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

Diana Sena Soares

Um estudo de bojos de galáxias: relação entre morfologia e população estelar

São Paulo
2015

Diana Sena Soares

Um estudo de bojos de galáxias: relação entre morfologia e população estelar

Monografia final do Trabalho de Conclusão de Curso apresentada ao Departamento de Astronomia do Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo como requisito parcial para conclusão do bacharelado em Astronomia.

Área de Concentração: Astronomia
Orientador(a): Prof.ª Dr.ª Paula R. T. Coelho

São Paulo
2015

À Auzerina,

"*Toda a nossa ciência, comparada com a realidade, é primitiva e infantil - e, no entanto, é a coisa mais preciosa que temos.*"

Albert Einsten

"*A normalidade é uma rua pavimentada: conveniente para caminhar, mas nela não cresce nenhuma flor.*"

Vincent Van Gogh

Resumo

Visualmente, bojos de galáxias aparentam ser estruturas simples, mas uma análise mais detalhada revela que eles podem diferir muito entre si. Um estudo de suas propriedades cinemáticas indica diferenças entre bojos sustentados por rotação ou sustentados por dispersão de velocidades, enquanto que uma análise de suas morfologias permite diferenciar bojos com barras, sem barras, que apresentam ou não anéis e ainda galáxias nas quais bojos inexistem.

O objetivo desse estudo é fazer uma análise comparativa entre diversas morfologias de bojos de galáxias, através da comparação de suas propriedades morfológicas e das propriedades de suas populações estelares.

Foi usado o pacote R para o estudo das correlações entre parâmetros estruturais (referentes à morfologia da galáxia) e parâmetros de população estelar (idade e metalicidade) do bojo. Foi implementado um código que aplica os testes estatísticos de Pearson e de Spearman, e retorna o valor dos coeficientes de correlação e seus respectivos valores-p. A partir desses valores-p, apenas as correlações significativas (que possuam valor-p menor que 0.05) do teste de Spearman foram selecionadas e analisadas.

A partir desse estudo, podemos destacar os resultados mais importantes como sendo: 1) a fraca correlação encontrada entre Idade com parâmetros estruturais das galáxias, 2) o fato que grupos sem barra apresentam correlações mais fortes que grupos com barra, e 3) o efeito que AGNs causam quando estão presentes na amostra, tornando as correlações entre parâmetros mais fracas. Alguns resultados que ainda estão sendo discutidos pelo grupo também são apresentados, como a interessante forma com que se dá a relação entre a razão bojo-total e as distribuições de metalicidade do bojo. O presente estudo aponta algumas características bastante interessantes sobre bojos de galáxias que ainda não haviam sido notadas, e procura servir como preliminar para abordagens futuras.

Lista de Figuras

1.1 Objetos da amostra analisada no presente estudo, que seriam classificados como "objetos difusos" por Messier. Crédito: Gadotti (2009). 12

1.2 Estutura de uma galáxia espiral. Adaptado de: http://umanitoba.ca/faculties/science/astronomy/jwest/projects/umbrau59/spiralgalaxy.html 13

1.3 Diferença entre bojos. Obtido de http://users.obs.carnegiescience.edu/kmd/Teaching/Exgal/Exgal2014_Aula9_ToShare.pdf 15

2.1 Resumo dos principais testes estatísticos citados. A figura é de autoria própria. . 24

2.2 Diagrama ilustrando a avaliação qualitativa dos coeficientes de correlação, válidos tanto para Pearson quanto Spearman. Adaptado de https://www.eecis.udel.edu/~portnoi/classroom/prob_estatistica/2006_1/lecture_slides/aula20.pdf 25

3.1 Matriz de correlação de Pearson para toda a amostra. 29

3.2 Matriz de correlação de Pearson para bojos com barra, os que possuem AGN à esquerda e os sem AGN à direita. 29

3.3 Matriz de correlação de Pearson para bojos sem barra, os que possuem AGN à esquerda e os sem AGN à direita. 30

3.4 Matriz de correlação de Spearman para toda a amostra. 30

3.5 Matriz de correlação de Spearman para bojos com barra, os que possuem AGN à esquerda e os sem AGN à direita. 31

3.6 Matriz de correlação de Spearman para bojos sem barra, os que possuem AGN à esquerda e os sem AGN à direita. 31

3.7 Porcentagem das correlações significativas (valor p menor que 0.05) para cada subgrupo e para a amostra completa, comparando Pearson e Spearman. 32

3.8 Valores das correlações mais fortes entre [Fe/H] e parâmetros morfológicos, mostrados na primeira coluna. Valores obtidos com o teste de Spearman. 33
3.9 Valores dos parâmetros morfológicos que se correlacionam mais fortemente com a Idade. Valores obtidos com o teste de Spearman. 33

4.1 Dados referentes à amostra completa. O quadrado superior esquerdo mostra a distribuição de [Fe/H], o inferior direito mostra a distribuição da razão B/T. O número no canto superior direito é o coeficiente de correlação de Spearman entre os dois parâmetros, e o quadrado inferior esquerdo mostra dispersão dos dados. A linha em vermelho mostra a tendência estatística que a dispersão dos dados segue. 36
4.2 Distribuição de [Fe/H] para bojos clássicos (esquerda) e pseudo bojos (direita). . . 36
4.3 Dados referentes à amostra completa. O quadrado superior esquerdo mostra a distribuição de [Fe/H], o inferior direito mostra a distribuição da razão Massa Luminosidade do disco. O número no canto superior direito é o coeficiente de correlação de Spearman entre os dois parâmetros, e o quadrado inferior esquerdo mostra dispersão dos dados. A linha em vermelho mostra a tendência estatística que a dispersão dos dados segue. 37

C.1 Matriz de correlação de Spearman para toda a amostra. 49
C.2 Matriz de correlação de Spearman para bojos com barra e com AGN. 50
C.3 Matriz de correlação de Spearman para bojos com barra e sem AGN. 50
C.4 Matriz de correlação de Spearman para bojos sem barra mas com AGN. 51
C.5 Matriz de correlação de Spearman para bojos sem barra e sem AGN. 51

Sumário

1. Introdução 11

1.1 Estudo de bojos de galáxias 14

1.1.1 Bojos Clássicos 14

1.1.2 Pseudo bojos 14

1.2 Embasamento do projeto 16

1.2.1 Determinação dos parâmetros morfológicos 16

1.2.2 Determinação dos parâmetros de população estelar 18

2. Metodologia 20

2.1 Amostra e parâmetros 20

2.2 Estatística 22

3. Resultados 28

4. Discussão 35

4.1 Estatística dos valores-p 35

4.2 Correlação entre Razão bojo-total e [Fe/H] 35

4.3 Correlação entre Massa-Luminosidade do disco e [Fe/H] 37

5. Conclusão 38

Referências 40

Apêndice 43

A. Tabelas da significância dos valores-p 45

A.1 Tabela Spearman 45

A.2 Tabela Pearson 46

B. Script para matrizes de correlação em R 47

B.1 Script para matrizes 47

B.2 Função para formatar as matrizes 48

C. Matrizes de correlação Spearman, organizadas por componentes 49

Capítulo 1

Introdução

Galáxias são estruturas constituídas por estrelas, gás, poeira e outras partículas, como fótons, raios cósmicos e matéria escura. Estudar a história das galáxias implica estudar a história do próprio Universo, pois são elas que o povoam em suas maiores escalas e existem nos mais diferentes ambientes. Tratam-se de estruturas complexas, e para galáxias luminosas elas são classificadas em espirais, elípticas ou irregulares dependendo das suas características morfológicas (Hubble, 1926).

Nesse trabalho serão apresentados os resultados obtidos com o estudo das propriedades morfológicas e de população estelar das galáxias. No presente capítulo será dada uma apresentação histórica e uma breve introdução teórica sobre classificação e estrutura das galáxias, além de comentar brevemente os dois estudos prévios que serviram de base para esse trabalho. O Capítulo 2 evidencia os métodos utilizados para a seleção dos dados da amostra, além de fazer uma introdução teórica aos testes estatísticos que serão utilizados e reiterar os motivos pelos quais foram escolhidos. O Capítulo 3 apresenta os resultados obtidos e comenta a respeito de possíveis interpretações, e o Capítulo 4 mostra alguns resultados interessantes que ainda estão sendo discutidos pelo grupo. O Capítulo 5 apresenta um resumo das conclusões obtidas, e nos apêndices são encontrados tabelas de significância para os testes apresentados (Apêndice A), o script em R para os testes estatísticos (Apêndice B) e matrizes de correlação de Spearman (Apêndice C).

Diferente do que se possa imaginar, mesmo se tratando de estruturas muito grandes e, portanto, mais fáceis de serem observadas, as galáxias só vieram a ser consideradas sistemas independentes da Via Láctea no início do século XX (Schneider, 2006). Até então, a ideia de que as conhecidas "nebulosas espirais" não passavam de objetos dentro da nossa própria Galáxia era a hipótese mais comumente aceita na academia, tendo sido catalogadas por Charles Messier como "objetos difusos" que ainda não eram compreendidos (Messier, 1781). Na Figura 1.1 vemos, de cima para baixo, galáxias elípticas, galáxias com bojos clássicos, galáxias com pseudo-bojos e galáxias sem bojo. Para as três últimas linhas, as três primeiras colunas da esquerda ilustram

galáxias com barras:

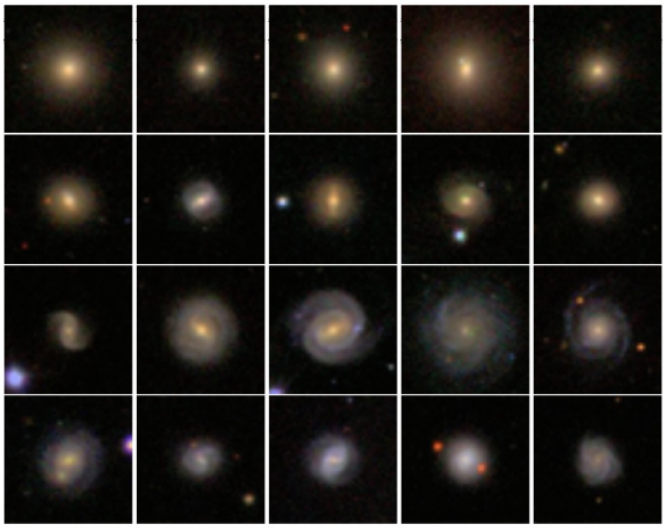

Figura 1.1: Objetos da amostra analisada no presente estudo, que seriam classificados como "objetos difusos" por Messier. Crédito: Gadotti (2009).

