Universidade de São Paulo
Instituto de Astronomia, Geofísica e Ciências Atmosféricas
Departamento de Astronomia

Alexandre Henriques da Silva

Calibração da relação massa-riqueza de aglomerados de galáxias via lentes gravitacionais fracas

São Paulo
2020

Alexandre Henriques da Silva

Calibração da relação massa-riqueza de aglomerados de galáxias via lentes gravitacionais fracas

Trabalho de Conclusão de Curso apresentado ao Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo como requisito parcial para a obtenção do título de Bacharel em Astronomia.

Vertente: Pesquisa Básica
Orientador: Prof. Dr. Eduardo S. Cypriano

São Paulo
2020

Dedico este trabalho a Vanderlei Henriques, meu avô e amigo. Eu nunca te esquecerei.

Agradecimentos

Agradeço à minha família, que me apoiou para que eu me mantivesse na Universidade, em todos os sentidos;

À Rafaela Felix, o amor da minha vida, presente em todas as horas, nos momentos bons e ruins. Você despertou em mim o interesse pela vida acadêmica e nada do que construí seria possível sem você;

Ao meu Orientador, Eduardo Cypriano. Sua paciência e sabedoria foram fundamentais para realizar este trabalho;

Ao meu amigo e parceiro, Pedro Ticiani. Você proporcionou boas risadas mas também estudamos muito juntos. Agradeço aos demais colegas, sem exceção, que foram essenciais na minha trajetória acadêmica e nos momentos além da Universidade;

Agradeço à pequena Luna, minha apadrinhada. Seu sorriso puro de criança trouxe energia para superar os desafios que a vida acadêmica traz;

Às Professoras Jane e Vera, que guiaram com maestria a turma para obter excelência na formatação dos TG's;

À FAPESP, pelo apoio financeiro, sob o projeto n^o: 2019/13581-8;

Por fim, agradeço ao Instituto de Astronomia, Geofísica e Ciências Atmosféricas, para mim o ambiente mais aconchegante de toda a USP.

Esta tese/dissertação foi escrita em LaTeX com a classe IAGTESE, para teses e dissertações do IAG.

" *A ciência não é apenas compatível com a espiritualidade; é uma fonte profunda de espiritualidade.*"

Carl Sagan

"*Canto tanto minhas penas quanto minhas esperanças. Umas canto pras morenas, outras canto pras crianças. [...]*"

Dori Caymmi (Voz de Mágoa)

Resumo

O presente projeto tem por objetivo determinar a relação massa-observável de aglomerados de galáxias, sendo que o observável é o parâmetro de riqueza λ. Para isso, é necessário o emprego da técnica de lentes gravitacionais, com dados provenientes do levantamento *COnstrain Dark Energy with X-Ray* (CODEX). Frequentemente, esse tipo de calibração é feito usando a técnica de empilhamento, que utiliza dados combinados de diversos aglomerados para analisar um perfil médio destes. É proposta uma abordagem que evita esta discretização, baseado no ajuste simultâneo e individual de todos os sistemas através de técnicas bayesianas. Para isto, utiliza-se um perfil teórico de Navarro-Frenk-White (NFW), como exposto em Navarro et al. (1997) e Wright & Brainerd (2000), e um *prior* gaussiano para a concentração, presente em Dutton & Macciò (2014). Propõe-se uma calibração linear no espaço log, sendo $\log(M) = \log(M_0) + \alpha \log(\lambda/\lambda_0)$. A partir da escolha $\lambda_0 = 20$, foram obtidos $\log(M_0) = 14.36^{+0.21}_{-0.20}$ e $\alpha = 1.04^{+0.34}_{-0.34}$, mediante ajustes de todos os perfis de cisalhamento de 21 aglomerados de galáxias da amostra. Entretanto, não foi possível descartar a compatibilidade, ao nível de 3σ, entre calibrações realizadas através dos perfis de densidade superficial projetada e ajustes diretamente dos pares massa-riqueza. A proposta mostrou-se viável, embora haja a possibilidade de aperfeiçoar o modelo analítico pela inclusão de termos extras ao perfil NFW e a abordagem possa ser estendida para uma elevada quantidade de aglomerados.

Abstract

This project aims to determine the mass-observable relation of galaxy clusters, the observable being the richness parameter λ. For this, it is necessary to use the gravitational lens technique, with data from the *COnstrain Dark Energy with X-Ray* (CODEX) survey. This type of calibration is usually done using a stacking technique, which uses combined data from several clusters to analyze an average profile of these. An approach has been proposed that avoids this discretization, based on the simultaneous and individual adjustment of all systems through bayesian techniques. For this, a theoretical Navarro-Frenk-White (NFW) profile is used, as shown in Navarro et al. (1997) and Wright & Brainerd (2000), and a Gaussian *prior* for the concentration, present in Dutton & Macciò (2014). Linear calibration in the log space is proposed, with $\log(M) = \log(M_0) + \alpha \log(\lambda/\lambda_0)$. From the choice $\lambda_0 = 20$, we obtained $\log(M_0) = 14.36^{+0.21}_{-0.20}$ and $\alpha = 1.04^{+0.34}_{-0.34}$, by adjusting all the shear profiles of 21 galaxy clusters in the sample. However, it was not possible to rule out the compatibility, at the level of 3σ, between calibrations performed through the projected surface density profiles and adjustments directly from the mass-richness pairs. The proposal proved to be feasible, although there is a possibility to improve the analytical model by adding extra terms to the NFW profile and the approach can be extended to a high number of clusters.

Lista de Figuras

1.1 Regimes de Lentes Gravitacionais. 20

2.1 Configuração típica. 24
2.2 Cisalhamento tangencial 28
2.3 Eficiência de lentes em função do *redshift*. 29

5.1 Posição dos aglomerados da amostra. 40
5.2 Exemplo de ajuste NFW. 42
5.3 Posterior do exemplo de ajuste. 43

6.1 Parâmetro A. 48
6.2 Parâmetro B. 49
6.3 Posterior dos parâmetros A, B. 50
6.4 Calibração massa-riqueza por M2. 51
6.5 Calibração massa-riqueza por M1 e M2. 51

Lista de Tabelas

5.1 Propriedades dos aglomerados selecionados 40
5.2 Massa e concentração dos 21 aglomerados pelo método MCMC 44

6.1 Parâmetros linear e angular em função de λ_0 50

Sumário

1. Introdução 19
1.1 O fenômeno de lentes gravitacionais 19
1.2 Deflexão da luz na Relatividade Geral 20

2. Lentes Gravitacionais 23
2.1 Geometria de lentes gravitacionais 23
2.2 Elipticidade das fontes e o cisalhamento 26
2.3 Importância do redshift 29

3. Perfil radial de densidade 31
3.1 Lentes fracas e o perfil NFW 31

4. Técnicas bayesianas e o método MCMC 35
4.1 Inferência Bayesiana 35
4.2 Computação Bayesiana 37

5. Dados e metodologia 39
5.1 Levantamento CODEX e dados de riqueza 39
5.2 Massa e concentração pelo método MCMC 41

6. Resultados 47
6.1 Calibração massa-riqueza 47

7. Conclusões 53

Referências 55

Capítulo 1

Introdução

1.1 O fenômeno de lentes gravitacionais

A trajetória de um fóton proveniente de um objeto celeste é perturbada por inomogeneidades na distribuição de matéria. Em analogia com um feixe de luz observado após uma mudança no índice de refração do meio, a imagem de uma galáxia ou aglomerado pode sofrer mudanças na posição aparente ou até mesmo magnificações por um objeto massivo ao longo da linha de visada. A partir do formalismo da Teoria da Relatividade Geral (TRG), o fenômeno é bem explicado considerando a luz imersa em um campo gravitacional e sujeita a efeitos de maré locais.

É neste cenário que surge a técnica de lentes gravitacionais: um tratamento adequado da geometria e de dados observacionais pode levar a estimativas importantes sobre a massa projetada.

Na Figura 1.1 são apresentados os regimes de lentes gravitacionais forte e fraco. Próximo à cruz verde domina o regime de lentes fortes caracterizado pela formação de imagens múltiplas e arcos. Estes eventos estão restritos a regiões internas e são mais raros. Em grandes raios, torna-se mais comum o regime de lentes fracas, exemplificado no quadro em destaque. Neste caso, a distorção amplifica as fontes na direção tangencial à linha que as liga ao centro da lente e só pode ser observada tomando-se médias sobre várias fontes.

No destaque da Figura 1.1, a linha em azul indica a orientação média das galáxias de fundo, enquanto a linha vermelha exibe a orientação introduzida pelo cisalhamento (*shear*). A diferença nas direções de ambas as linhas se deve à distribuição de orientações e de elipticidade das galáxias. Como o sinal de cisalhamento depende tanto da distribuição quanto da massa total do defletor, aglomerados de galáxias são ótimos objetos para serem

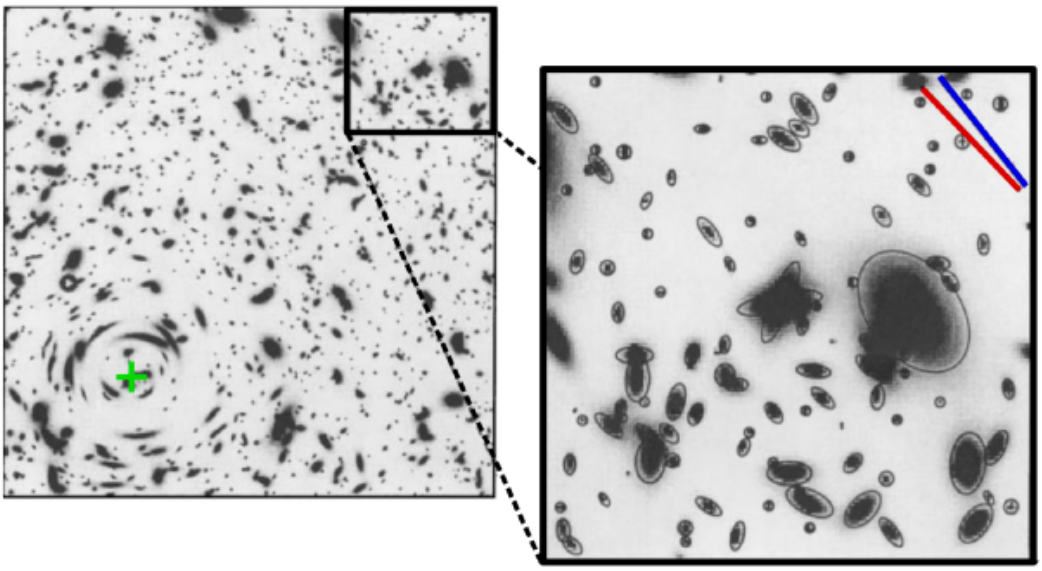

Figura 1.1: Ilustração dos dois regimes de lentes gravitacionais. À esquerda foi modelado um aglomerado de galáxias a *redshift* $z = 0.15$ e fontes a *redshift* $z = 1$. Na região central (borda inferior esquerda) arcos tangenciais e radiais são facilmente identificáveis. O aumento da distância radial a este centro é acompanhada de diminuição do sinal de cisalhamento. Do lado direito está destacada a região onde predomina o efeito de lentes fracas: a elipticidade introduzida pelo cisalhamento é menor que a elipticidade intrínseca das galáxias. Extraída de Mellier (1999).

estudados pela técnica.

