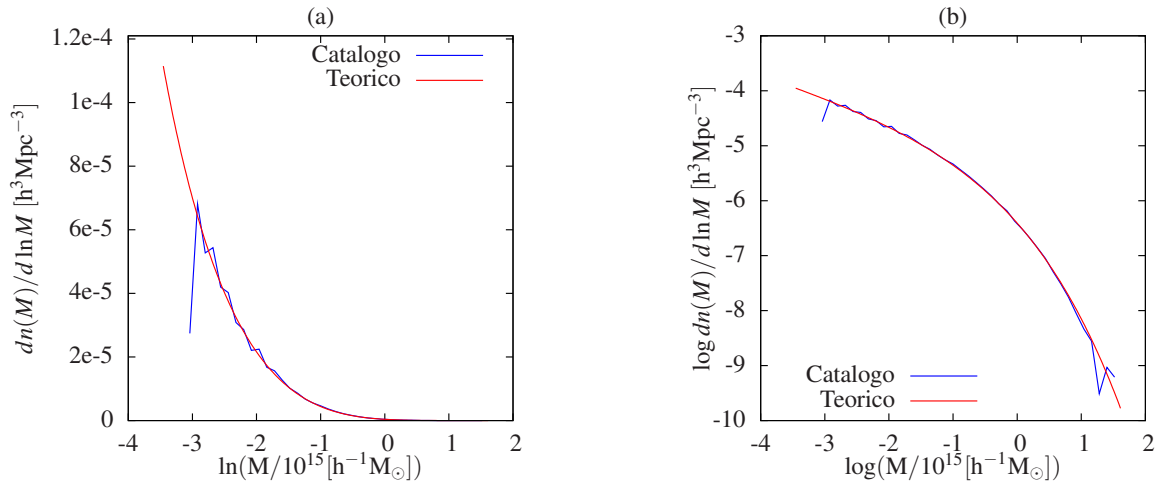


Figura 35: Funções de massa para $\Delta = 200$ e $z = 0$ obtidas a partir do catálogo de aglomerados Hubble Volume (curva azul) e da função multiplicidade de Jenkins et al. [8] (curva vermelha) em função do logaritmo neperiano (painel (a)). O painel (b) mostra as mesmas funções de massa em função da massa, porém em escala logarítmica. Equivalente à figura 5 da referência [4].