Dentre os astrofísicos da época as opiniões sobre a existência de galáxias eram divididas: Harlow Shapley defendia a teoria tradicional, e pelo contraste com o ideal de Heber Curtis de que tais nebulosas seriam na verdade outras galáxias tão grandes e complexas quanto a nossa própria foi que se deu o Grande Debate de 1920 (Sparke e Gallagher, 2007). A teoria de que esses objetos nebulosos observados eram na verdade outros "universos-ilha" como o nosso foi proposto por Kant (1755), que afirmava

> *"Se o tamanho de um sistema planetário em que a Terra é quase visto como um grão de areia enche o entendimento com espanto, como deliciosamente surpreendidos nós seremos quando examinarmos a multidão infinita dos mundos e dos sistemas que preenchem a totalidade da Via Láctea. Mas quanto mais seremos maravilhados quando descobrirmos que todos esses arranjos imensuráveis de estrelas criam uma outra unidade, cujo fim não sabemos e que é talvez, como a Via Láctea, inconcebivelmente grande e ainda, mais uma vez, apenas uma unidade em uma infinidade de outros sistemas."* (The Universal Natural History and Theory of Heaven, Immanuel Kant 1755, tradução livre)

Ainda que o resultado do debate tenha sido inconclusivo na época, o cálculo da distância entre a nossa Galáxia e Andrômeda feito por Edwin Hubble em 1923, envolvendo as variáveis Cefeidas, promoveu as evidências necessárias para provar que existiam outros "universos-ilha" que não estavam situados dentro da nossa Galáxia. Vê-se que a descoberta desses sistemas não-

dinamicamente ligados com a Via Láctea é um resultado consideravelmente recente, podendo-se concluir que mesmo com as resoluções espaciais que possuímos nos telescópios hoje em dia, muitas vezes o estudo detalhado da estrutura das galáxias pode não ser uma tarefa trivial.

A estrutura das galáxias varia muito dependendo de sua classificação, e o foco do estudo aqui apresentado são as galáxias espirais. Em específico, analisaremos uma de suas componentes: o bojo, e como este se comporta em relação às outras componentes. A morfologia de galáxias espirais se divide em bojo, disco, braços espirais (que podem ou não apresentar barras), e o halo, como mostra a Figura 1.2, e elas podem ser ainda subdivididas nas categorias Sa, Sb, Sc ou Sd, dependendo do grau de desenvolvimento dos braços espirais e em quão proeminente é o núcleo em relação ao disco (Hubble, 1926).

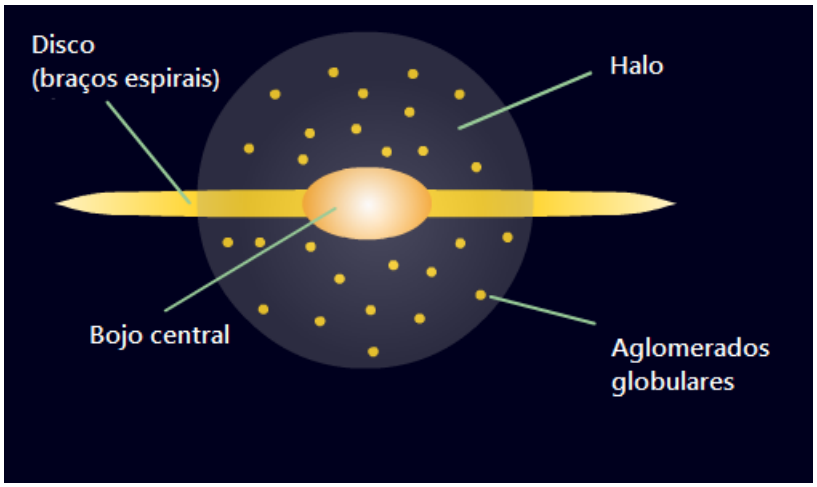

Figura 1.2: Estutura de uma galáxia espiral. Adaptado de: http://umanitoba.ca/faculties/science/astronomy/jwest/projects/umbrau59/spiralgalaxy.html

Em Sandage e Bedke (1994), as características dos bojos de galáxias espirais são apresentados de forma que o tamanho do bojo, que é melhor percebido em galáxias vistas pela borda, decresce progressivamente, enquanto que a taxa de formação estelar aumenta, desde galáxias Sa até Sd. Assim como é discutido em Carollo et al. (2000), a definição acima levanta a ideia que o bojo se trata de uma componente central das galáxias e que tem suas propriedades bastantes comparáveis às dos discos. Bojos são menores do que os discos das galáxias, mas apresentam brilho superficial muito mais alto. Também são sistemas dinamicamente quentes e possuem densidade de estrelas muito alta (10^4 vezes maior do que a da vizinhança solar).

De Carroll e Ostlie (2007) sabe-se que uma grande dificuldade encontrada ao observar as propriedades dos bojos, principalmente na nossa própria galáxia, é a extinção que os comprimentos de onda no visível sofrem, devido à grande presença de poeira nessas regiões. Ainda, alguns estudos sobre sua formação (Trager et al., 1998) obtiveram indícios, a partir das características

da sua população estelar (como idade e metalicidade), que bojos podem apresentar mais de um mecanismo de formação. Essas teorias vem sendo aprimoradas e possibilitaram classificar os bojos em clássicos ou pseudo bojos.

1.1 Estudo de bojos de galáxias

1.1.1 Bojos Clássicos

Bojos clássicos são componentes esferoidais, que tendem a ter populações estelares mais velhas do que as populações do disco, e geralmente são comparados a galáxias elípticas. Segundo a definição de Renzini (2006), é legítimo comparar bojos clássicos a galáxias elípticas que, de alguma forma, possuem um disco proeminente ao seu redor, e galáxias elípticas como bojos clássicos que perderam a oportunidade de gerar ou manter o disco proeminente, baseados na classificação de Hubble-Sandage-de Vaucouleurs (Kormendy e Kennicutt, 2004). São formados através de processos violentos como *minor mergers* ou durante o próprio colapso da protogaláxia. De Fisher e Drory (2011), bojos de galáxias que possuem índice de Sérsic maior do que 2 são classificados como clássicos, e a frequência de bojos clássicos aumenta fortemente com a massa estelar, sendo dominante em galáxias com massa acima de $10^{15} M_{\odot}$. Inicialmente, todos os bojos eram classificados como clássicos, mas as diferentes morfologias e características de populações estelares se tornaram evidências para diferentes processos de formação, dando origem aos pseudo bojos.

1.1.2 Pseudo bojos

Pseudo bojos são formados a partir de um processo longo de instabilidades e de evolução secular, dados pelo fluxo gradual de material das regiões do disco para as regiões mais internas da galáxia (Gadotti, 2009). Eles tendem a apresentar populações estelares mais jovens do que bojos clássicos, possuem cinemática suportada por rotação e um perfil de brilho superficial menos concentrado, que é mais próximo dos observados em discos e parametrizados pelo índice de Sérsic, como também é feito para bojos clássicos e discos. Quando é possível observar galáxias inclinadas, pseudo bojos se mostram mais achatados do que bojos clássicos, também se assemelhando mais à discos, e caso a galáxia seja vista pela borda, pseudos bojos tendem a apresentar uma forma de amendoim, o que acredita-se ser as partes internas de barras que evoluíram e saíram do plano do disco. Como pseudo bojos quase sempre possuem barras e suas galáxias hospedeiras tendem a apresentar braços espirais, é necessário que os mesmos sejam observados em alta resolução espacial, para que seja possível a identificação dessas estruturas.

Em resumo, de acordo com Kormendy e Kennicutt (2004), a lista preliminar abaixo trata de algumas características principais de pseudo bojos. Se a galáxia em estudo apresenta alguma das característica abaixo muito bem desenvolvida, é seguro afirmar que sua componente central se trata de um pseudo bojo.

1. O candidato a pseudo bojo é semelhante a discos quando visto em imagens, e seu achatamento aparente é similar ao disco externo.
2. Ele é, ou possui, barras nucleares (em galáxias vistas frontalmente).
3. Ele possui formato de caixa (*"box shaped"*, em galáxias vistas pela borda).
4. Possui índice de Sérsic aproximadamente entre 1 e 2.
5. É dominado por estrelas de população I (estrelas jovens, gás e poeira), e não há sinal de *mergers* em processo.

Resultados recentes (Kormendy et al., 2011) também mostram que as propriedades dos pseudo bojos não se correlacionam com a massa do buraco negro central. A diferença estrutural entre bojos clássicos e pseudo bojos pode ser observada na Figura 1.3.