Lentes gravitacionais constituem um método de determinação de massa independente da luminosidade e atuam como "telescópios cósmicos", magnificando fontes distantes e fracas, conforme Cibirka (2017). Além disso, a dependência das distâncias angulares com parâmetros cosmológicos como H_o, Ω_M e Ω_λ pode fornecer restrições importantes para a cosmologia, de acordo com Mellier (1999).

1.2 Deflexão da luz na Relatividade Geral

Partindo da mecânica clássica e tratando a luz como uma partícula, muitos físicos fizeram suposições de que a luz poderia sofrer interações gravitacionais. Cita-se frequentemente a suposição de Laplace de que poderia haver um corpo tão massivo que seu campo gravitacional impediria que a luz escapasse.

Em cartas trocadas entre John Mitchel e Johaan von Soldner no final dos anos 1700 e início dos anos 1800, menciona-se a possibilidade de que a luz seja defletida (ao pas-

sar próximo de um corpo pontual de massa M, como uma estrela) de um ângulo $\hat{a}_N = 2GM/(c^2\xi)$, como apontado em Schneider (2006).

A formulação da TRG forneceu uma base sólida para o estudo do comportamento da luz em um campo gravitacional. Em um artigo publicado em 1911, Einstein recalculou resultados obtidos por Mitchell e Soldner para o ângulo de deflexão, encontrando um valor diferente do "newtoniano" por um fator 2:

$$\hat{a} = \frac{4GM}{c^2\xi}. \tag{1.1}$$

Um corpo massivo como o Sol pode defletir a luz de estrelas próximas à sua superfície em relação à linha de visada e mudar a posição aparente delas, mas estas só são observadas na ocorrência de um eclipse solar total. A observação do eclipse solar em Sobral-CE no ano de 1919 com precisão suficiente permitiu confirmar o valor obtido por Einstein e promover fortemente sua teoria.

Capítulo 2

Lentes Gravitacionais

2.1 Geometria de lentes gravitacionais

Considere a situação apontada na Figura 2.1. Um corpo extenso contido no plano da lente é colocado próximo à linha de visada de uma fonte S, supondo que a extensão da distribuição de massa ao longo dessa linha é bem menor que as demais distâncias envolvidas. D_s, D_l e D_{ls} indicam, respectivamente, as distâncias observador-fonte, observador-lente e lente-fonte.

Denotando vetores em negrito, o ângulo $\boldsymbol{\beta}$ indica a posição real da fonte dado que S é visto na posição observada $\boldsymbol{\theta}$, já que a lente redireciona os feixes de luz. A trajetória da luz é defletida por um ângulo $\hat{\boldsymbol{\alpha}}$. Definindo o ângulo de deflexão reduzido $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}\frac{D_{ls}}{D_s}$, através da geometria temos

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{2.1}$$

que é a conhecida equação de lentes.

Observe que a Equação (2.1) pode admitir mais de uma solução. Podemos interpretar fisicamente esta característica no contexto de lente gravitacionais fortes, onde são observadas múltiplas imagens da mesma fonte.

A densidade superficial de massa correspondente à projeção da distribuição de massa no plano da lente é escrita como

$$\Sigma(\boldsymbol{\xi}) = \int \rho(\boldsymbol{\xi}, z)dz, \tag{2.2}$$

onde $\boldsymbol{\xi}$ é o vetor no plano da lente e z é a coordenada ortogonal ao plano. A deflexão pode ser calculada como a superposição das deflexões causadas por elementos infinitesimais de

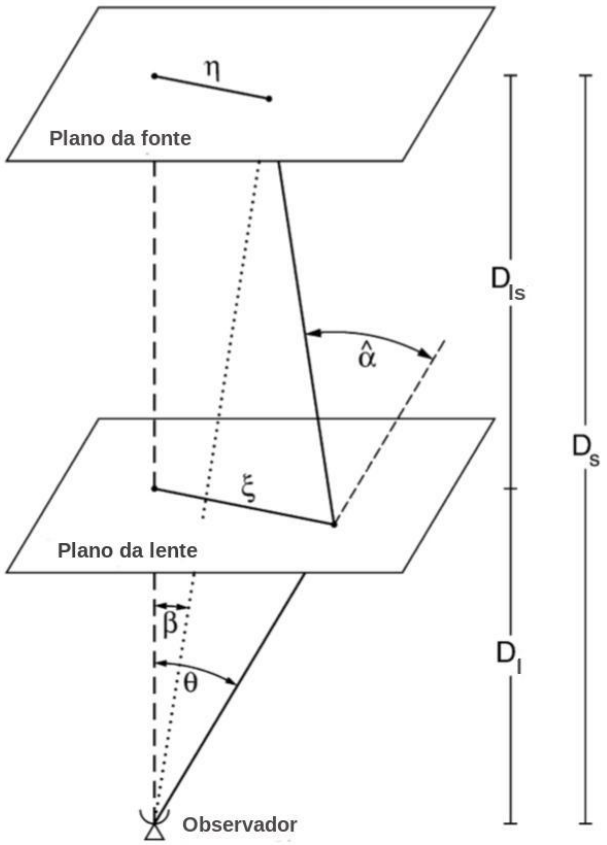

Figura 2.1: Esquema típico da configuração de lentes gravitacionais. Adaptado de Bartelmann & Schneider (2001).

massa distribuídas ao longo do plano

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int \Sigma(\boldsymbol{\xi}) \frac{\boldsymbol{\xi} - \boldsymbol{\xi'}}{|\boldsymbol{\xi} - \boldsymbol{\xi'}|^2} d^2\xi'. \tag{2.3}$$

Mas o lado direito da Equação (2.3) pode ser reescrito se levarmos em conta que o potencial gravitacional escalar projetado,

$$\psi(\boldsymbol{\xi}) = 2G \int \Sigma(\boldsymbol{\xi}) \ln(\boldsymbol{\xi} - \boldsymbol{\xi'}) d^2\xi', \tag{2.4}$$

satisfaz a Equação de Poisson

$$\nabla_\xi^2 \psi(\boldsymbol{\xi}) = 4\pi G \Sigma(\boldsymbol{\xi}), \tag{2.5}$$

de forma que

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{2}{c^2}\nabla_{\xi}^2\psi(\boldsymbol{\xi}). \tag{2.6}$$

Para o caso de uma lente com simetria esférica, podemos escolher a origem das coordenadas para coincidir com o centro de simetria, com $\boldsymbol{\xi}$ o vetor que aponta do centro da distribuição à imagem (veja Figura 2.1). Assim, obtemos

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2}\frac{M(\xi)}{\xi}, \tag{2.7}$$

onde $M(\xi)$ é a massa dentro desse raio

$$M(\xi) = 2\pi\int_0^{\xi}\Sigma(\boldsymbol{\xi}')\xi' d\xi'. \tag{2.8}$$

Supondo que a densidade de massa superficial é isotrópica, o ângulo de deflexão reduzido pode ser escrito como função da posição

$$\alpha(\theta) = \frac{4\pi G\Sigma}{c^2}\frac{D_{ls}D_l}{D_s}\theta, \tag{2.9}$$

pois $\xi = D_l\theta$. A proporcionalidade entre α e θ pode ser melhor avaliada ao definirmos

$$\Sigma_{crit} = \frac{c^2 D_s}{4\pi G D_{ls} D_l}, \tag{2.10}$$

a densidade crítica, que depende unicamente da geometria do sistema. Alternativamente, por incluir as distâncias, pode-se entender a densidade crítica dependente do *redshift* z. Dessa forma, define-se uma densidade superficial adimensional como a razão entre a densidade superficial projetada e a crítica

$$\kappa(\theta) = \frac{\Sigma(\theta)}{\Sigma_{crit}}. \tag{2.11}$$

A razão κ é chamada de convergência e atua como um indicador do regime em que o sistema se encontra. A situação de lente forte equivale a dizer que a convergência excedeu a unidade, i.e., $\kappa > 1$ (e consequentemente $\Sigma > \Sigma_{crit}$). Situações em que $\kappa << 1$ marcam o domínio de lentes fracas.

A densidade superficial de massa admite um potencial gravitacional escalar Ψ que satisfaz a equação de Poisson e tal que $2\kappa = \nabla^2\Psi$. Desta forma, a matriz de distorção escreve-se

$$\mathcal{A} = \delta_{ij} - \frac{\partial^2\psi}{\partial x_i \partial x_j} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1+\kappa-\gamma_1 \end{pmatrix} = (1-\kappa)\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}, \quad (2.12)$$

onde introduzimos o cisalhamento complexo $\gamma = \gamma_1 + i\gamma_2$ e definimos o cisalhamento reduzido, $g_i = \gamma_i/(1-\kappa)$, sendo que o segundo expressa a distorção anisotrópica da fonte. Relaciona-se o cisalhamento com o potencial de deflexão via

$$\gamma_1 = \frac{1}{2}\left(\frac{\partial^2\psi}{\partial x_1^2} - \frac{\partial^2\psi}{\partial x_2^2}\right);\ \gamma_2 = \frac{\partial^2\psi}{\partial x_1 \partial x_2}. \quad (2.13)$$

O efeito do tensor de magnificação $\mathcal{M} = \mathcal{A}^{-1}$ é transformar uma fonte circular em uma fonte elíptica cuja razão entre os eixos é aproximadamente $(1-|g|)/(1+|g|)$ e rotacionada pelo ângulo $\tan^{-1}(g_2/g_1)/2$. A magnificação escalar é a razão entre o fluxo observado e o fluxo da fonte, calculada por

$$\mu = \frac{1}{det\mathcal{A}} = \frac{1}{(1-\kappa)^2(1-|g|^2)}. \quad (2.14)$$

Uma característica importante de lentes gravitacionais é que o brilho superficial é conservado. Por outro lado, percebe-se que a magnificação altera apenas formas e tamanhos angulares observados. Neste sentido, uma fonte não resolvida pode ter seu brilho amplificado, tornando-se detectável.