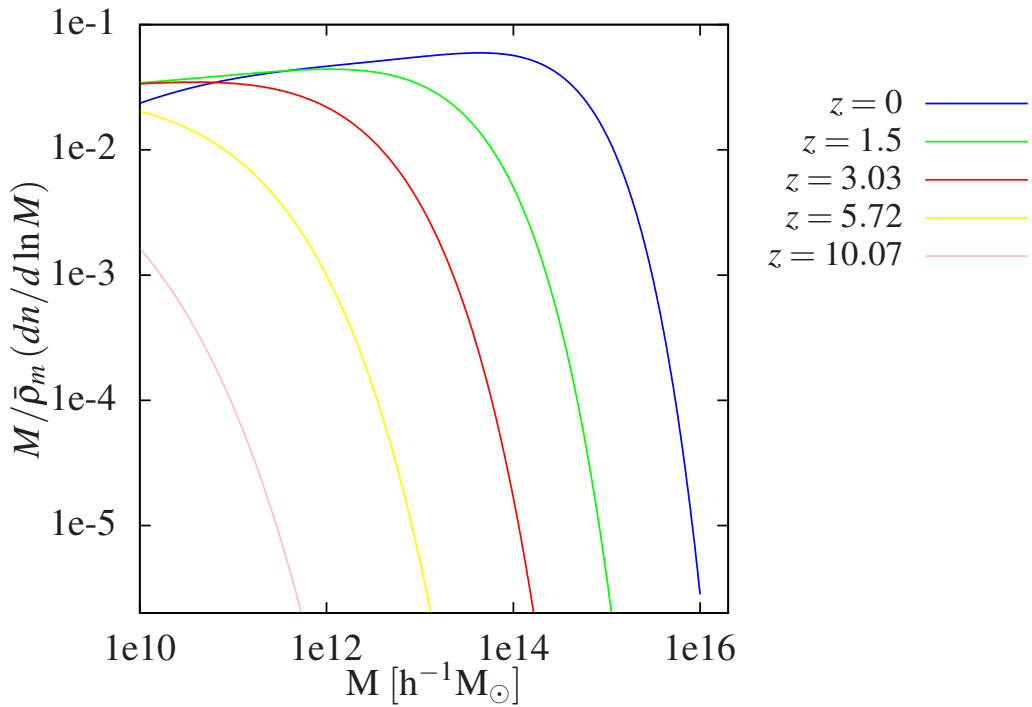


Figura 36: Função de massa de Jenkins et al. [8] em função da massa calculada nos desvios para o vermelho z conforme indica a legenda. Equivalente à figura 2 do artigo [5].

Referências

- [1] BAHCALL, N. A. Large-scale structure in the universe indicated by galaxy clusters. *ARA&A*, v. 26, p. 631, 1988.
- [2] VOIT, G. M. Tracing cosmic evolution with clusters of galaxies. *Rev. Mod. Phys.*, v. 77, p. 207–258, 2005.
- [3] FRENK, C. S. et al. The santa barbara cluster comparison project: a comparison of cosmological hydrodynamics solutions. *Astrophys. J.*, v. 525, p. 554–582, 1999.
- [4] EVRARD, A. E. et al. Galaxy clusters in hubble volume simulations: Cosmological constraints from sky survey populations. *Astrophys. J.*, v. 573, p. 7–36, 2002.
- [5] SPRINGEL, V. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, v. 435, p. 629–636, 2005.
- [6] HEITMANN, K. et al. Robustness of cosmological simulations i: Large scale structure. *Astrophys. J. Suppl.*, v. 160, p. 28–58, 2005.
- [7] HEITMANN, K. et al. The cosmic code comparison project. *Comput. Sci. Dis.*, v. 1, p. 015003, jun. 2008.
- [8] JENKINS, A. et al. Mass function of dark matter halos. *Mon. Not. Roy. Astron. Soc.*, v. 321, p. 372–384, 2001.
- [9] WARREN, M. et al. Precision determination of the mass function of dark matter halos. *Astrophys. J.*, v. 646, p. 881–885, 2006.
- [10] TINKER, J. et al. Toward a halo mass function for precision cosmology: The limits of universality. *Astrophys. J.*, v. 688, p. 709–728, 2008.
- [11] SHETH, R. K.; TORMEN, G. Large scale bias and the peak background split. *Mon. Not. Roy. Astron. Soc.*, v. 308, p. 119, 1999.
- [12] SHETH, R. K.; MO, H. J.; TORMEN, G. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *Mon. Not. Roy. Astron. Soc.*, v. 323, p. 1–12, 2001.
- [13] DIEMAND, J.; MOORE, B.; STADEL, J. Convergence and scatter of cluster density profiles. *Mon. Not. Roy. Astron. Soc.*, v. 353, p. 624, 2004.
- [14] MERRITT, D. et al. Empirical models for dark matter halos. i. nonparametric construction of density profiles and comparison with parametric models. *Astron. J.*, v. 132, p. 2685, 2006.