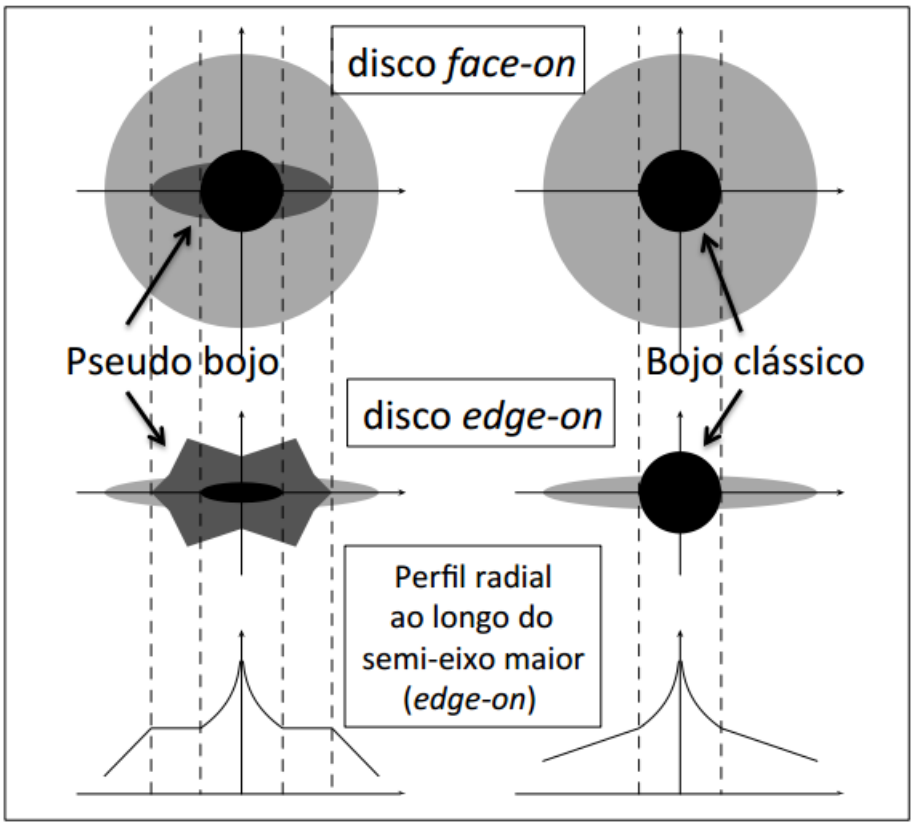

Figura 1.3: Diferença entre bojos. Obtido de http://users.obs.carnegiescience.edu/kmd/Teaching/Exgal/Exgal2014_Aula9_ToShare.pdf

Tendo em vista a complexidade do cenário em que a formação e o desenvolvimento dos bojos

de galáxias estão inseridos, fica clara a necessidade de estudar se esse desenvolvimento dos bojos é afetado pelo desenvolvimento do restante da galáxia, ou se ele se dá de forma independente. Essa é a questão que o presente estudo procura responder, e para isso, foram feitas análises dos parâmetros de bojos de galáxias obtidos em estudos anteriores, Gadotti (2009) e Coelho e Gadotti (2011). A sessão a seguir discute os dois estudos citados e mostra qual a metodologia adotada para a obtenção dos dados que serão utilizados nesse trabalho, além de fornecer mais detalhes sobre as diferenças entre bojos clássicos e pseudo-bojos.

1.2 Embasamento do projeto

1.2.1 Determinação dos parâmetros morfológicos

Gadotti (2009) fez a decomposição de imagem de uma amostra de quase 1000 componentes do Sloan Digital Sky Survey, contendo bojos, discos e barras. Concluiu que é possível utilizar a relação de Kormendy (relação entre o brilho superficial efetivo e o raio efetivo de galáxias) para distinguir bojos clássicos de pseudo bojos. Pela análise da distribuição de massa estelar de galáxias elípticas, bojos clássicos, pseudo bojos, discos e barras, e de seus parâmetros estruturais, foi possível inferir que 36% da massa estelar de galáxias massivas no universo local é contida em discos, 32% nas galáxias elípticas, 25% em bojos clássicos, 4% nas barras e 3% em pseudo bojos. Com o estudo, também foi possível apontar diferenças estruturais entre bojos clássicos e pseudo bojos, mostrando que seus processos de formação se deram de forma diferente, embora possam ter ocorrido concomitantemente.

Ao se fazer a comparação entre os diferentes meios de formação e evolução de galáxias, são estudadas questões que se mantém em aberto, como as diferenças estruturais entre galáxias elípticas, bojos clássicos e pseudo bojos, com que frequência galáxias disco abrigam pseudo bojos se comparadas aos bojos clássicos, como é dada a distribuição de massa e em quais condições pseudo bojos são importantes. A fim de obter a resposta para algumas dessas perguntas, Gadotti (2009) analisou detalhadamente uma amostra de 963 bojos do SDSS e, através desse trabalho, obteve o índice de Sérsic, comprimentos de escala e brilho superficial efetivo para galáxias elípticas, bojos clássicos e pseudo bojos, além de comprimentos de escala e brilho superficial central para discos, levando em conta a presença ou não de barras. Também foi determinada a razão bojo-total para as bandas g, r e i, além de cores e massas estelares para cada componente das galáxias.

No trabalho foi realizada a decomposição de imagem nas componentes estruturais, onde foi necessária a inclusão de componentes de barras uma vez que, caso elas não sejam levadas em conta na modelagem de galáxias com barras, as propriedades do bojo são afetadas conforme mostrado

em Gadotti (2008). A estimativa da massa estelar foi feita a partir da razão massa-luminosidade, obtida através das cores estimadas para cada componente, garantindo maior precisão nos dados. Para fazer a distinção entre bojos clássicos e pseudo bojos, foi usado um método diferente do índice de Sérsic: usualmente, se classificam em pseudo bojos aqueles que possuem o índice de Sérsic abaixo do que é esperado para bojos clássicos. Para se evitar ambiguidades nessa classificação e fazer um estudo mais confiável e com maior sentido físico, foi usada a relação de Kormendy (Kormendy e Kennicutt, 2004), uma vez que bojos clássicos seguem a mesma relação estabelecida para galáxias elípticas, diferentemente de pseudo bojos, e isso não depende do seu índice de Sérsic. Isso mostra que bojos clássicos contém algumas similaridades estruturais com galáxias elípticas que pseudo bojos não apresentam.

A amostra para o estudo foi selecionada para ser adequada tanto para uma análise estrutural com decomposição de imagens quanto para representar as populações de galáxias no universo local. Foram escolhidas galáxias com redshift $0.02 \leq z \leq 0.07$, o que provém um número estatisticamente grande de dados com uma resolução espacial adequada. Com o intuito de estudar galáxias mais massivas, foram escolhidas aquelas com massa superior a $10^{10} M\odot$, e com razão axial $b/a \geq 0.9$ para garantir que as galáxias estejam sendo vistas de frente, sendo que os parâmetros a e b foram medidos na isofota de 25 mag arcsec^{-2} na banda g. Ao garantir que as galáxias sejam vistas de frente, os efeitos causados pela projeção e pela poeira são atenuados, possibilitando a distinção das barras e uma determinação mais confiável dos parâmetros estruturais de cada componente das galáxias. Da amostra selecionada, foram excluídas galáxias que estejam sendo deturpadas por uma galáxia vizinha ou por um merger, galáxias muito fracas, irregulares ou que sejam muito pequenas para uma decomposição de imagem detalhada. A amostra final contém 963 galáxias e, dentre elas, 407 são AGNs. Ainda, a partir do redshift da amostra, é possível inferir que o espectro obtido representa a parte central da galáxia, dominada pelo bojo, com uma pequena contaminação do disco.

Dentre os resultados obtidos por Gadotti (2009), é importante ressaltar as diferenças encontradas entre os bojos, como o fato de que pseudo-bojos estão passando por uma intensa atividade de formação estelar, de forma que são sempre localizados na região azul do diagrama de cor-magnitude, enquanto que bojos clássicos quase sempre se encontram na região vermelha. Ainda, foi encontrada uma correlação entre o índice de Sérsic e a razão bojo-total que é seguida apenas por bojos clássicos, evidenciando que pseudo-bojos são formados por mecanismos diferentes. Também, o lugar ocupado por pseudo-bojos no plano fundamental é diferente do ocupado por bojos clássicos e mais próximo do ocupado pelos discos, e sua relação massa-tamanho é mais similiar à de barras do que à de bojos clássicos. Todos esses resultados corroboram a teoria de

que pseudo-bojos são formados por processos lentos e não violentos, como instabilidades do disco, enquanto bojos clássicos se devem a processos violentos, como *mergers*. Todos os parâmetros estruturais obtidos pelo estudo foram disponibilizados juntamente com seus valores típicos, e serão utilizados para o desenvolvimento do projeto aqui apresentado.

1.2.2 Determinação dos parâmetros de população estelar

Coelho e Gadotti (2011) obtiveram, através do ajuste espectral de 575 bojos de galáxias com espectros disponíveis no Sloan Digital Sky Survey, parâmetros de idades e metalicidades para galáxias com e sem barra e que apresentam ou não atividade de núcleo ativo.

O estudo foi realizado devido ao interesse em se estudar o impacto da presença ou ausência de barras na evolução dos bojos. Com um estudo aprofundado das propriedades de galáxias com barras é possível observar a ocorrência do fenômeno de rejuvenescimento da população estelar central em galáxias disco, o que se deve ao fato de que as barras são capazes de coletar o gás do disco e transferi-lo para as regiões centrais rapidamente, contribuindo para a aceleração da formação dos bojos pela maior taxa de formação estelar central. Tal fenômeno foi observado anteriormente uma vez que as linhas de emissão de Hα são mais proeminentes em galáxias com barras se comparadas com o espectro de galáxias que não apresentam barra, no entanto, não havia sido encontrado evidência direta desse fenômeno a partir das idades de populações estelares em bojos, quando se leva em consideração apenas galáxias disco de tipo precoce (early-type). O estudo de Coelho e Gadotti (2011) foi realizado utilizando técnicas de análise espectral em uma amostra bem definida de ambos os tipos de galáxias, nos quais foram possíveis utilizar uma alta significância estatística.

A amostra utilizada em Coelho e Gadotti (2011) é uma sub amostra da apresentada em Gadotti (2009), a qual contém as galáxias do SDSS Data Release 2 com massas maiores do que $10^{10} M\odot$, redshift entre $0.02 \leq z \leq 0.07$, e razão axial $b/a \geq 0.9$, uma vez que ao se escolher galáxias face-on os efeitos de poeira e projeção são minimizados, facilitando a identificação das barras.

Para verificar se as galáxias da amostra possuem ou não possuem barras, foram procuradas assinaturas típicas da presença de barras nas imagens das galáxias, nos contornos de isofotas, e pelo perfil pixel a pixel de intensidade radial. Para a amostra de Coelho e Gadotti (2011) foram selecionadas galáxias disco da amostra anterior, sendo no total 251 galáxias com barras e 324 galáxias sem barras, das quais 187 são AGNs de tipo 2. Ainda, de acordo com o teste de Kolmogorov-Smirnov, a diferença de distribuição de massa estelar em galáxias com barras e sem barras não é maior do que 1σ, de forma que possíveis vieses nos dados devido a massa estelar das

galáxias são descartados.