Mesmo que o cisalhamento e a convergência sejam efeitos observáveis, eles diferem em essência. Por um lado, a convergência é simplesmente calculada pela Equação (2.11) em determinado raio, enquanto γ não tem uma interpretação tão simples. Efetivamente, o valor de γ depende da distribuição de massa em todo o espaço e seu valor esperado é de maior interesse, como será visto na Seção 2.2.

2.2 Elipticidade das fontes e o cisalhamento

As galáxias de fundo não são perfeitamente circulares: elas possuem uma elipticidade intrínseca. Se o semi-eixo maior é a e o menor é b, escrevemos $\epsilon^{(I)} = \frac{a-b}{a+b}$. Decorre que observa-se uma combinação linear de sua elipticidade intrínseca e do cisalhamento reduzido na forma

$$\epsilon^{(O)} = \epsilon^{(I)} + \frac{\gamma}{1-\kappa}. \quad (2.15)$$

Em geral, $\epsilon^{(I)}$ é desconhecido para uma galáxia específica e a Equação (2.15) só é uma boa estimativa se o cisalhamento superar a elipticidade intrínseca. Portanto, a estratégia é supor que as galáxias de fundo em uma região estão distribuídas em orientações aleatórias, de modo que $\left\langle \epsilon^{(I)} \right\rangle = 0$ e

$$\left\langle \epsilon^{(O)} \right\rangle = \left\langle \frac{\gamma}{1-\kappa} \right\rangle \ (\approx g, \ \text{ se } \kappa << 1), \tag{2.16}$$

com os *brakets* indicando média sobre muitas fontes. Fica claro que tal abordagem é estatística. Contudo, o principal limitador da precisão com que podemos medir γ é o ruído de forma (*shape noise*), que pode ser estimado como a largura da distribuição de elipticidade das fontes, σ_e.

Na prática, alguns autores (Hoekstra et al. (2000) e Leauthaud et al. (2007)) apontam que as galáxias tem uma elipticidade média de 0.25 por componente, sendo $e_1 = \epsilon^{(I)} \cos(2\varphi)$ e $e_2 = \epsilon^{(I)} \sin(2\varphi)$ as componentes cartesianas da elipticidade, onde φ é o ângulo de posição na Figura 2.2. Neste sentido, a incerteza estatística na medida de γ_i seria aproximadamente

$$\sigma_{\gamma_i} = \frac{\sqrt{\langle e_i^2 \rangle}}{\sqrt{N}} \approx \frac{0.25}{\sqrt{N}}, \tag{2.17}$$

sendo N é o número de galáxias sobre a qual foi tomada a média. Em primeira análise a incerteza poderia diminuir conforme aumenta-se N, mas N é limitado tanto pela profundidade quanto pela área da observação.

Existem ainda fenômenos observacionais que estão presentes, como a *point spread function* (que sucintamente descreve a transformação de uma imagem pontual em uma imagem estendida) introduzida pela atmosfera terrestre e limitações do aparato experimental, como telescópios e lentes. Felizmente, eles contribuem levemente para a incerteza na medida da elipticidade e existem maneiras de contorná-los.

Como já discutido, o efeito de lentes fracas tende a introduzir um alinhamento tangencial coerente nas imagens em relação ao centro da lente. Se o ângulo azimutal entre o eixo horizontal e o vetor que liga o centro da lente ao centro da imagem é ϕ (veja Figura 2.2), podemos definir uma distorção tangencial média como

$$\gamma_T = -(\gamma_1 \cos(2\phi) + \gamma_2 \sin(2\phi)), \tag{2.18}$$

e uma distorção cruzada

$$\gamma_\times = \gamma_1 \sin(2\phi) - \gamma_2 \cos(2\phi). \tag{2.19}$$

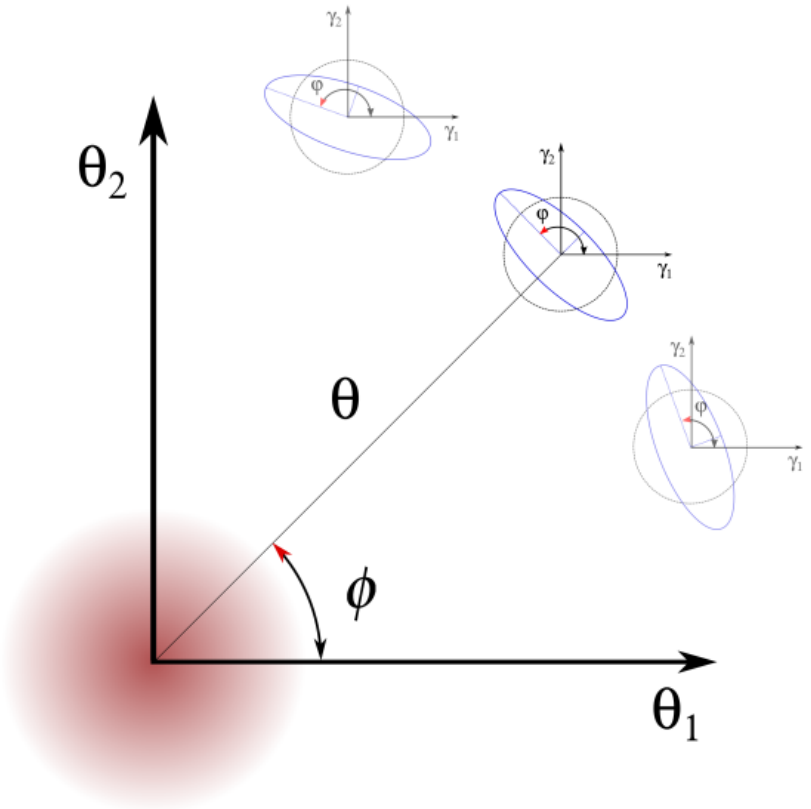

Figura 2.2: Configuração do cisalhamento tangencial γ_T e demais parâmetros. θ é a distância da fonte (galáxia circular) à origem do sistema de coordenadas, que coincide com o centro do aglomerado, em vermelho. Em azul é indicado o remapeamento da matriz $\mathcal{M}$ sobre as fontes. Extraída de Oliveira (2016).

Se $\gamma_T > 0$, diz-se que a galáxia está tangencialmente ou radialmente orientada; caso contrário, ao passo que $\gamma_\times$ aparecerá sempre deslocado de $\pi/4$ em relação a γ_T. Fisicamente, o cisalhamento reduzido induzido por uma distribuição de massa produz somente elipticidades tangenciais de modo que $\gamma_\times = 0$, de acordo com Oliveira (2016). O trabalho de Miralda-Escude (1991) identifica que é possível relacionar esta distorção tangencial média com um contraste de massa

$$\langle \gamma_T \rangle = \frac{\overline{\Sigma}(< r) - \overline{\Sigma}(r)}{\Sigma_{crit}} = \frac{\Delta\overline{\Sigma}(r)}{\Sigma_{crit}} = \overline{\kappa}(< r) - \overline{\kappa}(r), \tag{2.20}$$

ou seja, a distorção média induzida sobre as formas das fontes à distância r do centro do aglomerado é proporcional à diferença entre a densidade de massa superficial interior a

r e a densidade superficial de massa à distância r. Como estamos considerando apenas a distorção tangencial, identifica-se a densidade superficial diferencial como $\Delta\Sigma_+$. Em particular, observa-se que uma superfície de densidade constante não induz efeito de cisalhamento.

2.3 Importância do redshift

A Equação (2.10), que relaciona a propriedades físicas com as propriedades da lente, tem uma forte dependência com as distâncias envolvidas. Se definirmos o parâmetro $\beta = \frac{D_{ls}}{D_s}$, notamos que o sinal de lente gravitacional decai com o aumento do *redshift*. A Figura 2.3 compara o valor médio de β para dois aglomerados a *redshift* $z = 0.2$ e $z = 0.8$. Nota-se que os valores assintóticos para β são maiores para o objeto a menor *redshift*.

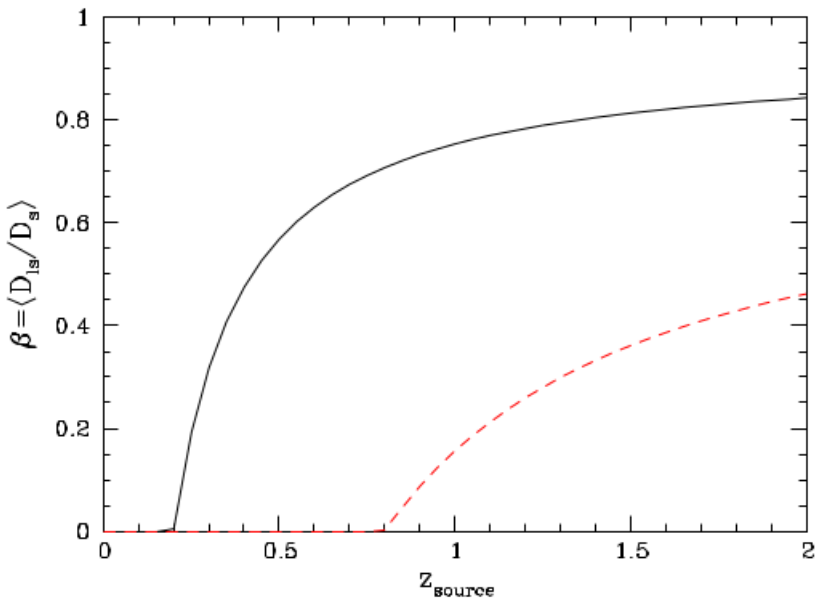

Figura 2.3: Parâmetro β em função do *redshift* para dois aglomerados a $z = 0.2$ (em preto) e a $z = 0.8$ (em vermelho). Extraída de Hoekstra et al. (2013).

Primeiramente, deve-se notar que muitas fontes que seriam adequadas para medidas de lentes fracas envolvem um custo elevado na determinação de seu *redshift* a partir de técnicas espectroscópicas, dado que se encontram em grandes quantidades e posuem baixo fluxo. Além disso, galáxias membro do aglomerado podem diluir o sinal de cisalhamento, contribuindo para a subestimativa da massa. Esta contaminação pode alcançar frações consideráveis nas regiões mais internas, onde a densidade numérica de galáxias é maior.