- [15] GAO, L. et al. The redshift dependence of the structure of massive Λ CDM halos. *Mon. Not. Roy. Astron. Soc.*, v. 387, p. 536, nov. 2007.
- [16] HOST, O.; HANSEN, S. H. A detailed statistical analysis of the dark matter mass profiles of galaxy clusters. jul. 2009.
- [17] BOYARSKY, A. et al. New evidence for dark matter. nov. 2009.
- [18] COE, D. Dark matter halo mass profiles. maio 2010.
- [19] COLLABORATION, T. D. E. S. The dark energy survey. *astro-ph/0510346*, 2007.
- [20] ANNIS, J. et al. *A Proposal to NOAO for the Dark Energy Survey*. 2004. Disponível em: <https://des.fnal.gov/main_documents/A_Proposal_to_NOAO.pdf>.
- [21] KRAVTSOV, A. V.; VIKHLININ, A.; NAGAI, D. A new robust low-scatter x-ray mass indicator for clusters of galaxies. *Astrophys. J.*, v. 650, p. 128–136, 2006.
- [22] VIKHLININ, A. et al. Chandra sample of nearby relaxed galaxy clusters: mass, gas fraction, and mass-temperature relation. *Astrophys. J.*, v. 640, p. 691, 2006.
- [23] CARLSTROM, J. E.; HOLDER, G. P.; REESE, E. D. Cosmology with the sunyaev - zel'dovich effect. *Annu. Rev. Astron. Astrophys.*, v. 40, p. 643–680, 2002.
- [24] SHELDON, E. S. et al. Weak lensing measurements of 42 sdss/rass galaxy clusters. *Astrophys. J.*, v. 554, p. 881–887, 2001.
- [25] SHELDON, E. S. et al. Cross-correlation weak lensing of sdss galaxy clusters i: Measurements. *Astrophys. J.*, v. 703, p. 2217–2231, set. 2009.
- [26] CARLBERG et al. Galaxy cluster virial masses and Ω . *Astrophys. J.*, v. 462, p. 32, 1996.
- [27] RINES, K.; DIAFERIO, A.; NATARAJAN, P. The virial mass function of nearby sdss galaxy clusters. *Astrophys. J.*, v. 657, p. 183–196, 2007.
- [28] ROZO, E. et al. Improvement of the richness estimates of maxbcg clusters. *Astrophys. J.*, v. 703, p. 601–613, 2009.
- [29] OYAIZU, H. et al. Photometric redshift error estimators. *Astrophys. J.*, v. 689:709, 2008, nov. 2007.
- [30] LIMA, M. et al. Estimating the redshift distribution of photometric galaxy samples. *Mon. Not. Roy. Astron. Soc.*, v. 390:118, 2008, jan. 2008.
- [31] CUNHA, C. E. et al. Estimating the redshift distribution of photometric galaxy samples ii. applications and tests of a new method. *Mon. Not. Roy. Astron. Soc.*, v. 396:2379–2398, 2009, out. 2008.
- [32] GIT. *Version Control System*. Disponível em: <<http://git-scm.com/>>.
- [33] REFREGIER, A. Shapelets i. a method for image analysis. *Mon. Not. R. Astron. Soc.*, v. 338, p. 35–47, 2003.

- [34] LASDAMAS. *Simulação Carmen (Large Suite of Dark Matter Simulations - LasDamas)*. Disponível em: <<http://lss.phy.vanderbilt.edu/lasdamas/simulations.html>>.
- [35] LASDAMAS. *Wiki das simulações LasDamas*. Disponível em: <http://coop.vanderbilt.edu/lasdamas/LASDAMAS_Wiki_Home>.
- [36] BUSHA, M. T.; WECHSLER, R. H. Making mock galaxy catalogs with addgals. *Moriond.in2p3.fr/J08/proceedings/busha.pdf*.
- [37] SCOVILLE, N.; AUSSEL, H.; BRUSA, M. The cosmic evolution survey (COSMOS): OVERVIEW. *Astrophys. J. Suppl. Ser.*, v. 172, p. 1–8, 2007.
- [38] GIRARDI, L. et al. Star counts in the Galaxy - Simulating from very deep to very shallow photometric surveys with the TRILEGAL code. *A&A*, v. 436, p. 895–915, 2005.
- [39] FERREIRA, P. da C. *Simulações de arcos gravitacionais em imagens e estudo de sua razão axial*. Dissertação (Mestrado) — Universidade Federal do Rio de Janeiro, 2010.
- [40] HANISCH R. J., e. a. Definition of the Flexible Image Transport System (FITS). *A&A*, v. 376, p. 359–380, 2001.
- [41] KOESTER, B. P. et al. Maxbcg: A red-sequence galaxy cluster finder. *Astrophys. J.*, v. 660, p. 221, 2007.
- [42] JOHNSTON, D. E. et al. Cross-correlation weak lensing of sdss galaxy clusters ii: Cluster density profiles and the mass–richness relation. set. 2007.
- [43] ROZO, E. et al. Cosmological Constraints from the SDSS maxBCG Cluster Catalog. *Astrophys. J.*, v. 708, p. 645–660, fev. 2010.
- [44] SDSS. Disponível em: <<http://www.sdss.org/>>.
- [45] SDSS-III. Disponível em: <<http://www.sdss3.org/>>.
- [46] BAHCALL, N. A.; CEN, R. The mass function of clusters of galaxies. *Astrophys. J.*, v. 407, p. L49, 1993.
- [47] GIRARDI, M. et al. The observational mass function of nearby galaxy clusters. *Astrophys. J.*, v. 506, p. 45, 1998.
- [48] RAHMAN, N.; SHANDARIN, S. Constraining cosmological models by the cluster mass function. *Astrophys. J.*, v. 550, p. L121, 2001.
- [49] LUCCHIN, F.; MATARRESE, S. The effect of non-gaussian statistics on the mass multiplicity of cosmic structures. *Astrophys. J.*, v. 330, p. 535–544, 1988.
- [50] RIBEIRO, A. L. B.; WUENSCHÉ, C. A.; LETELIER, P. S. Small deviations from gaussianity and the galaxy cluster abundance evolution. *Astrophys. J.*, v. 539, p. 4, 2000.
- [51] GAZTAÑAGA, E.; LOBO, J. A. Nonlinear gravitational growth of large-scale structures inside and outside standard cosmology. *Astrophys. J.*, v. 548, p. 47–59, 2001.