Para análise dos dados foi utilizado o código de ajuste espectral STARLIGHT, de forma análoga à desenvolvida em Cid Fernandes et al. (2005), calibrados aos modelos solares de Vazdekis et al. (2010).

Coelho e Gadotti (2011) determinaram que idades estelares médias de bojos de galáxias com barras e sem núcleo ativo são significantemente mais baixas se comparadas às de galáxias sem barra, com uma significância estatística de quase 4σ. Para bojos de galáxias massivas, com barras e que não apresentam núcleo ativo ($\log M_{bojo} > 10^{10.1} M\odot$), a distribuição de idade estelar é dada por uma gaussiana bimodal, com picos em 4.7 e 10.6 Gyr. Tal bimodalidade não existe em galáxias sem barra, fato que corrobora a teoria dos bojos serem formados por processos induzidos pelas barras. É importante ressaltar ainda que as amostras utilizadas, ainda que tenham distribuições de massa estelar muito similares, apresentam distribuições de massa de bojos significantemente diferentes, no sentido de que bojos de galáxias com barras tem menor massa. Isso é contra-intuitivo, uma vez que é esperado que o disco tenha menor massa, já que provêm dele o gás utilizado pelas barras para formar os bojos. Para esclarecer tal ocorrência é necessário um estudo teórico e observacional mais aprofundado.

Capítulo 2

Metodologia

A metodologia empregada no estudo consiste na ampla análise estatística da amostra, com base em testes de correlação aplicados tanto para a amostra completa quanto para 4 subgrupos. A amostra foi dividida baseando-se na presença ou não de barras e na presença ou ausência de atividade AGN. Isso é necessário para que seja possível identificar se alguma dessas estruturas afeta a evolução do bojo da galáxia. Uma vez que o objetivo do estudo é verificar se há relação entre a população estelar do bojo e a morfologia da galáxia como um todo, calcular o grau de associação entre os parâmetros estruturais e de população estelar se torna uma ferramenta muito útil. Isso se torna possível ao se calcular o coeficiente de correlação entre eles, que podem ser obtidos por meio de testes estatísticos. As seções abaixo descrevem a amostra utilizada nesse estudo e os testes estatísticos que serão aplicados.

2.1 Amostra e parâmetros

Inicialmente, foi feito um estudo da literatura a fim de compreender a fundo os parâmetros e os dados a serem analisados. Gadotti (2009) obteve, a partir de decomposição de imagem para 4 componentes (bojo, disco, barras e região central), os parâmetros estruturais para uma amostra de 963 galáxias, sendo que a escolha da amostra já foi detalhada na Seção 1.2.1. Os parâmetros estruturais obtidos em Gadotti (2009) e que foram utilizados nesse estudo são: brilho superficial central do disco, comprimento de escala disco, brilho superficial efetivo, raio efetivo e índice de Sérsic[1] do bojo; brilho superficial efetivo, raio efetivo, elipsidade, índice de Sérsic, semieixo maior e boxidade da barra; razão bojo total, razão disco-total, razão barra-total, redshift, razão M/L do bojo, disco e barra, e também a massa das 3 componentes.

A respeito da decomposição de imagem, Gadotti (2009) fez uso do código BUDDA v2.1, que é capaz de ajustar até quatro componentes em uma só imagem da galáxia. Com os ajustes dados

[1] Relação que mostra como a intensidade da galáxia varia com o raio, obtida a partir do perfil radial de brilho superficial das galáxias.

pelo BUDDA, foram calculadas as magnitudes absolutas para cada componente da galáxia, para todas as galáxias da amostra, nas bandas g, r e i. Para ser capaz de ajustar uma galáxia de cada vez, foi adotado o procedimento de, primeiramente, obter a imagem a ser ajustada e limpá-la de qualquer contaminação por estrelas em primeiro plano. As imagens foram ajustadas sem subtrair a imagem de fundo do céu, o que propicia um cálculo mais adequado de valores de chi^2. Em segundo lugar, foram dados os valores iniciais dos parâmetros a serem ajustados: os parâmetros geométricos foram definidos visualmente a partir do contorno de isofotas, enquanto que os perfis de brilho superficial foram estimados por uma inspeção pixel a pixel do perfil de intensidade radial da galáxia.

Vale ressaltar que, durante a comparação dos resultados preliminares encontrados pelo nosso grupo, percebemos a importância de incluir entre nossos parâmetros a dispersão de velocidade para a amostra, e a obtivemos diretamente do SDSS. Isso foi feito utilizando os parâmetros de identificação (plate, mjd e fiber) das galáxias da amostra e obtendo os valores de dispersão de velocidades que o SDSS disponibiliza para essas galáxias. A Tabela 2.1 apresenta a legenda dos nomes abreviados dos parâmetros, que foram utilizados nos eixos dos gráficos, e seus nomes completos.

Abreviação	Parâmetro
V_{disp}	Dispersão de velocidades
Age	Idade
[Fe/H]	Metalicidade
M_{bulge}	Massa do bojo
B/T	Razão bojo-total
Sersic (bulge)	Índice de Sersic do bojo
$r_{eff}(bulge)$	Raio efetivo do bojo
M_{disk}	Massa do disco
D/T	Razão disco-total
h^{disk}_{length}	Comprimento de escala disco
a_{bar}	Semmieixo maior da barra
$r_{eff}(bar)$	Raio efetivo da barra
Box(bar)	"*Boxiness´´* da barra
$B_{eff}(bulge)$	Brilho superficial efetivo do bojo
$B_{eff}(disk)$	Brilho superficial efetivo do disco
$B_{eff}(bar)$	Brilho efetivo da barra
M/L_{bulge}	Razão massa-luminosidade do bojo
M/L_{disk}	Razão massa-luminosidade do disco
M/L_{bar}	Razão massa-luminosidade da barra
Bar/T	Razão barra-total
Sersic (bar)	Índice de Sérsic da barra
e_{bar}	Elipcidade da barra
M_{bar}	Massa da barra
z	Redshift

Tabela 2.1 - Legenda para os parâmetros.

Com relação à população estelar, Coelho e Gadotti (2011) obtiveram os parâmetros de idade e metalicidade para uma sub-amostra de Gadotti (2009), a qual contém todas as galáxias disco. A amostra contém 575 bojos de galáxias, das quais 251 são galáxias com barras e 324 são galáxias sem barras, e dentre elas há 187 AGNs de tipo 2, sendo essa a mesma amostra utilizada nesse estudo.

2.2 Estatística

A escolha dos métodos adequados para a análise estatística deve ser baseada no objetivo do estudo, podendo requerer a comparação entre grupos, análise de erros, correlações e regressões lineares ou mesmo comprovar a normalidade da distribuição dos dados. A estatística permite que os resultados obtidos em um estudo possam ser validados ou contestados, comprovando que

a escolha correta dos métodos de análise é uma etapa significativa no desenvolvimento de uma pesquisa.

De Bussab e Morettin (1987), sabe-se que os testes estatísticos podem se dividir em dois grupos: paramétricos e não paramétricos. A maior diferença entre os dois é que os testes paramétricos têm como requisito a distribuição normal em amostras pequenas, enquanto que testes não paramétricos não fazem tal diferenciação, além de no geral não requerer parâmetros populacionais como média ou variância. Entretanto, testes não paramétricos não são tão eficientes quanto os paramétricos, podendo não apontar tantas diferenças entre os dados quando que elas realmente existem, necessitando amostras maiores ou maiores diferenças para rejeitar uma hipótese nula. Como algumas estatísticas pressupõem a distribuição normal em amostras pequenas (dimensão igual ou inferior a 30), pode ser necessário comprovar se a amostra possui tal distribuição. O teste de Kolmogorov-Smirnov[2] testa a aderência dos dados à normalidade, levando em consideração que as distribuições são contínuas e os pontos de máximo das mesmas (Wall e Jenkins, 2012).

Sobre testes paramétricos, é possível comparar a diferença entre as médias de dois grupos a partir do Teste-t de Student, mesmo quando a variância da população é desconhecida (Magalhães e Lima, 2004). Sobre testes não paramétricos, podem ser citados o teste de Mann-Whitney e o teste de Wilcoxon, que são alternativas não-paramétricas ao teste t, sendo o primeiro para variáveis independentes e o segundo para as dependentes. Há também o teste de Qui-quadrado, que avalia a dispersão entre variáveis e a associação entre elas, entre outros. Na Figura 2.1 a seguir é possível visualizar os casos adequados para cada teste estatístico citado e suas respectivas correspondências paramétricas e não-paramétricas.

[2] O teste K-S é mais conhecido como um teste de aderência, usado para analisar se as distribuições de duas populações diferem uma da outra. Com uma correção, chamada de correção de Lilliefors, tal teste pode ser adaptado para distribuições normais (texto adaptado de http://www.mat.uc.pt/~cmtm/ECwww/TestesNP.pdf)

Objetivo	Medida (de populações com distribuição normal)	Medida (de populações com distribuições não normais)
Descrever um grupo	Média e desvio-padrão	Mediana e amplitude interquartil
Comparar um grupo a um valor hipotético	Teste t para uma só amostra	Teste de Wilcoxon
Comparar dois grupos independentes	Teste t para duas amostras independentes	Teste de Mann-Whitney
Comparar dois grupos emparelhados	Teste t para amostras emparelhadas	Teste de Wilcoxon
Quantificar a associação entre duas variáveis	Correlação de Pearson	Correlação de Spearman
Prever valores a partir de outra variável de medida	Regressão linear simples	Regressão não paramétrica

Figura 2.1: Resumo dos principais testes estatísticos citados. A figura é de autoria própria.