Dessa forma, pode-se levar em conta estatisticamente a diluição do sinal[1] ou realizar um refinamento do catálogo a partir de algum critério, como exclusão de galáxias do aglomerado em um diagrama cor-magnitude, dentre outras abordagens.

[1] O "*boost factor correction*", por exemplo, visa corrigir o efeito de diluição na forma $(1+n_{non-bkg}/n_{bkg})$ supondo que as galáxias de fundo são radial e uniformemente distribuídas, como discutido por Pratt et al. (2019). O uso de *redshift* fotométrico de alta qualidade, como o do J-PAS, seria uma alternativa viável.

Capítulo 3

Perfil radial de densidade

3.1 Lentes fracas e o perfil NFW

A massa e estrutura interna de aglomerados de galáxias refletem as propriedades das perturbações de densidade primordiais e a natureza da matéria escura, de acordo com Pratt et al. (2019). Os perfis de densidade radiais podem ser relacionados com o efeito de lentes gravitacionais, como apontado pela Equação (2.20).

A partir de simulações numéricas de N-corpos, foi identificado um perfil de massa que poderia descrever halos de matéria escura, que recebeu o nome de perfil Navarro-Frenk-White (NFW), como exposto em Navarro et al. (1997):

$$\rho_{NFW}(r) = \frac{\delta_c \rho_c}{\left(\frac{r}{r_s}\right)\left(1+\frac{r}{r_s}\right)^2}, \tag{3.1}$$

onde definimos um raio de escala $r_s = R_\Delta/c$ que indica o ponto em que a inclinação do perfil muda e c é o parâmetro de concentração, que mensura o quão concentrada a matéria está em direção ao centro.

Os efeitos de c são visíveis nas partes centrais e externas do aglomerado, que são regiões onde o sinal de cisalhamento é fraco. Na região central, a pequena quantidade de galáxias para obter medidas resulta em sinal muito ruidoso. Já afastado do centro o sinal é fraco pois a intensidade decai com a distância, além de halos dos aglomerados próximos contaminarem as medidas. Em todas as regiões ainda há o impacto da diluição do *shear* por conta das galáxias pertencentes ao aglomerado.

A densidade de escala é definida como

$$\delta_c = \frac{\Delta}{3}\frac{c^3}{\ln(1+c) - c/(1+c)}. \tag{3.2}$$

Além disso, ρ_c é a densidade crítica do Universo do aglomerado em questão

$$\rho_c(z) = \frac{3H^2(z)}{8\pi G}, \tag{3.3}$$

pois $H(z)$ é o parâmetro de Hubble no *redshift* z. O raio do virial, R_Δ, é definido como o raio em que a densidade em seu interior é Δ vezes a densidade crítica, correspondente à uma massa

$$M_\Delta = M(R_\Delta) = \frac{4\pi}{3}\rho_c R_\Delta^3. \tag{3.4}$$

O trabalho de Bryan & Norman (1998) fornece a expressão $\Delta_{vir} = (18\pi^2 + 82\xi - 39\xi^2)/\Omega(z)$, com $\xi = \Omega(z) - 1$. Assim, tomando $\Omega(z) = 1$, adotamos $\Delta = 200$, de forma que M_{200} é aproximadamente a massa do virial num universo Einstein-de Sitter.

Expressões analíticas para a densidade superficial projetada e as relações com o cisalhamento e convergência foram desenvolvidas por Wright & Brainerd (2000). Por conveniência, escolhendo $x = r/r_s$ um raio adimensional, podemos escrever

$$\Sigma_{NFW}(x) = \begin{cases} \frac{2r_s\delta_c\rho_c}{(x^2-1)}\left[1 - \frac{2}{\sqrt{1-x^2}}\operatorname{arctanh}\sqrt{\frac{1-x}{1+x}}\right], & \text{se } x < 1 \\ \frac{2r_s\delta_c\rho_c}{3}, & \text{se } x = 1\,, \\ \frac{2r_s\delta_c\rho_c}{(x^2-1)}\left[1 - \frac{2}{\sqrt{x^2-1}}\arctan\sqrt{\frac{x-1}{1+x}}\right], & \text{se } x > 1 \end{cases} \tag{3.5}$$

enquanto a densidade média num círculo de raio x é

$$\overline{\Sigma}_{NFW}(x) = \begin{cases} \frac{4r_s\delta_c\rho_c}{x^2}\left[\frac{2}{\sqrt{1-x^2}}\operatorname{arctanh}\sqrt{\frac{1-x}{1+x}} + \ln(\frac{x}{2})\right], & \text{se } x < 1 \\ 4r_s\delta_c\rho_c\left[1 + \ln\frac{1}{2}\right], & \text{se } x = 1\,. \\ \frac{4r_s\delta_c\rho_c}{x^2}\left[\frac{2}{\sqrt{x^2-1}}\arctan\sqrt{\frac{x-1}{1+x}} + \ln(\frac{x}{2})\right], & \text{se } x > 1 \end{cases} \tag{3.6}$$

As Equações (3.5) e (3.6) podem ser combinadas para reescrever a Equação (2.11)

$$\kappa_{NFW}(x) = \frac{\Sigma(x)}{\Sigma_{crit}} \tag{3.7}$$

e a Equação (2.20)

$$\gamma_{NFW} = \frac{\overline{\Sigma}_{NFW}(x) - \Sigma_{NFW}(x)}{\Sigma_{crit}}. \tag{3.8}$$

Portanto, ao ajustar um perfil NFW, $\gamma\Sigma_{crit}$, para a densidade de massa da lente é possível obter M_{200} para o aglomerado em questão, embora o modelo possa ser aperfeiçoado.

Tratando-se de um perfil radial, decorre que seria necessário identificar corretamente o seu centro: surge, então, a necessidade de incluir um termo que leve em conta um possível erro na determinação do centro do aglomerado, o chamado erro de centragem.

Note também que lentes gravitacionais são sensíveis à toda matéria ao longo da linha de visada, de modo que qualquer estrutura em larga escala (*large scale structure*, LSS) contribui introduzindo uma alteração na forma das galáxias, conhecido como *cosmic shear*. Ou seja, a massa total medida seria a soma da massa do aglomerado com a massa de qualquer estrutura não relacionada, como apontado em Hoekstra et al. (2013). Além disso, aponta-se que um ajuste simples NFW poderia ter as incertezas dobradas nos parâmetros obtidos (massa e concentração), tratando apenas de erros estatísticos.

Por fim, o processo de formação de estruturas no Universo tende a reunir os aglomerados de galáxias. Assim, o modelo NFW simples pode ser discrepante dos dados para raios grandes, indicando que existe algum halo nas vizinhança contribuindo para o cisalhamento. Com uma dificuldade mais elevada, a correção surge no chamado termo *2-halo*, como discutido em Mandelbaum et al. (2005).

Capítulo 4

Técnicas bayesianas e o método MCMC

4.1 Inferência Bayesiana

Em muitas situações é desejável conhecer o valor de um conjunto de parâmetros $\boldsymbol{\theta} = \{\theta_1, \theta_2, \cdots, \theta, \cdots\}$ de um modelo teórico. Embora não seja possível determiná-los com exatidão, a Estatística se baseia no princípio de restringir o desconhecimento acerca destes parâmetros. Diferentes abordagens para um mesmo problema podem levar a diferentes resultados para o conjunto $\boldsymbol{\theta}$, tratando-se do modelo escolhido ou da inferência adotada.

Ocupam lugar de destaque as inferências frequentista e bayesiana. Na primeira, cada parâmetro θ é fixo durante o estudo e os dados são realizações aleatórias de uma amostra - indicando a escolha do nome "frequência" - de modo que a melhor estimativa para θ é tal que maximiza a probabilidade de todo o conjunto de dados. Na segunda abordagem, os parâmetros $\boldsymbol{\theta}$ são tratados como distribuições de probabilidades, mantendo os dados observados fixos durante o estudo.

Em outras palavras, ao observar um conjunto $\mathbf{D}$ de N dados, para a inferência frequentista θ é estimado ao maximizar a *verossimilhança*, que é o produto das probabilidades individuais dos dados $\mathcal{L} = \prod_{i=1}^{N} P(D_i|\theta)$, quando os dados são independentes entre si. Quando os dados do experimento são obtidos de uma distribuição Gaussiana, a log-verossimilhança se reduz ao usual χ^2.

Para a inferência bayesiana, dada a probabilidade de um conjunto de dados ser descrito por certo parâmetro $P(D_i|\theta)$ e um conhecimento prévio dos parâmetros $P(\theta)$, a distribuição de probabilidades do parâmetro θ é dada por

$$P(\theta|\mathbf{D}) = \frac{P(\mathbf{D}|\theta)P(\theta)}{P(\mathbf{D})}, \tag{4.1}$$

que é a regra de Bayes para a informação em θ. Para um conjunto de dados fixo vale $P(\mathbf{D}|\theta) = \mathcal{L}$, indicando a plausabilidade de que os dados sejam descritos pelo modelo com parâmetro θ e $P(\theta)$ é a distribuição *a priori* do parâmetro.

Note que o *prior* não depende dos dados do experimento, sendo apenas probabilidades dos parâmetros em relação ao modelo estudado. Podem ser originados de teorias ou dados de experimentos anteriores, mas nunca devem depender do estudo ao qual ele faz parte.

Priores normalmente dividem-se em duas classes: informativos e não-informativos. Apesar do nome, ambos carregam informação, mas em quantidade diferente. Um *prior* não-informativo pode ser exemplificado por uma distribuição uniforme para o parâmetro, $P(\theta) = \text{Unif}\,(\theta_1, \theta_2) = I_{\theta_1 \leq \theta \leq \theta_2}$, onde I é a função indicadora e qualquer valor no intervalo é equiprovável. Em certas circunstâncias, também é possível que um *prior* plano seja informativo .

Uma escolha de *prior* informativo seria $P(\theta) = \mathcal{N}(\mu, \sigma^2)$, em que a realização da variável θ é bem descrita por uma distribuição normal de média μ e variância σ^2 conhecidas. Este *prior* restringe fortemente a crença sobre o valor de θ, já que a Gaussiana favorece quantidades próximas à média.