- [52] COORAY, A.; SHETH, R. Halo models of large scale structure. *Phys.Rept.*, v. 372, p. 1–129, 2002.
- [53] LACEY, C.; COLE, S. Merger rates in hierarchical models of galaxy formation. *Mon. Not. R. Astron. Soc.*, v. 262, p. 627–649, 1993.
- [54] NAVARRO, J. Notas de aula da escola avançada de astrofísica do iag. In: . [S.l.: s.n.], 2000. Mangaratiba, Rio de Janeiro, 2 a 6 de setembro de 2000.
- [55] PRESS, W. H.; SCHECHTER, P. Formation of galaxies and clusters of galaxies by self-similar gravitational condensation. *Astrophysical Journal*, v. 187, p. 425–438, 1974.
- [56] GUNN, J. E.; III, J. R. G. On the infall of matter into clusters of galaxies and some effects on their evolution. *Astrophys. J.*, v. 176, p. 1, 1972.
- [57] EKE, V. R.; COLE, S.; FRENK, C. S. Cluster evolution as a diagnostic for omega. *MNRAS*, v. 282, p. 263–280, 1996.
- [58] ZHOU, W. et al. The largest virialized dark halo in the universe. *Int.J.Mod.Phys. D*, v. 11, p. 1189–1194, 2002.
- [59] BOND, J. R. et al. Excursion set mass functions for hierarchical Gaussian fluctuations. *Astrophys. J.*, v. 379, p. 440–460, 1991.
- [60] JEONG, D. Mass function and bias from halo model. <https://webpace.utexas.edu/dj955/www/notes/halomodel.pdf>. 2009.
- [61] POPOLO, A. D. On the cosmological mass function theory. *Astron.Rep.*, v. 51, p. 709–734, 2007.
- [62] MAKLER, M.; KODAMA, T.; CALVÃO, M. O. On local approximations to the nonlinear evolution of large-scale structure. *Astrophys. J.*, v. 557, p. 88–101, 2001.
- [63] BARDEEN, J. M. et al. The statistics of peaks of gaussian random fields. *Astrophys. J.*, v. 304, p. 15–61, 1986.
- [64] WEINBERG, S. *Cosmology*. [S.l.]: Oxford University Press, 2008.
- [65] TAYLOR, A. N.; HAMILTON, A. J. S. Non-linear cosmological power spectra in real and redshift space. *Mon. Not. R. Astron. Soc.*, v. 282, p. 767–778, 1996.
- [66] BUNN, E. F.; WHITE, M. The 4 year COBE normalization and large-scale structure. *Astrophys. J.*, v. 480, p. 6–21, 1997.
- [67] DODELSON, S. *Modern Cosmology*. [S.l.]: Academic Press, 2003.
- [68] SELJAK, U.; ZALDARRIAGA, M. A line-of-sight integration approach to cosmic microwave background anisotropies. *Astrophys. J.*, v. 469, p. 437, 1996.
- [69] ZALDARRIAGA, M.; SELJAK, U.; BERTSCHINGER, E. Integral solution for the microwave background anisotropies in nonflat universes. *Astrophys. J.*, v. 494, p. 491, 1998.