Em estatística, a análise de correlação mede a intensidade e a direção de relações lineares ou não lineares entre variáveis, quando é preciso atender à necessidade de estabelecer a existência ou inexistência de uma relação entre medidas sem que seja necessário o ajuste de uma função matemática. No presente estudo, foram usados dois testes de correlação: o teste paramétrico de Pearson para correlações e seu equivalente não paramétrico, o teste de Spearman.

O coeficiente de correlação de Pearson, representado por r, é aplicado em variáveis quantitativas para medir o grau de relação linear entre elas, e segue os requisitos que as variáveis sejam aleatórias e apresentem uma distribuição normal bivariada conjunta, ou seja, para cada X dado, a variável Y é normalmente distribuída, caracterizando-o como um teste paramétrico. O coeficiente é definido pela equação a seguir, sendo x e y as duas variáveis que se deseja comparar, $x_{médio}$ e $y_{médio}$ são seus valores médio encontrados e r é o coeficiente de correlação de Pearson.

$$r = \frac{\sum (x_i - x_{médio})(y_i - y_{médio})}{\sqrt{(\sum (x_i - x_{médio})^2)(\sum (y_i - y_{médio})^2)}}. \tag{2.1}$$

É interessante notar que o valor de r não varia mesmo que os valores das variáveis sejam convertidos para uma escala diferente, e que ele obedece às regras de comutação. O coeficiente r é adimensional e possui valores entre -1 e 1, que reflete a intensidade da correlação como mostra a Figura 2.2.

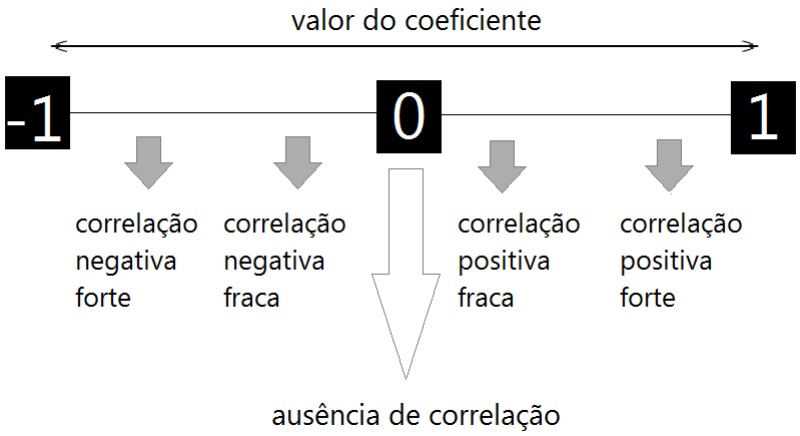

Figura 2.2: Diagrama ilustrando a avaliação qualitativa dos coeficientes de correlação, válidos tanto para Pearson quanto Spearman. Adaptado de https://www.eecis.udel.edu/~portnoi/classroom/prob_estatistica/2006_1/lecture_slides/aula20.pdf

Entretanto, a interpretação do valor de r depende muito dos objetivos pelos quais será utilizado, e nesse estudo será usada a definição dada em Callegari-Jacques (2003) tanto para r quanto para o coeficiente ρ de Spearman, que avalia os coeficientes qualitativamente em:

Se $0 < | \text{r} | < 0.30$, a correlação linear é fraca.

Se $0.30 \leq | \text{r} | < 0.60$, a correlação linear é moderada.

Se $0.60 \leq | \text{r} | < 0.90$, a correlação linear é forte.

Se $0.90 \leq | \text{r} | < 1$, a correlação linear é muito forte.

Vale ressaltar que o teste de Pearson mede apenas a intensidade de relações lineares, não servindo para analisar comportamentos não-lineares.

O coeficiente de correlação de Spearman é o mais antigo teste estatístico baseado em postos, ou seja, que leva em consideração a ordenação das variáveis e não seus valores observados ((Zar, 1999)). Por isso, ele não é sensível a assimetrias nas distribuições e nem a *outliers*, não exigindo o requisito de que as variáveis provenham de uma distribuição normal e se caracterizando como um teste não-paramétrico. Ele foi introduzido por Spearman em 1904, e conforme apresentado em Siegel (1975), foi derivado a partir do próprio estimador do coeficiente de correlação linear de Pearson. Ainda, segundo Guilford (1950), é o teste mais adequado a ser utilizado quando se deseja analisar a não-linearidade dos parâmetros, em contraste com o teste de Pearson.

O único requisito exigido pelo coeficiente de Spearman é que os elementos das amostras em estudo formem duas séries ordenadas, sendo representado por ρ e calculado pela relação a seguir:

$$\rho = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}, \tag{2.2}$$

onde d é a diferença entre as ordenações das variáveis e n é o número de pares de ordenações. O

coeficiente ρ também varia de -1 a 1, e segue as mesmas intensidades mostradas para o teste de Pearson (Figura 2.2).

Nesse estudo, a qualidade da correlação obtida pelo teste de Spearman será filtrada pela significância das correlações que é dada a partir do nível descritivo amostral, conhecido como valor-p e associado a cada uma das correlações encontradas. Os valores de significância associados ao valor-p se encontram na Tabela A.1 (para Spearman) e na tabela A.2 (para Pearson), ambas no apêndice A. No caso dos testes de correlação utilizados, a hipótese nula define que não existe correlação entre os parâmetros estudados, e o valor-p representa o menor nível de significância (ver apêndice) com que não se rejeita a hipótese nula. Em outras palavras, se o valor p for muito pequeno, a probabilidade do resultado obtido ser verídico é improvável, levando à rejeição da hipótese nula, que no caso é o desejado. É importante que o valor p não seja muito grande, pois isso pode levar a aceitação da hipótese nula quando a mesma é falsa, conhecido como erro de tipo II (Gavreau, 1994). É importante notar que o valor-p é bastante influenciado pelo tamanho da amostra, sendo que amostras grandes tendem a apresentar menores valores de valor-p. É preciso levar isso também em consideração quando se faz a escolha do valor-p adequado, impossibilitando a existência de um valor padrão que seja seguro a qualquer estudo. Levando esses aspectos em consideração, decidiu-se selecionar apenas as correlações que apresentassem valor-p menor que 0.05, considerado o valor tradicional.

No presente estudo não foi feito um aprofundamento pelo uso da análise de regressão uma vez que não era de interesse analisar o modelo teórico seguido pelos parâmetros da amostra, e sim provar quantitativamente a existência de relação entre eles. Ainda, não foi preciso realizar um teste de normalidade (como o teste de Shapiro-Willks (Magalhães e Lima, 2004) ou o teste de Kolmogorov-Smirnov para duas amostras independentes [3] para provar a normalidade dos dados pois a amostra é suficientemente grande para se comportar como uma normal, como é justificado pelo Teorema Central do Limite [4] para distribuições multivariadas (Johnson, 1988).

Foi desenvolvido um código no programa R [5] para aplicar os testes estatísticos nas amostras (no apêndice B). O código calcula as matrizes de correlação, os valores-p, exporta as matrizes calculadas e gera os gráficos para uma melhor visualização, tanto correlogramas quanto correlações

[3] Razali e Wah (2011) mostram como o teste K-S pode ser utilizado para analisar a distribuição normal de dados.

[4] O Teorema do Limite Central mostra que a distribuição das médias amostrais de variáveis independentes e aleatoriamente distribuídas tenderão a uma distribuição normal à medida que o tamanho da amostra cresce (Magalhães e Lima, 2004). É um dos mais importantes teoremas da estatística e é de fundamental importância para as áreas de inferência estatísticas, e sua demonstração pode ser comprovada em http://www.math.uah.edu/stat/sample/CLT.html

[5] https://www.r-project.org/

em mapa. Foi também feita uma função a parte para formatar as matrizes de correlação antes de exportá-las.

Capítulo 3

Resultados

Neste trabalho, utilizamos resultados da literatura ((Gadotti, 2009) e (Coelho e Gadotti, 2011)) para realizar um estudo comparativo entre propriedades estruturais e populações estelares em galáxias espirais. Tal desenvolvimento está descrito a seguir.

A fim de estudar os parâmetros descritos na Seção 1.1, foi feita inicialmente uma análise quantitativa e preliminar das correlações entre eles, utilizando o teste de Pearson. A amostra foi dividida em 4 subgrupos, separados pela presença ou ausência de barras e por possuir ou não atividade AGN. As matrizes de correlação mostradas nas Figuras 3.1, 3.2 e 3.3 apresentam todos os parâmetros do estudo, sem nenhuma seleção dos parâmetros que apresentem maior correlação ou melhores p-valores em relação aos outros. Com isso, o intuito era justamente estudar como todos os parâmetros se comportam em relação aos outros e analisar se alguma diferença significativa é notada quando as correlações são comparadas às encontradas pelo teste de Spearman. É importante ressaltar que nas matrizes de correlação abaixo, foi utilizada a ferramenta "*Hierarquical clustering*" (hclust) no código em R, para que todas as correlações mais fortes ficassem sempre agrupadas e fosse mais fácil nota-las. Isso muda a ordem dos parâmetros nos eixos, e as matrizes calculadas sem essa ferramenta e com os eixos agrupados por componentes (separadas em bojo, disco e barra) se encontram no apêndice A.