O termo no denominador da Equação (4.1), chamado de *evidência dos dados*, é calculado como $P(\mathbf{D}) = \int P(\mathbf{D}|\theta)P(\theta)d\theta$ e atua como uma constante normalizadora, de forma que as probabilidades sempre somem-se a um. Os três termos no lado direito da Equação (4.1), mediante a regra de Bayes, produzem o *posterior* de θ ou distribuição *a posteriori*.

É conveniente substituir a igualdade na Equação (4.1) por um sinal de proporcionalidade para omitir o termo $P(\mathbf{D})$ do denominador. A forma $P(\theta|\mathbf{D}) \propto P(\mathbf{D}|\theta)P(\theta)$ é particularmente útil em situações que envolvam a determinação de parâmetros de modelos analíticos, pois a evidência dos dados apenas normaliza a distribuição *a posteriori* e pode ser muito difícil de calcular na prática. A evidência é importante, por exemplo, na seleção de modelos, i.e., qual dos M modelos existentes melhor descreve o fenômeno de interesse.

Trotta (2017) aponta algumas vantagens para a abordagem bayesiana. Em muitos casos é comum a presença de *nuisance parameters*[1] aos dados e a inferência frequentista não possui uma forma eficiente de lidar com eles. A teoria bayesiana permite calcular a distribuição dos parâmetros de interesse e marginalizar sobre os *nuisance parameters*.

Deve-se ressaltar que na maioria dos problemas um *prior* é importante e, segundo

[1] Parâmetros que não são o objetivo do estudo mas devem ser levados em conta na análise.

Trotta (2017), sua omissão pode resultar em estimativas incorretas. O presente projeto, por exemplo, carece de um *prior* quanto à massa do aglomerado, restringindo os valores fisicamente aceitáveis. Note também que a prática frequentista supõe distribuições de dados que podem não ter sido obtidos, enquanto a teoria bayesiana utiliza apenas dados que foram realmente observados.

Pode ser mostrado que a escolha de *priores* impróprios (não normalizáveis) em certas situações leva aos mesmos resultados obtidos pelo método frequentista, desde que o posterior não seja impróprio. Ainda que resultados frequentistas possam ser recuperados a partir da inferência bayesiana em situações simples, esta continua sendo uma ferramenta que permite resolver situações mais complexas.

4.2 Computação Bayesiana

O posterior dos parâmetros pode ser bastante complexo, não possuindo expressões analíticas, além de conter muita informação. Neste sentido, existem diversas formas de sumarizá-lo computacionalmente e ocupa lugar de destaque o método *Markov Chain Monte Carlo* (MCMC).

MCMC baseia-se na amostragem do posterior da distribuição de interesse a partir da comparação entre modelo e dados. Na prática, a estimativa de parâmetros por esta ferramenta possui incertezas mais robustas - pouco sensíveis à presença de dados " contaminantes" - permite estimativas de grande quantidade de parâmetros e análise mais simples de covariâncias entre eles, além de lidar com os *nuisance parameters* facilmente.

Uma das implementações de maior destaque no meio científico foi desenvolvida por Foreman-Mackey et al. (2013) e é a biblioteca EMCEE na linguagem PYTHON, adotada neste trabalho.

É definida pelo usuário uma função que calcula o modelo com os n_{dim} parâmetros livres. Deve-se então estabelecer um conjunto de N *walkers*, comportados numa matriz de dimensão $N \times n_{dim}$ contendo propostas para os valores iniciais dos parâmetros. Tais propostas podem ser geradas aleatoriamente, desde que dentro dos limites estabelecidos pelos *priores* do problema.

Cada *walker*, em uma iteração, explora o espaço de parâmetros através de um "passo", ou seja, há uma atualização do parâmetro de $\theta_0 \rightarrow \theta$. Este novo parâmetro é avaliado

quanto a possibilidade de melhor conciliar os dados e o modelo teórico por meio de um teste semelhante ao usual χ^2, já que supomos que os erros são gaussianos. Formalmente, esta exploração ocorre por meio das cadeias de Markov, processos estocásticos em que a proposta futura da variável depende apenas do valor atual.

O MCMC utiliza funções-proposta para gerar valores dos parâmetros e aceitá-los quando a probabilidade destes descreverem a amostra aumentar. Em outras palavras, se a proposta θ produz um modelo mais compatível com a amostra, este valor é armazenado e o processo é repetido para cada *walker* por n_{step} iterações.

Ao final do processo, o produto é a distribuição (amostrada) do posterior , conhecida como "cadeia". Ao longo do processo, os N *walkers* armazenam os valores aceitos para os parâmetros e o posterior para cada conjunto de parâmetros.

Alguns fatores têm de ser levados em conta nesta abordagem: é comum que nas execuções sejam atribuídas probabilidades muito próximas de zero (da ordem da precisão da máquina) ou que realizem-se cálculos com números grandes (tão grandes quanto o maior número representado pelo computador), acarretando na perda de precisão e comprometendo os resultados. Desta maneira, é conveniente trabalhar com o logaritmo natural das distribuições, já que os mesmos parâmetros que maximizam as distribuições maximizam seus logaritmos. É fácil perceber que, ao tomar o logaritmo natural nos dois lados da Equação (4.1), obtém-se $\ln P(\theta|\mathbf{D}) \propto \ln P(\mathbf{D}|\theta) + \ln P(\theta)$, que além de resolver a questão numérica simplifica o problema ao substituir produtos por somas.

Note também que problemas multi-modais podem apresentar convergência espúria ou não-convergência. Deve-se levar em consideração também a convergência da cadeia em relação à forma que o método é construído. Por iniciarem em diversos pontos, algumas cadeias podem apresentar lentidão (em relação a n_{step}) para convergir para regiões de maior densidade. Assim, é comum que seja adotado um *burn-in*, que consiste no descarte das iterações iniciais visando remover a região que pode influenciar negativamente na análise estatística.

Capítulo 5

Dados e metodologia

5.1 Levantamento CODEX e dados de riqueza

De acordo com Cibirka (2017), o levantamento CODEX contém uma amostra de aglomerados de galáxias selecionados com base em observações no óptico e raios-X por apresentarem um excesso de fótons de 4σ na área de superposição do *ROSAT All Sky Survey* e do *Sloan Digital Sky Survey*.

Kiiveri et al. (2020) aponta que cada fonte detectada é considerada um candidato a aglomerado e o algoritmo *redMaPPer* - como descrito em Rykoff et al. (2014) - reporta, nas regiões de superdensidade, estimativas do *redshift* fotométrico, riqueza, posição refinada e fluxo em raios-X.

De acordo com Rykoff et al. (2014), se $\mathbf{x}$ é um vetor que carrega propriedades de uma galáxia (por exemplo, magnitude, cor ou posição), a distribuição projetada dentro e ao redor de um aglomerado pode ser modelada como $S(\mathbf{x}, \lambda) = \lambda u(\mathbf{x}, \lambda) + b(\mathbf{x})$, sendo $u(\mathbf{x}, \lambda)$ o perfil de densidade do aglomerado e $b(\mathbf{x})$ a densidade de fundo, correspondendo às galáxias não pertencentes ao aglomerado.

Desta forma, a riqueza λ define a quantidade de galáxias do aglomerado. A probabilidade de que uma galáxia pertença ao aglomerado é dada por $p(\mathbf{x}) = \lambda u(\mathbf{x}, \lambda)/(\lambda u(\mathbf{x}, \lambda) + b(\mathbf{x}))$. Além disso, o número total de galáxias deve satisfazer a relação $\lambda = \sum p(\mathbf{x}, \lambda) = \sum_{r<R_c} \lambda u(\mathbf{x}, \lambda)/(\lambda u(\mathbf{x}, \lambda) + b(\mathbf{x}))$ para algum raio *cut-off* $R_c = R_c(\lambda)$. Em resumo, Rykoff et al. (2014) aponta que a determinação da riqueza, que envolve a resolução numérica destas equações e demais melhorias, foi o objetivo por trás do desenvolvimento do *redMaPPer*.

A Figura 5.1 exibe as coordenadas dos 21 aglomerados do levantamento CODEX utilizados neste projeto. Estes aglomerados encontram-se na amostra primária, exposta no

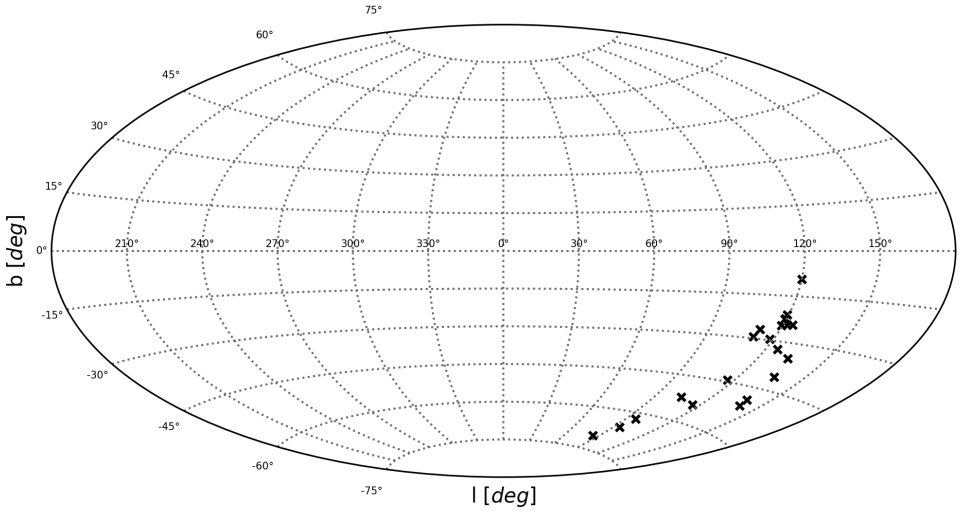

Figura 5.1: Posições, em coordenadas galácticas, dos 21 aglomerados escolhidos para a calibração massa-riqueza.

trabalho de Cibirka (2017), consistindo de aglomerados de alta riqueza e *redshift* intermediário ($\lambda > 60$ e $z > 0.4$). Cibirka (2017) realiza um empilhamento dos aglomerados para determinação de suas concentrações pela técnica de lentes gravitacionais fracas.

Cibirka (2017) destaca que a sub-amostra de lentes fracas foi construída com 36 aglomerados da amostra primária. A Tabela 5.1 sintetiza as principais características dos aglomerados estudados, combinados dos trabalhos de Cibirka (2017) e Kiiveri et al. (2020).