- [70] ZALDARRIAGA, M.; SELJAK, U. CMBFAST for spatially closed universes. *Astrophys. J. Suppl. Ser.*, v. 129, p. 431–434, 2000.
- [71] SUGIYAMA, N. Cosmic background anisotropies in cdm cosmology. *Astrophys. J. Suppl. Ser.*, v. 100, p. 281, 1995.
- [72] EISENSTEIN, D. J.; HU, W. Baryonic features in the matter transfer function. *Astrophys. J.*, v. 496, p. 605, 1998.
- [73] ALAM, U.; SAHNI, V.; STAROBINSKY, A. A. Reconstructing cosmological matter perturbations using standard candles and rulers. *Astrophys. J.*, v. 704, p. 1086–1097, 2009.
- [74] LINDER, E. V.; JENKINS, A. Cosmic structure growth and dark energy. *MNRAS*, v. 346, p. 573–583, 2003.
- [75] GROSSI, M.; SPRINGEL, V. The impact of early dark energy on non-linear structure formation. *MNRAS*, v. 394, p. 1559–1574, 2009.
- [76] UZAN, J. P. The acceleration of the universe and the physics behind it. *Gen. Rel. Grav.*, v. 39, p. 307–342, 2007.
- [77] NESSERIS, S.; PERIVOLAROPOULOS, L. Testing Λ CDM with the growth function $\delta(a)$: Current constraints. *Phys. Rev. D*, v. 77, p. 023504, 2008.
- [78] KOMATSU, E. et al. Five-year wilkinson microwave anisotropy probe (wmap) observations: Cosmological interpretation. *Astrophys. J. Suppl. Ser.*, v. 180, p. 330–376, mar. 2009.
- [79] EINASTO, J. et al. *Mon. Not. R. Astron. Soc.*, v. 206, p. 529, 1984.
- [80] DAVIS, M. et al. *Astrophys. J.*, v. 292, p. 371, 1985.
- [81] LACEY, C.; COLE, S. Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations. *Mon. Not. R. Astron. Soc.*, v. 271, p. 676, 1994.
- [82] REED, D. et al. Evolution of the mass function of dark matter haloes. *Mon. Not. R. Astron. Soc.*, v. 346, p. 565–572, 2003.
- [83] PIERPAOLI, E.; SCOTT, D.; WHITE, M. Power spectrum normalization from the local abundance of rich clusters of galaxies. *Mon. Not. R. Astron. Soc.*, v. 325, p. 77–88, 2001.
- [84] EVRARD, A. E. *Galaxy clusters as probes of cosmology and astrophysics*. [S.l.]: Cambridge University Press, 2004. 1-21 p. (Carnegie Observatories Astrophysics Series, v. 3). Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution.
- [85] LUKIC, Z. et al. The structure of halos: implications for group and cluster cosmology. *Astrophys. J.*, v. 692, p. 217–228, 2009.
- [86] CONNOLLY, A. J. et al. Slicing through multicolor space: Galaxy redshifts from broadband photometry. *Astronomical Journal*, v. 110, n. 6, p. 2655, dez. 1995.