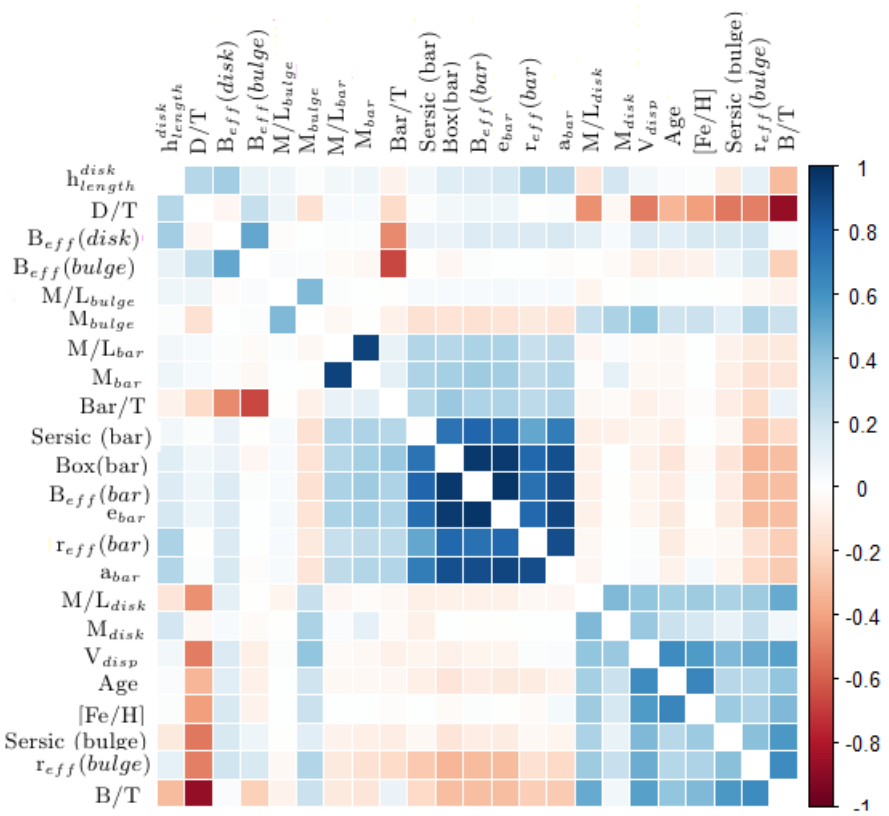

Figura 3.1: Matriz de correlação de Pearson para toda a amostra.

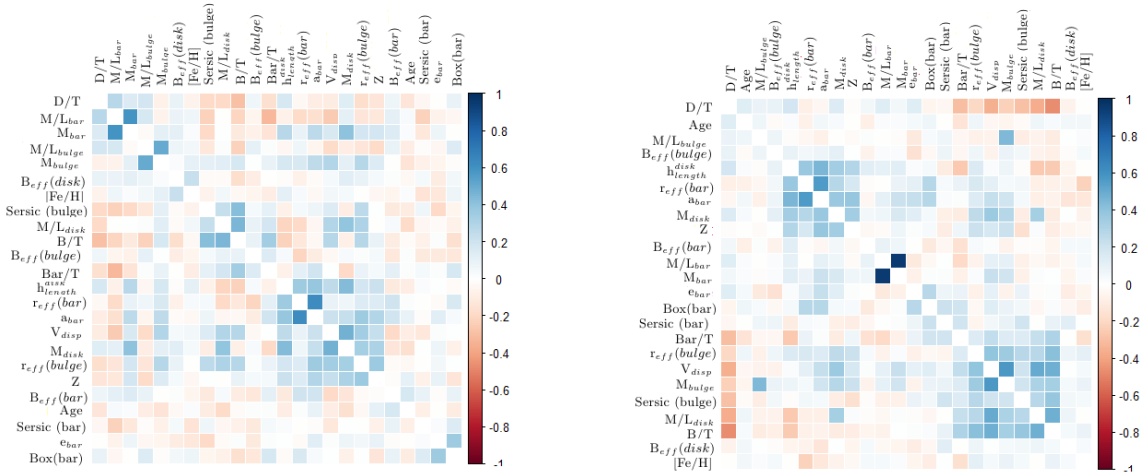

Figura 3.2: Matriz de correlação de Pearson para bojos com barra, os que possuem AGN à esquerda e os sem AGN à direita.

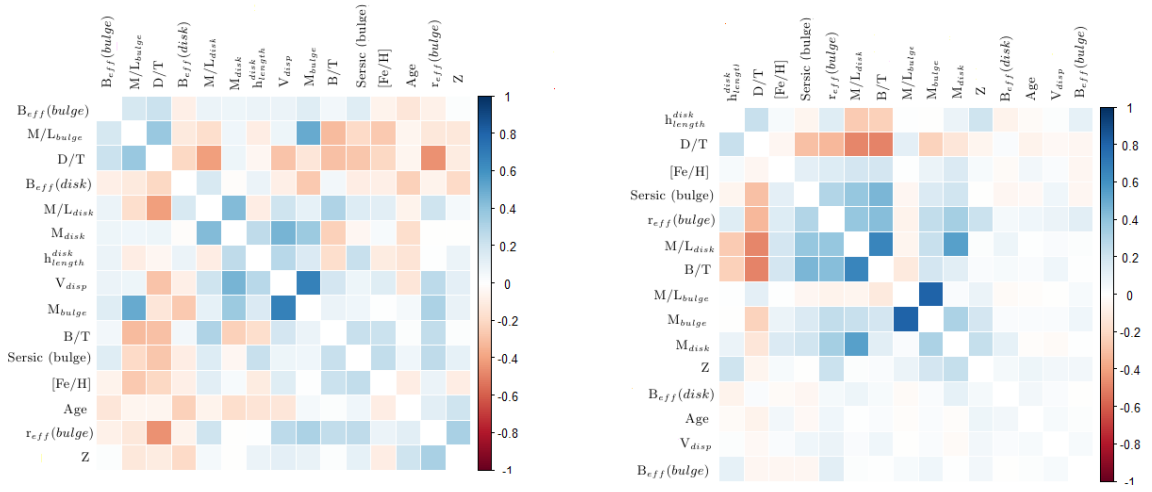

Figura 3.3: Matriz de correlação de Pearson para bojos sem barra, os que possuem AGN à esquerda e os sem AGN à direita.

Com o objetivo de analisar as correlações não lineares entre os parâmetros, foi realizado também o teste de Spearman. A partir dos valores das correlações obtidas por esse teste, foram selecionados apenas os parâmetros que apresentam correlações significativas (com p-valor menor do que 0.05). As matrizes de correlação resultantes são apresentados nas Figuras 3.4 a 3.6.

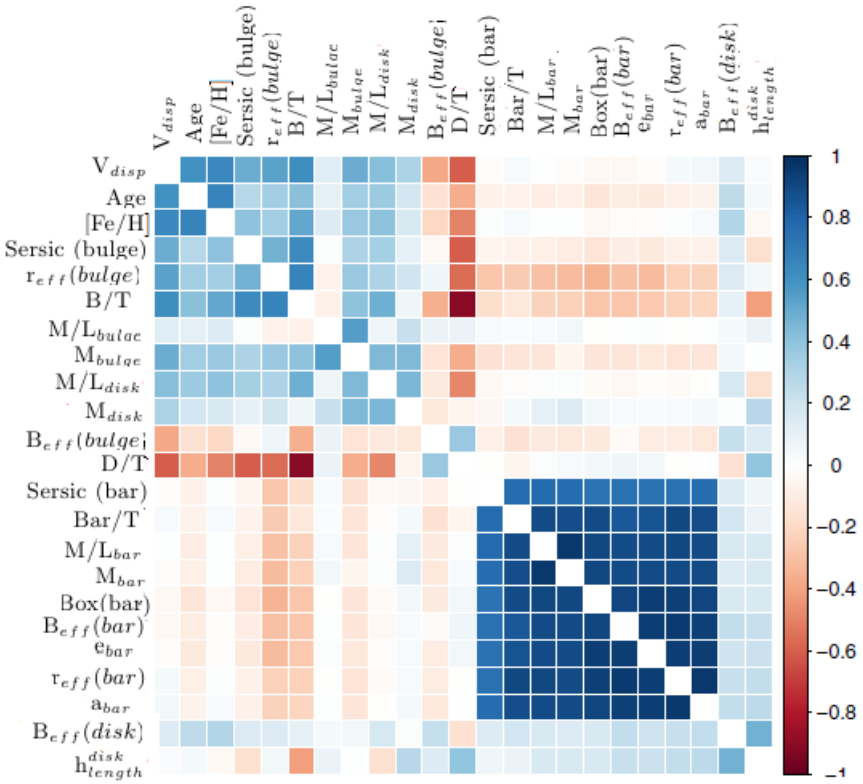

Figura 3.4: Matriz de correlação de Spearman para toda a amostra.

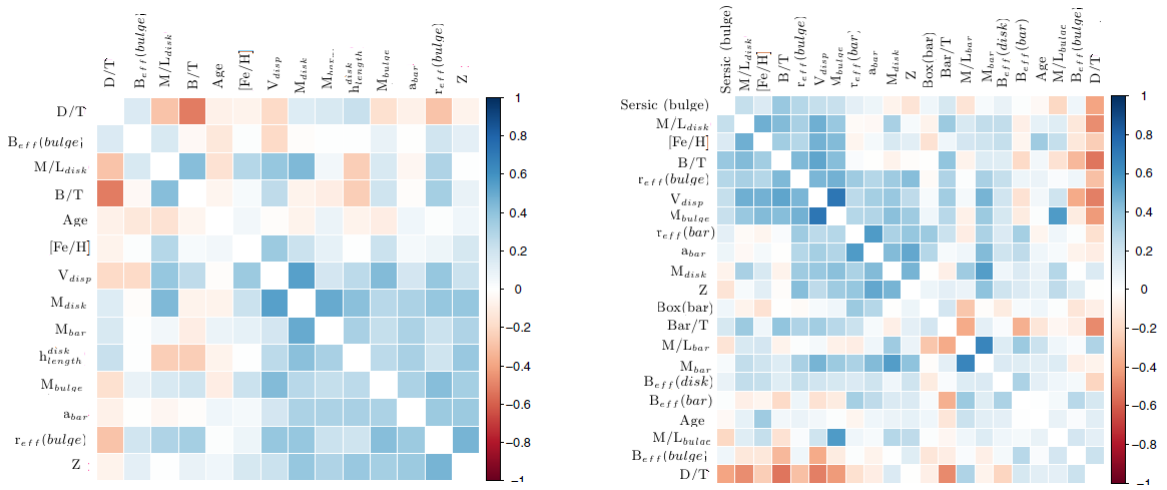

Figura 3.5: Matriz de correlação de Spearman para bojos com barra, os que possuem AGN à esquerda e os sem AGN à direita.