Tabela 5.1 - Propriedades dos aglomerados selecionados. Os *redshifts* e riquezas referem-se aos determinados pelo levantamento CFHT.

ID	R.A	Dec.	z	λ	σ_λ
12451	08:04:38	53:25:24	0.583	46	3
24865	08:22:45	41:28:09	0.487	55	2
24872	08:25:59	40:15:19	0.407	116	4
24877	08:24:40	40:06:53	0.593	71	4
24981	08:56:14	37:55:52	0.407	107	3
25424	11:31:01	38:24:42	0.510	69	3
25953	14:03:42	38:27:38	0.478	88	3
27940	00:20:10	34:53:36	0.463	89	3
27974	00:08:55	32:11:12	0.469	75	3
29284	08:03:30	33:02:06	0.541	68	3
35361	14:56:13	30:21:12	0.411	98	3
36818	22:20:21	27:21:00	0.559	64	3

Continua na próxima página...

Tabela 5.1 - Continuação

ID	R.A	Dec.	z	λ	σ_λ
41843	23:40:45	20:53:02	0.436	75	3
43403	08:10:20	18:15:13	0.423	94	3
46649	01:35:17	08:48:14	0.587	128	5
47981	08:40:02	08:38:04	0.54	69	3
50492	23:16:46	12:47:12	0.525	105	3
50514	23:32:14	10:35:32	0.475	73	3
55181	00:45:10	-01:51:49	0.532	97	4
59915	01:25:01	-05:31:53	0.472	98	3
64232	00:42:32	-11:04:07	0.553	66	3

5.2 Massa e concentração pelo método MCMC

Em princípio, o ajuste NFW pode ser realizado sobre apenas um aglomerado, mas em situações de baixa massa e devido a incertezas como o *shape noise*, pode ser muito difícil obter dados com alta resolução. De acordo com Cibirka (2017), empilhar os dados contribui por: *i)* aumentar a razão sinal-ruído à medida que multiplicamos o número de fontes pela quantidade de aglomerados da pilha; *ii)* diminuir os efeitos de estruturas em larga escala e da triaxilidade, que geram erros sistemáticos aos dados e *iii)* permite estender o estudo para aglomerados de baixas massas[1].

Embora eficiente, a abordagem culmina na análise de um perfil médio sobre os aglomerados relacionados. É proposto evitar o empilhamento, onde cada aglomerado possui um conjunto de parâmetros individuais e as análises subsequentes serão simultâneas.

A fim de estabelecer uma relação entre a riqueza e a massa de aglomerados, é necessário tratar os dados disponibilizados pelo *survey* escolhido. Neste sentido, o primeiro passo foi desenvolver uma rotina em PYTHON capaz de realizar a leitura do conjunto de dados de interesse e implementar as equações apresentadas nas Seções 2.1 e 2.2.

Adotando, no plano do céu, o centro do aglomerado com coordenadas[2] (x_c, y_c) e coordenadas genéricas (x, y), cada galáxia i é indexada com (x_i, y_i), além das componentes da

[1] Da ordem de $10^{14} M_\odot$, em acordo com Hoekstra et al. (2013).

[2] É conveniente adotar coordenadas polares com centro coincidente com o centro do aglomerado.

elipticidade[3] e do *redshift*. Então, calcula-se $\Delta\Sigma_+$[4] a partir das Equações (2.10) e (2.20). Para ilustrar como é feito o ajuste NFW, o perfil de densidade superficial diferencial para o campo *CODEX 12451*, do *Canada France Hawaii Telescope* - CFHT, é exibido na Figura 5.2.

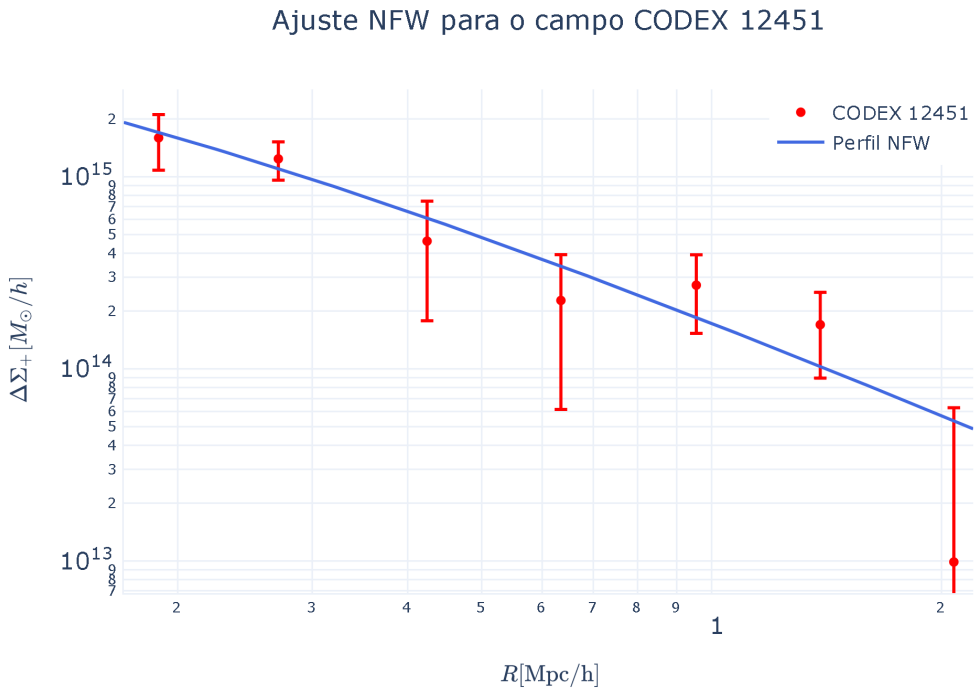

Figura 5.2: Exemplo de ajuste de perfil NFW, sem correções extras, a uma amostra a partir da redução de dados.

Os parâmetros M_{200} e c que ajustam o perfil NFW na Figura 5.2 são escolhidos como os que maximizam o posterior das distribuições, sendo este construído pelo algoritmo EMCEE, já discutido na Seção 4.2. A Figura 5.3 exibe o espaço de parâmetros para o campo *CODEX 12451*.

Fica evidente, na Figura 5.3, que não é possível, com os dados utilizados neste trabalho, determinar a concentração a partir do posterior de forma apropriada, já que não houve convergência para este parâmetro (observe o posterior da concentração no lado direito da Figura 5.3) e não faz sentido calcular estatísticas associadas. Os valores de c são limitados em cerca de $c = 500$ para visualização, embora valores tão altos não tenham sentido físico.

[3] O sinal de cisalhamento pode ser tomado como a média para um subintervalo radial.

[4] O sinal + indica o uso do cisalhamento tangencial.

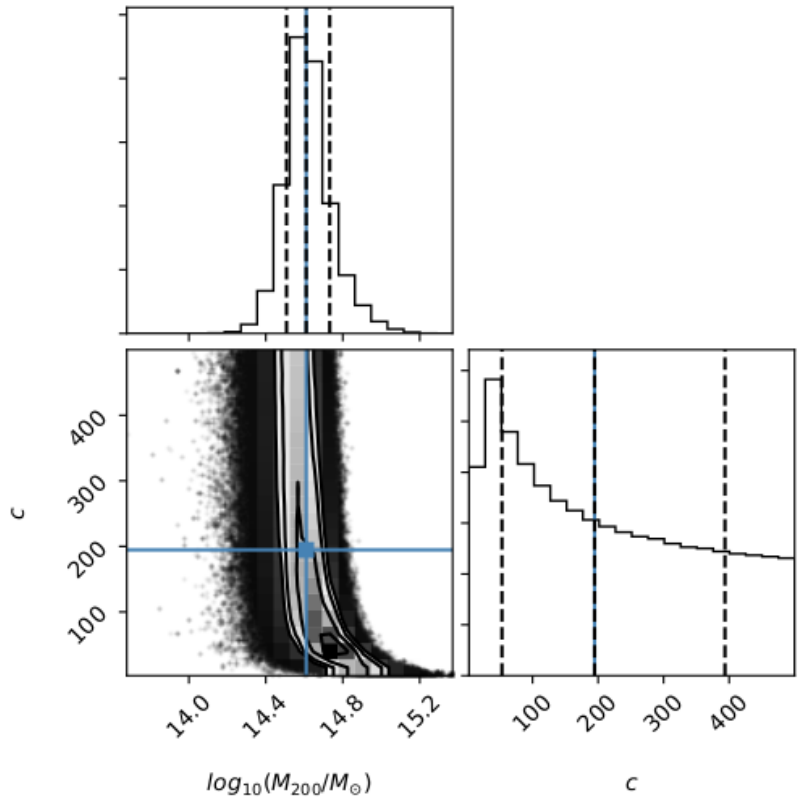

Figura 5.3: Espaço de parâmetros para o campo estudado. As linhas verticais tracejadas indicam a média e os intervalos de confiança de 95%.

Hoekstra et al. (2013) argumentam que, observacionalmente, a concentração não pode ser bem determinada para aglomerados individuais apenas usando lentes fracas ou fortes. Por outro lado, estudos - Duffy et al. (2008), Dutton & Macciò (2014) - apontam que os parâmetros c e M_{200} não são independentes, mas sim que $c = c(M_{200}, z)$, apresentando expressões analíticas para tal quantidade.

Para o estudo subsequente, optou-se por utilizar a relação exposta em Dutton & Macciò (2014) para a concentração em função da massa e *redshift*, obtida empiricamente através de simulações numéricas, que é a lei de potência

$$\log_{10} c = a + b \log_{10} \left(M / \left[10^{12} h^{-1} M_\odot \right] \right), \tag{5.1}$$

sendo $a = 0.520 + (0.905 - 0.520) \exp\left(-0.617 z^{1.21}\right)$ e $b = -0.101 + 0.026z$.

Desta forma, a concentração é relacionada com a massa e *redshift*. Na prática, c deixa de ser um parâmetro livre e o método MCMC pode ser empregado apenas para a variável M, adotando um *prior* do tipo gaussiano para a concentração. A Tabela 5.2 exibe o logaritmo

da massa e concentração a partir do método MCMC aplicado aos perfis de cisalhamento individuais.

Tabela 5.2 - Concentração sob o *prior* de Dutton & Macciò (2014) e log-massa dos 21 aglomerados do estudo, determinados via método MCMC a partir dos perfis de densidade individuais.