- [87] WITTMAN, D.; RIECHERS, P.; MARGONINER, V. E. Photometric redshifts and photometry errors. *Astrophys. J.*, v. 671, p. L109–L112, 2007.
- [88] CYPRIANO, E. S. et al. Gemini and chandra observations of abell 586, a relaxed strong-lensing cluster. *Astrophys. J.*, v. 630, p. 38–49, 2005.
- [89] CYPRIANO, E. S. et al. Weak-lensing mass distributions for 24 x-ray abell clusters. *Astrophys. J.*, v. 613, p. 95–108, 2004.
- [90] COMERFORD, J. M.; NATARAJAN, P. The observed concentration-mass relation for galaxy clusters. *Mon. Not. R. Astron. Soc.*, v. 379, p. 190–200, 2007.
- [91] JOHNSTON, D. E. Cross-correlation lensing: Determining galaxy and cluster mass profiles from statistical weak-lensing measurements. *Astrophys. J.*, v. 656, p. 27–41, 2007.
- [92] RYDEN, B. *Introduction to Cosmology*. [S.l.]: Addison Wesley, 2003.
- [93] CHEVALLIER, M.; POLARSKI, D. Accelerating universes with scaling dark matter. *Int. J. Mod. Phys. D*, v. 10, p. 213–224, 2001.
- [94] LINDER, E. V. Exploring the expansion history of the universe. *Phys. Rev. Lett.*, v. 90, p. 091301, 2003.
- [95] DES. *Dark Energy Task Force*. Disponível em: <https://des.fnal.gov/survey_documents/DES-DETF/>.
- [96] CASH, W. Parameter estimation in astronomy through application of the likelihood ratio. *Astrophys. J.*, v. 228, p. 939–947, 1979.
- [97] HOLDER, G.; HAIMAN, Z.; MOHR, J. J. Constraints on Ω_m , Ω_Λ , and σ_8 from Galaxy Cluster Redshift Distributions. *Astrophys. J.*, v. 560, p. L111–L114, 2001.
- [98] HU, W.; KRAVTSOV, A. V. Sample variance considerations for cluster surveys. *Astrophys. J.*, v. 584, p. 702–715, 2003.
- [99] LIMA, M.; HU, W. Self-calibration of cluster dark energy studies: Counts in cells. *Phys.Rev. D*, v. 70, p. 043504, 2004.
- [100] PIERRE, M. et al. Precision cosmology with a wide area xmm cluster survey. set. 2010.
- [101] BARLOW, R. J. *Statistics: A guide to the use of statistical methods in the physical sciences*. [S.l.]: John Wiley & Sons, 1999. (The Manchester Physics Series).
- [102] COE, D. Fisher matrices and confidence ellipses: A quick-start guide and software. *eprint arXiv:0906.4123*.
- [103] ABAZAJIAN, K. N. et al. The seventh data release of the sloan digital sky survey. *Astrophys. J. Suppl. Ser.*, v. 182, p. 543–558, dez. 2009.
- [104] DR7. Disponível em: <<http://www.sdss.org/dr7/>>.
- [105] HAO, J. *Correspondência pessoal*.

- [106] LIMA, M.; HU, W. Photometric redshift requirements for self-calibration of cluster dark energy studies. *Phys. Rev. D*, v. 76, p. 123013, set. 2007.
- [107] CUNHA, C. Cross-calibration of cluster mass-observables. *Phys. Rev. D*, v. 79, p. 063009, 2009.
- [108] CUNHA, C. E.; EVRARD, A. E. Sensitivity of galaxy cluster dark energy constraints to halo modeling uncertainties. *Phys. Rev. D*, v. 81, p. 083509, 2010.
- [109] WU, H.-Y.; ZENTNER, A. R.; WECHSLER, R. H. The impact of theoretical uncertainties in the halo mass function and halo bias on precision cosmology. *Astrophys. J.*, v. 713, p. 856–864, out. 2010.
- [110] HU, W.; COHN, J. D. Likelihood methods for cluster dark energy surveys. *Physical Review D*, v. 73, p. 067301, 2006.
- [111] BOS, A. van den. *Parameter Estimation for Scientists and Engineers*. [S.l.]: John Wiley & Sons, 2007.
- [112] RODRÍGUEZ, G. *Generalized Linear Models*. Disponível em: <<http://data.princeton.edu/wws509/notes/a1s2.html>>.
- [113] HOEL, P. G. *Introduction to Mathematical Statistics*. [S.l.]: John Wiley & Sons, 1966.
- [114] COWAN, G. *Statistical Data Analysis*. [S.l.]: Oxford University Press, 1998.
- [115] CERN. *ROOT – A Data Analysis Framework*. Disponível em: <http://root.cern.ch/root/html/RooStats_LikelihoodInterval.html>.
- [116] BIERENS, H. J. *Introduction to the Mathematical and Statistical Foundations of Econometrics*. [S.l.]: Cambridge University Press, 2005.
- [117] HEESCH, D. van. *Doxygen: Source code documentation generator tool*. Disponível em: <<http://www.doxygen.org>>.
- [118] EGGERMONT, V. L. P. *Maximum Penalized Likelihood Estimation*. [S.l.]: Springer Science+Business Media, 2009. (Springer Series in Statistics, II: Regression).
- [119] CONTRIBUTORS, G. P. *GSL - GNU Scientific Library - GNU Project - Free Software Foundation (FSF)*. 2010. Disponível em: <<http://www.gnu.org/software/gsl/>>.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)