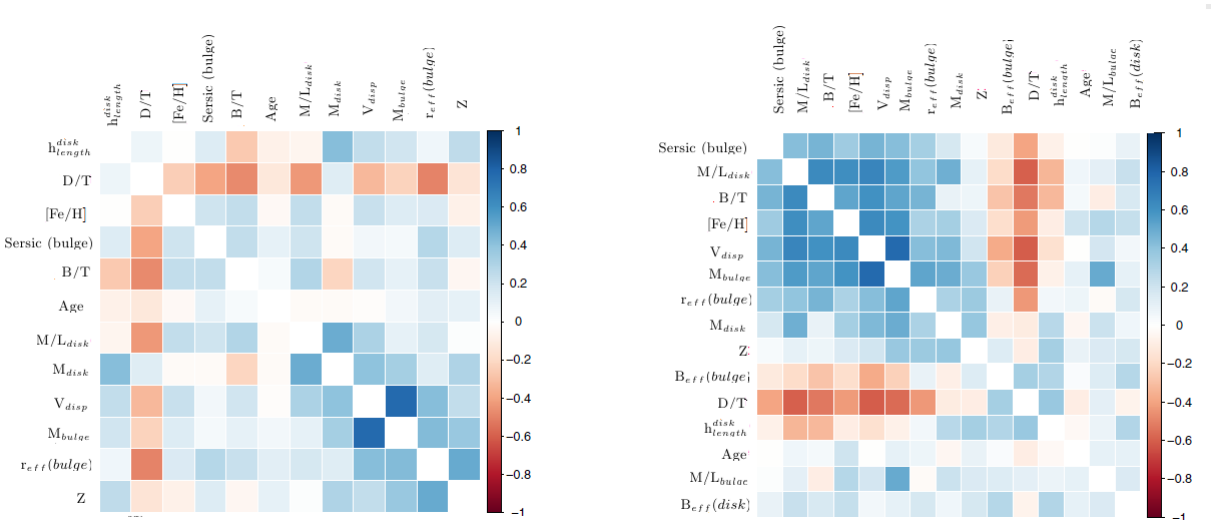

Figura 3.6: Matriz de correlação de Spearman para bojos sem barra, os que possuem AGN à esquerda e os sem AGN à direita.

Pela análise das matrizes de correlação, tanto as obtidas com o teste de Pearson quanto as pelo teste de Spearman, foi possível notar algumas correlações entre parâmetros que só apareceram nos testes de Spearman, e atribuímos esse resultado à presença de correlações não lineares em alguns grupos. Percebemos também que um número maior de correlações são obtidas com o teste de Spearman, mesmo após fazer a filtragem dos dados apenas para as correlações significativas, sobretudo nas subamostras que não apresentam AGN. Para a amostra completa, foram encontradas 155 correlações significativas (p-valor menor que 0.05) com Pearson, e 179 correlações significativas com Spearman. Para a amostra com barra e com AGN, foram 66 correlações com Pearson e 93 com Spearman, para barra e sem AGN foram 76 Pearson e 140 Spearman. Já para

os grupos que não possuem barra, foram 34 Pearson contra 83 Spearman para os que possuem AGN, e 38 Pearson contra 70 Spearman para os que não possuem AGN. O tamanho de cada submostra é diferente, então a Figura 3.7, dada em porcentagem, é uma melhor representação da verdadeira ocorrência de correlações não lineares em cada grupo.

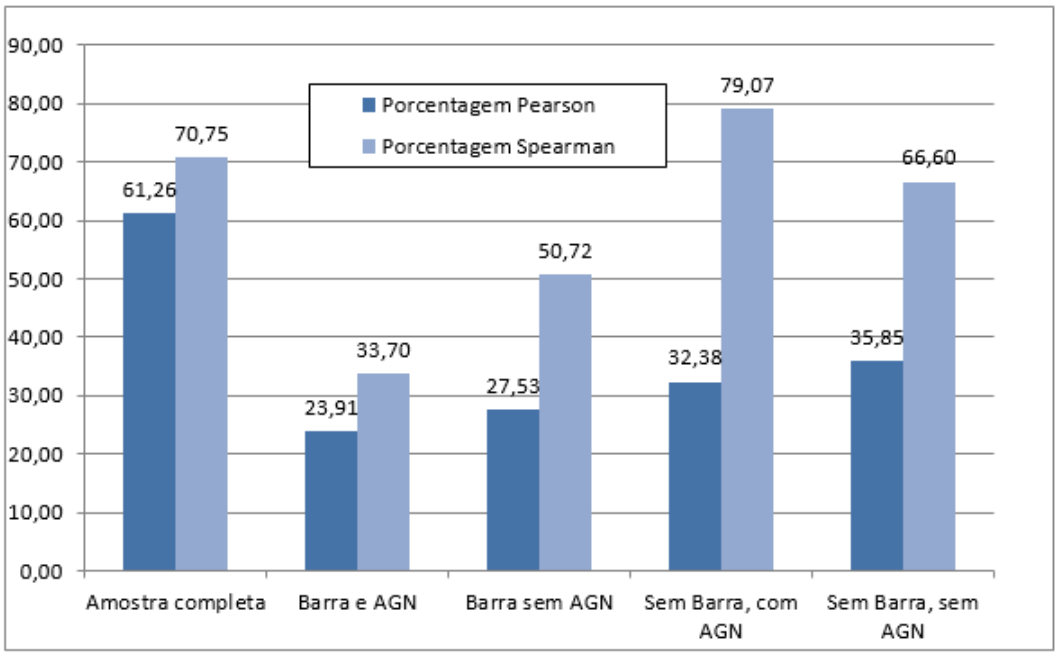

Figura 3.7: Porcentagem das correlações significativas (valor p menor que 0.05) para cada subgrupo e para a amostra completa, comparando Pearson e Spearman.

Há duas possíveis formas de interpretação desse gráfico:

1) Há maior ocorrência de correlação em grupos que não possuem barra, ou

2) Há um maior número de *outliers* nas amostras que não apresentam barra.

Ainda, a subamostra que não possui nem barra nem AGN apresenta correlações entre parâmetros morfológicos e o parâmetro [Fe/H] com valores muito mais altos do que nos outros grupos, sendo esse resultado revelado em ambos os testes estatísticos. A Figura 3.8 mostra os valores de correlação das 4 primeiras correlações mais fortes com relação ao [Fe/H], que são basicamente as mesmas para as 4 subamostras.

	Barra e AGN	Barra, sem AGN	Sem barra, AGN	Sem barra, sem AGN
Dispersão de Velocidades	0.36821	0.4726	0.22259	0.63159
Massa/Luminosidade disco	0.289177	0.484982	0.244557	0.619392
Massa do bojo	0.2173	0.438065	0.143878	0.597276
Razão Bojo/Total	0.034595	0.345186	0.247158	0.527848

Figura 3.8: Valores das correlações mais fortes entre [Fe/H] e parâmetros morfológicos, mostrados na primeira coluna. Valores obtidos com o teste de Spearman.

Pelo teste de Spearman, foi revelado que a amostra que possui barra e AGN é a única que não apresenta correlação forte entre o Índice de Sérsic do bojo e a dispersão de velocidades (com coeficiente r de Spearman 0.16 e p-valor 0.09). Ainda, notamos por meio desse teste que as correlações para os parâmetros morfológicos tem valores muito mais baixos com relação à idade do que com relação ao [Fe/H] e à dispersão de velocidades, que também apresentam p-valor mais alto, para todas as amostras. A Figura 3.9 abaixo mostra os parâmetros que mais se correlacionam com a Idade, na amostra completa, e mesmo esses parâmetros possuem correlações mais altas com [Fe/H] ou dispersão de velocidades do que com a própria idade.

	Idade	Dispersão de Velocidades	[Fe/H]
Razão bojo total	0.41	0.61	0.51
Massa Luminosidade do disco	0.36	0.43	0.40
Raio efetivo do bojo	0.34	0.53	0.35
Massa do bojo	0.34	0.38	0.50
Índice de Sérsic do bojo	0.28	0.49	0.40

Figura 3.9: Valores dos parâmetros morfológicos que se correlacionam mais fortemente com a Idade. Valores obtidos com o teste de Spearman.

Ainda, através das comparações entre as matrizes de correlação (apêndice A), foi notado que existe pouca correlação entre os parâmetros do bojo (incluindo idade e metalicidade, além dos parâmetros morfológicos) e os parâmetros das barras, podendo indicar que o bojo evolui com pouca influência do restante da galáxia. Através das matrizes também percebemos que quando a amostra estudada possui AGN, as correlações entre os parâmetros que são fortes em galáxias sem AGN ou somem ou se tornam muito fracas, sugerindo que a atividade AGN enfraquece a

relação evolutiva entre as componentes das galáxias.

Capítulo 4

Discussão

4.1 Estatística dos valores-p

Notamos que há um maior número de correlações, e também correlações mais fortes, quando a amostra completa é usada, ao invés das subamostras (com ou sem barra, e com ou sem AGN). Nesse contexto, é importante notar que o número de dados na amostra influencia fortemente o valor-p encontrado para as correlações, como já discutido na seção 2.2. Com um número de dados muito alto, os valores-p tendem a diminuir. Como a seleção de correlações significativas, nesse estudo, foi feita com base em valores-p menores do que 0.05, as amostras com um número de dados muito grande (como a amostra completa) pode apresentar um número maior de correlações que possuem valores-p menores do que 0.05. Isso pode ocorrer não só porque a amostra apresenta, de fato, maior grau de associação entre parâmetros, mas também pode ser devido a um viés estatístico causado pelo tamanho da amostra.

4.2 Correlação entre Razão bojo-total e [Fe/H]

Para visualizar qualitativamente as correlações entre os parâmetros, foram feitos gráficos das distribuições dos dados. A Figura 4.1 mostra um resultado relevante encontrado na distribuição não linear entre a razão bojo-total e o [Fe/H] do bojo. A existência da correlação já era um resultado esperado, mas ela se deu de uma forma interessante: valores baixos de [Fe/H] são encontrados para qualquer tamanho de bojo, mas valores altos de [Fe/H] só são encontrados em bojos com B/T alto, ou seja, bojos maiores.