ID	$\log(M_{200}/M_\odot)$	c
12451	$15.10^{+0.17}_{-0.14}$	$2.99^{+0.08}_{-0.08}$
24865	$14.83^{+0.21}_{-0.17}$	$3.22^{+0.10}_{-0.12}$
24872	$14.96^{+0.16}_{-0.14}$	$3.20^{+0.09}_{-0.10}$
24877	$15.39^{+0.12}_{-0.11}$	$2.88^{+0.07}_{-0.07}$
24981	$14.9^{+0.17}_{-0.14}$	$3.18^{+0.10}_{-0.10}$
25424	$14.33^{+0.41}_{-0.27}$	$3.39^{+0.15}_{-0.23}$
25953	$14.83^{+0.15}_{-0.13}$	$3.17^{+0.08}_{-0.08}$
27940	$15.03^{+0.14}_{-0.12}$	$3.12^{+0.07}_{-0.07}$
27974	$14.93^{+0.13}_{-0.11}$	$3.12^{+0.07}_{-0.08}$
29284	$15.78^{+0.07}_{-0.06}$	$2.79^{+0.05}_{-0.05}$
35361	$14.87^{+0.12}_{-0.10}$	$3.14^{+0.07}_{-0.07}$
36818	$14.82^{+0.12}_{-0.10}$	$3.10^{+0.08}_{-0.09}$
41843	$14.82^{+0.16}_{-0.13}$	$3.23^{+0.08}_{-0.09}$
43403	$14.83^{+0.14}_{-0.12}$	$3.20^{+0.07}_{-0.08}$
46649	$15.25^{+0.12}_{-0.10}$	$2.96^{+0.06}_{-0.06}$
47981	$15.09^{+0.23}_{-0.18}$	$3.04^{+0.10}_{-0.12}$
50492	$15.39^{+0.11}_{-0.10}$	$2.88^{+0.06}_{-0.06}$
50514	$14.69^{+0.18}_{-0.15}$	$3.27^{+0.09}_{-0.10}$
55181	$14.82^{+0.18}_{-0.16}$	$3.07^{+0.07}_{-0.08}$
59915	$15.08^{+0.09}_{-0.09}$	$3.00^{+0.06}_{-0.06}$
64232	$14.70^{+0.20}_{-0.18}$	$3.29^{+0.11}_{-0.12}$

Na Equação (4.1) θ é o parâmetro de interesse e na notação utilizada trata-se de uma variável única e real. Se há um conjunto de parâmetros a serem determinados, o posterior seria dado por

$$P(\boldsymbol{\theta}|\boldsymbol{D}) = P(\theta_1, \theta_2, \cdots, \theta_n|\boldsymbol{D}). \tag{5.2}$$

Como é de interesse a *probability density function* (PDF) de cada parâmetro, $P(\theta_i|\boldsymbol{D})$, marginaliza-se sobre os demais parâmetros, que refere-se à integração sobre todos os valores

possíveis. A regra do produto fornece que $P(A, B) = P(A|B)P(B)$, onde reconhecemos o formato do termo chamado *evidência* na regra de Bayes.

Desta forma, ainda que ocorra "degenerescência" no valor de c e não se utilize a relação da Equação (5.1), o posterior da massa não apresenta divergências já que o código EMCEE fornece os posteriores dos parâmetros individualmente e, consequentemente, marginalizados. Deve-se levar em conta, entretanto, que a existência de uma covariância entre M e c faz com que o aumento de uma das variáveis acarrete na diminuição da outra, especificamente na região $c < 100$ da Figura 5.3 . Isso se reflete na determinação da massa dos aglomerados de toda a amostra: os valores relativamente altos de concentração, da ordem de $c \gtrsim 10$, tendem a subestimar as massas dos aglomerados.

Capítulo 6

Resultados

6.1 Calibração massa-riqueza

Propõe-se que a massa relaciona-se com a riqueza por meio de uma relação linear (no espaço $\log_{10}$) da forma

$$\log M_{200} = \log M_0 + \alpha \log \lambda/\lambda_0 = A + B \log \frac{\lambda}{\lambda_0}, \tag{6.1}$$

onde λ_0 é um termo livre com valor tipicamente entre 20 e 40.

A etapa seguinte, que seria natural à primeira vista, para obter a calibração massa-riqueza seria colecionar as massas dos aglomerados, como exposto na Seção 5.2, e aplicar mais uma vez o método MCMC, agora no espaço $M \times \lambda$, para os parâmetros A e B. Esta abordagem será nomeada como Método 1 - **M1**.

Ocorre que as massas determinadas para os aglomerados podem apresentar vieses e grandes erros estatísticos, devido a todas as considerações feitas até o momento e por conta dos perfis de cisalhamento serem altamente ruidosos pela natureza dos dados de lentes fracas.

Uma forma de contornar este problema consiste em utilizar diretamente os perfis NFW na calibração, que será nomeada de Método 2 - **M2**. Os dois métodos possuem uma diferença conceitual. Em **M1**, a verossimilhança é dada pelo χ^2 entre a massa determinada através do MCMC e a massa obtida da Equação (6.1). Já a verossimilhança do **M2** é descrita pelo χ^2 entre os dados de cisalhamento e os perfis NFW obtidos com as massas da Equação (6.1) e a concentração dada pela Equação (5.1).

Desta forma, no Método 2, a verossimilhança é a soma[1] dos χ^2 do ajuste para cada

[1] Note que o uso do logaritmo faz com que produtos tornem-se somas.

aglomerado, enquanto pelo Método 1 a verossimilhança é a soma dos χ^2 entre a massa de cada aglomerado, estimada via MCMC, e a massa obtida da Equação 6.1.

Especificamente para **M2**, é utilizado um *prior* gaussiano para a concentração, com média dada pela relação de Dutton & Macciò (2014), Equação 5.1. Como já discutido, nestas situações não é necessário calcular as *evidências*, já que atuam apenas como constantes normalizadoras das distribuições *a posteriori*.

Usando as propriedades do logaritmo, podemos reescrever a Equação 6.1 para obter

$$\log M_{200} = A + B \log \lambda - B \log \lambda_0 = A' + B \log \lambda, \tag{6.2}$$

onde $A' = A - B \log \lambda_0$. Como λ_0 é um termo livre, os ajustes não devem depender deste parâmetro: o termo mais à direita na Equação 6.2 tem a vantagem de explicitar este fato, pois A' é o termo linear.

Para verificar que de fato A' é constante em relação a λ_0, escolhemos os valores de λ_0 como: 20, 25, 30, 35 e 40. Os gráficos nas Figuras 6.1 e 6.2 exibem, respectivamente a independência do parâmetro A' e B com λ_0, sendo as barras horizontais as médias obtidas. As flutuações em torno da média são naturais deste tipo de estudo, já que tratam-se de processos estocásticos (por envolverem o uso de cadeias de Markov).

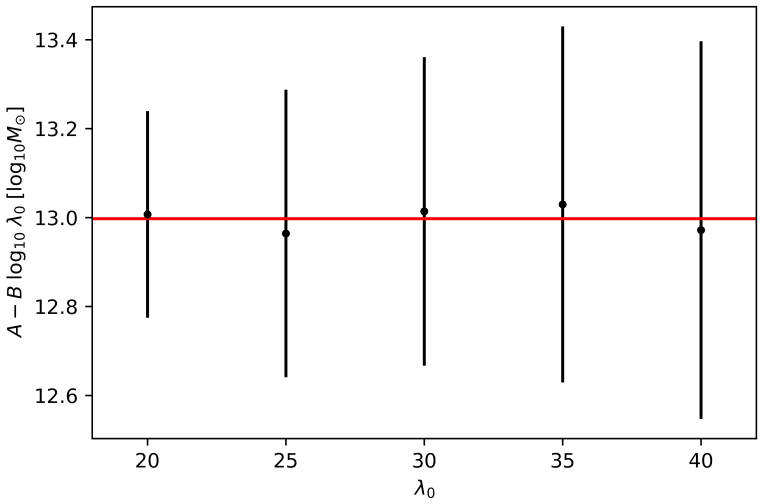

Figura 6.1: Independência do parâmetro $A' = A - B \log_{10} \lambda_0$ em relação a λ_0. A barra horizontal indica a média dos parâmetros obtidos.

A Tabela 6.1 indica a calibração obtida, pelo método **M2**, para as escolhas de λ_0.

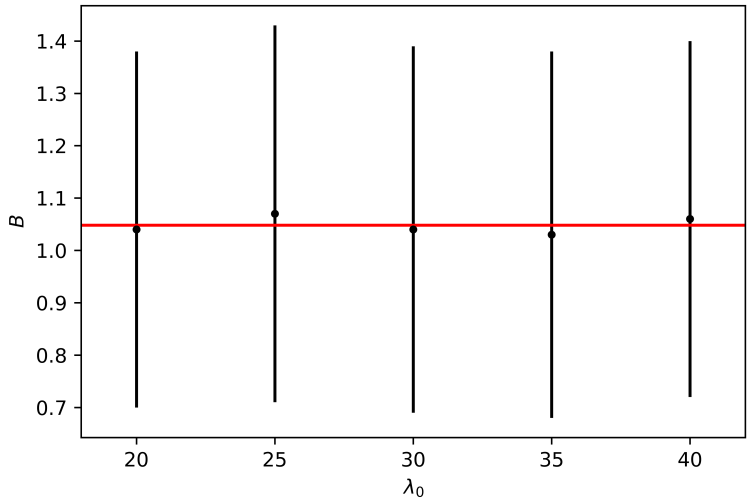

Figura 6.2: Independência do parâmetro B em relação a λ_0. A barra horizontal indica a média dos parâmetros obtidos.

Como a calibração não depende de λ_0, escolheremos o resultado para $\lambda_0 = 20$ para a análise subsequente.

A Figura 6.3 exibe o posterior amostrado para a escolha $\lambda_0 = 20$. Neste caso, houve convergência para ambos os parâmetros, diferentemente do caso massa-concentração. É interessante notar que há grande degenerescência no par A e B, devido à grande excentricidade das elipses de confiança. A cruz azulada indica os valores mais prováveis, enquanto as demais linhas tracejadas indicam os intervalos de confiança individuais. Posteriores semelhantes deram origem às calibrações exibidas na Tabela 6.1.

Foi determinada a massa dos aglomerados como discutido na Seção 5.2, mas por consistência adotando o *prior* da Equação 5.1 para a concentração. O resultado é exibido na Figura 6.4.

A curva mais escura da Figura 6.4 representa o ajuste mais provável, em termo dos parâmetros. Além disso, as curvas mais claras indicam 100 realizações, escolhidas aleatoriamente, da cadeia armazenada. Enquanto as incertezas exibidas na Tabela 6.1 representam as incertezas nos parâmetros, as realizações aleatórias atuam como estimadores da incerteza no modelo escolhido, ou seja, a relação dada pela Equação 6.1.

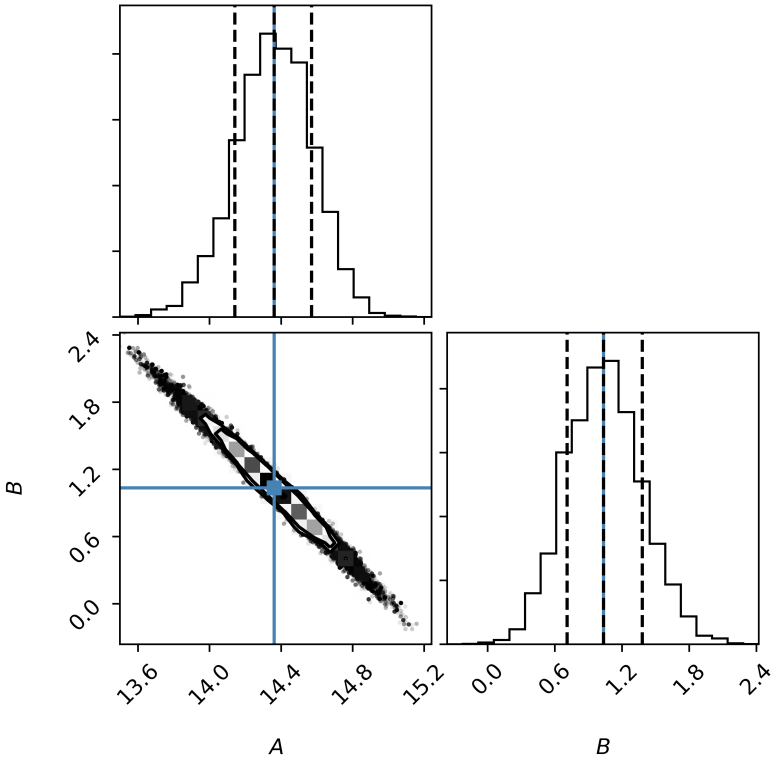

Figura 6.3: Posterior dos parâmetros amostrado via EMCEE, segundo a escolha $\lambda_0 = 20$.

Tabela 6.1 - Parâmetros A e B determinados pelo método MCMC para as escolhas de λ_0. As incertezas correspondem ao intervalo de confiança de 95%.

λ_0	A	B
20	$14.36^{+0.21}_{-0.20}$	$1.04^{+0.34}_{-0.34}$
25	$14.46^{+0.18}_{-0.17}$	$1.07^{+0.36}_{-0.36}$
30	$14.55^{+0.17}_{-0.15}$	$1.04^{+0.35}_{-0.33}$
35	$14.62^{+0.14}_{-0.13}$	$1.03^{+0.34}_{-0.36}$
40	$14.67^{+0.12}_{-0.12}$	$1.06^{+0.35}_{-0.34}$

Até o momento as calibrações foram realizadas pelo método **M2**, utilizando-se dos perfis de cisalhamento determinados. Para investigar se a proposta do método é realmente mais robusta, pode-se calibrar a relação diretamente das massas determinadas para os aglomerados. São exibidas na Figura 6.5 as calibrações massa-riqueza pelo método **M1** e **M2**.

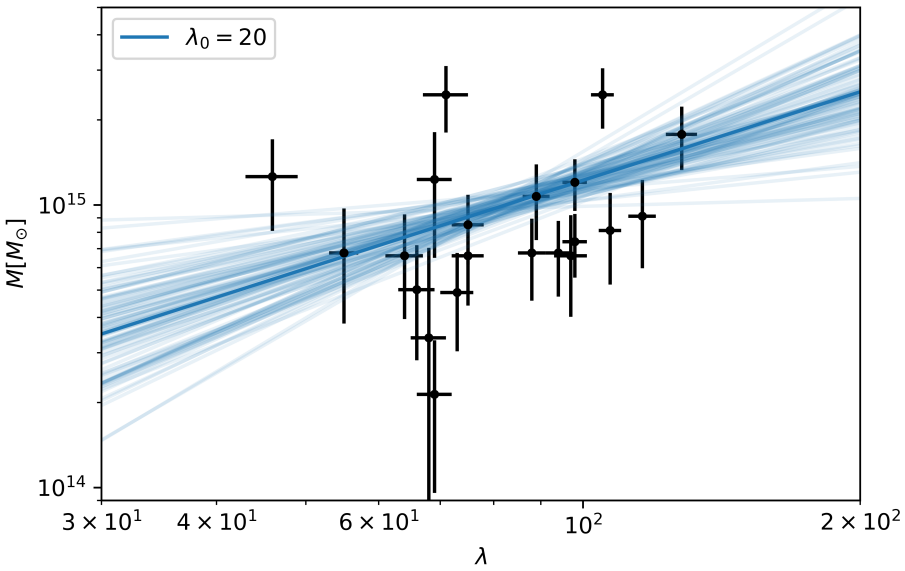

Figura 6.4: Calibração da relação massa-riqueza. As curvas em azul mais claro representam 100 realizações aleatórias do MCMC. Os dados em preto representam, para cada aglomerado, sua massa determinada com o *prior* da Equação 5.1 e a riqueza determinada pelo levantamento CFHT, expostos na Tabela 5.1. Note que as barras de erro não são simétricas por conta da escala logarítmica.

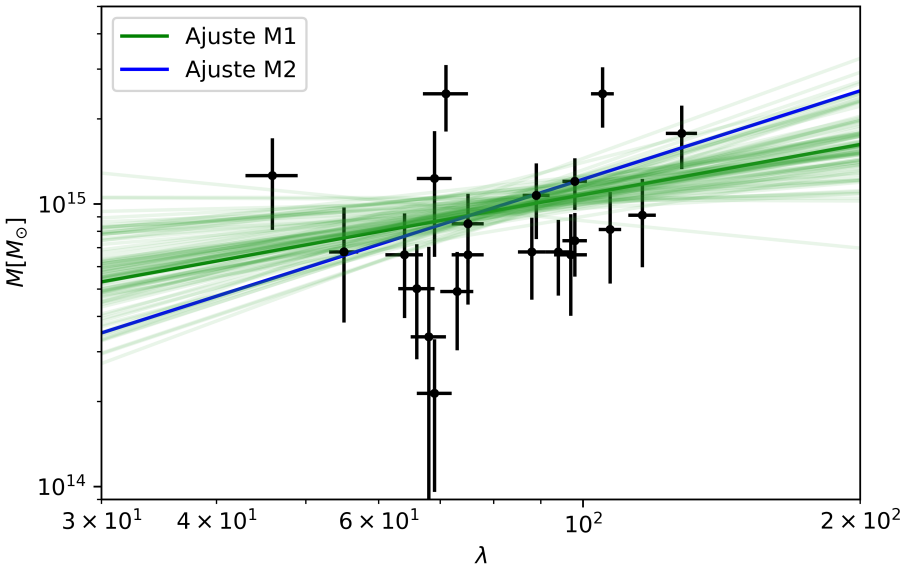

Figura 6.5: Calibração da relação massa-riqueza pelos dois métodos. As curvas em verde claro representam 100 realizações aleatórias do MCMC pelo método **M1**, enquanto a curva em azul é curva *best-fit* do método **M2**, já apresentada na Figura 6.4. Omitimos as realizações aleatórias do método **M2** para não sobrecarregar a visualização.

Capítulo 7

Conclusões

Deve-se ressaltar que a calibração via **M2** não depende explicitamente das massas dos aglomerados e, consequentemente, não deve ser afetada pela incertezas nas suas determinações apenas com o *prior* para a concentração e é esperado que esta seja uma calibração mais robusta. Em outras palavras, o ajuste em azul na Figura 6.4 não tem relação direta com os dados em preto, sendo obtido apenas dos perfis de cisalhamento.

Entretanto, é importante notar que, pela Figura 6.4, que os dois métodos não são totalmente incompatíveis, sendo o intervalo de 3σ delimitado pelas curvas em verde. Em contrapartida, isto não significa que são equivalentes e poderia se optar pelo método *M1*, dado que estamos utilizando um perfil teórico NFW sem termos extras, que podem afetar consideravelmente a estimativa da massa do aglomerado.

Por fim, a implementação do método foi bem-sucedida. É satisfatório que a proposta tenha mostrado potencial, por evitar o uso habitual do *stacking* para os perfis de cisalhamento e incorporar toda a amostra disponível.

Uma possível e natural continuação deste projeto seria a inclusão dos termos extras ao perfil analítico NFW, como os que contabilizam as incertezas na determinação do centro do aglomerado e a contribuição de halos vizinhos ao sinal de cisalhamento. Além disso, o esquema bayesiano adotado pode ser melhorado, com ênfase em performance computacional, podendo ser usado em *clusters high performance computing* e devem ser implementadas melhorias relacionadas à importação de dados - extremamente necessárias quando a abordagem do projeto for executada para um grande número de aglomerados de galáxias.

Referências

Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291

Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

Cibirka N. d. O., 2017, PhD thesis, Universidade de São Paulo

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, MNRAS, 390, 64

Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Hoekstra H., Franx M., Kuijken K., 2000, The Astrophysical Journal, 532, 88

Hoekstra H., Bartelmann M., Dahle H., Israel H., Limousin M., Meneghetti M., 2013, Space Sci. Rev., 177, 75

Kiiveri K., et al., 2020, MNRAS, artigo submetido para publicação.

Leauthaud A., et al., 2007, The Astrophysical Journal Supplement Series, 172, 219–238

Mandelbaum R., et al., 2005, Monthly Notices of the Royal Astronomical Society, 361, 1287–1322

Mellier Y., 1999, ARA&A, 37, 127

Miralda-Escude J., 1991, ApJ, 370, 1

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Oliveira R. M., 2016, PhD thesis, Universidade de São Paulo

Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, Space Sci. Rev., 215, 25

Rykoff E. S., et al., 2014, The Astrophysical Journal, 785, 104

Schneider P., 2006, Gravitational Lensing: Strong, Weak and Micro, p. 269–451

Trotta R., 2017, Bayesian Methods in Cosmology (arXiv:1701.01467)

Wright C. O., Brainerd T. G., 2000, ApJ, 534, 34