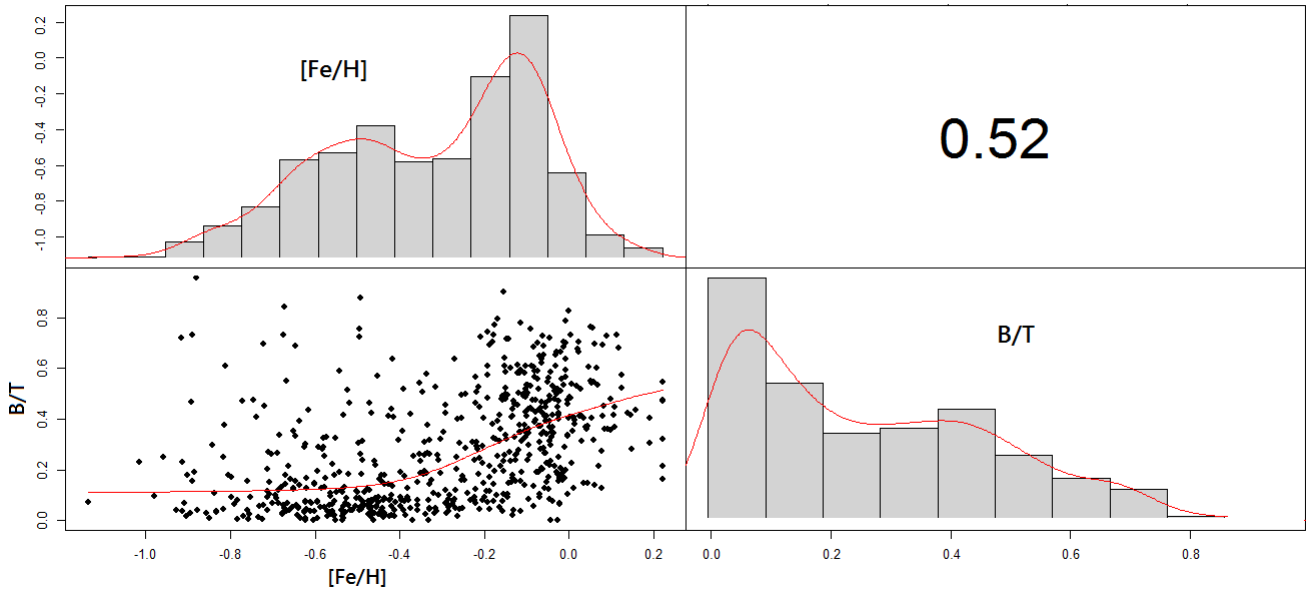

Figura 4.1: Dados referentes à amostra completa. O quadrado superior esquerdo mostra a distribuição de [Fe/H], o inferior direito mostra a distribuição da razão B/T. O número no canto superior direito é o coeficiente de correlação de Spearman entre os dois parâmetros, e o quadrado inferior esquerdo mostra dispersão dos dados. A linha em vermelho mostra a tendência estatística que a dispersão dos dados segue.

Bojos pequenos, no geral, são formados pela barra, sendo caracterizados como pseudo bojos (Seção 1.1.2). Para avaliar essa dependência entre o tamanho dos bojos, o [Fe/H] e suas classificações, foram feitos os histogramas da Figura 4.2. Era esperado, pela Figura 4.1, que pseudo bojos (no caso, bojos pequenos), apresentassem uma larga distribuição de [Fe/H], enquanto bojos clássicos (grandes) deveriam apresentar um pico maior em valores mais altos de [Fe/H]. Como é possível perceber nas Figuras abaixo, esses resultados foram confirmados. A classificação dos bojos da amostra em clássicos ou pseudo bojos foi feita por Gadotti (2009).

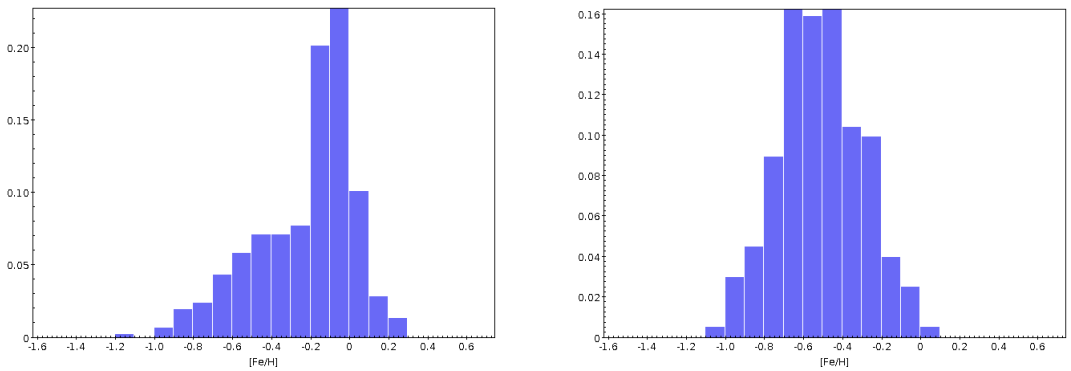

Figura 4.2: Distribuição de [Fe/H] para bojos clássicos (esquerda) e pseudo bojos (direita).

A próxima etapa consiste na comparação da distribuição de [Fe/H] para bojos com presença

ou ausência de barras, e tais resultados ainda serão analisados.

4.3 Correlação entre Massa-Luminosidade do disco e [Fe/H]

A correlação notada entre a Razão massa-luminosidade do disco e [Fe/H] se mostra muito interessante, uma vez que todos os valores de [Fe/H] foram calculados apenas para as componentes do bojos, e não para os discos ou para a galáxia como um todo. A Figura 4.5 mostra a relação entre os dois parâmetros, mas ainda não foi encontrado um motivo físico que explique esse resultado.

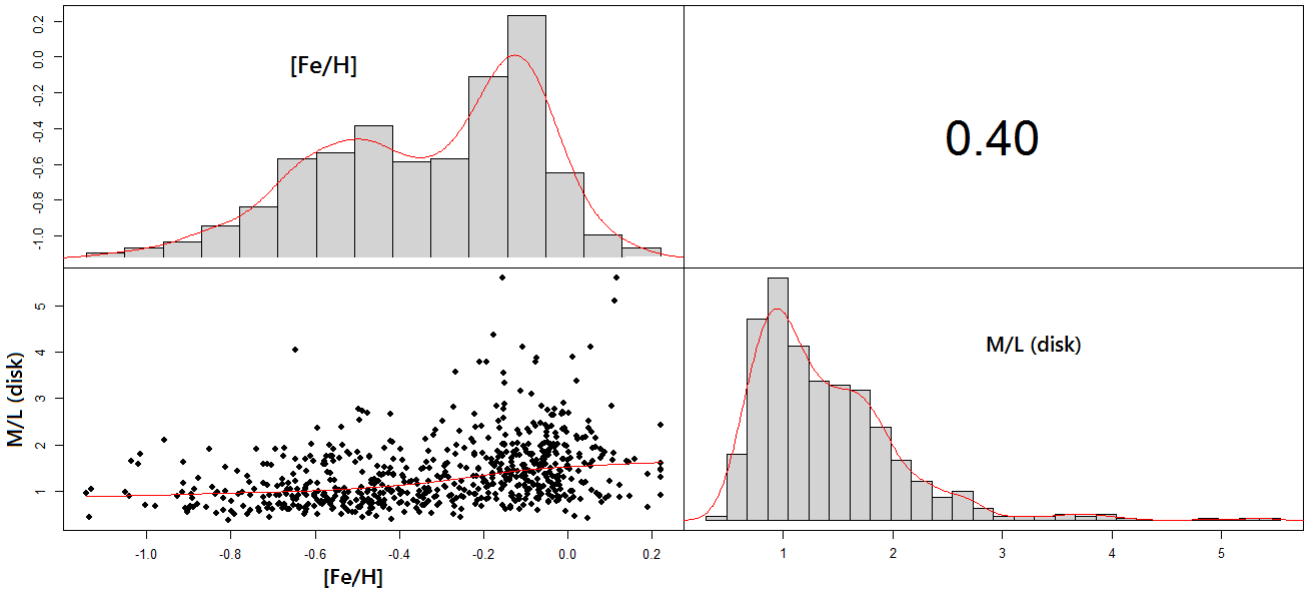

Figura 4.3: Dados referentes à amostra completa. O quadrado superior esquerdo mostra a distribuição de [Fe/H], o inferior direito mostra a distribuição da razão Massa Luminosidade do disco. O número no canto superior direito é o coeficiente de correlação de Spearman entre os dois parâmetros, e o quadrado inferior esquerdo mostra dispersão dos dados. A linha em vermelho mostra a tendência estatística que a dispersão dos dados segue.

Capítulo 5

Conclusão

A fim de encontrar correlações entre parâmetros estruturais e de população estelar em bojos de galáxias, a amostra completa e mais 4 subamostras, classificadas pela presença ou não de barras e de AGNs, foram submetidas a dois testes estatísticos: Pearson e Spearman.

Estatísticamente, pela análise das matrizes de correlação apresentadas tanto nas figuras de 3.1 a 3.6 quanto no Apêndice C, foi possível notar que algumas correlações significativas aparecem somente no teste de Spearman, concluindo que há correlações não lineares entre os parâmetros que só são notadas pelo teste não-paramétrico. Pela figura 3.7, também, podemos ver que um número maior de correlações é obtida quando se usa o teste de Spearman, se comparado ao número encontrado pelo teste de Pearson.

A respeito da correlação individual entre parâmetros, foi notado que os parâmetros morfológicos se correlacionam muito fracamente com a idade, enquanto apresentam correlações mais fortes com [Fe/H] e à dispersão de velocidades do bojo. Com base nas matrizes de correlação e na figura 3.7, vemos que há uma ocorrência maior de correlações em grupos que não possuem barra, sendo que esse resultado também pode ser causado caso haja um grande número de *outliers* na amostra. Ainda, pela figura 3.8, podemos perceber que o grupo que não apresenta barra nem AGN possui correlações entre parâmetros estruturais e os valores de [Fe/H] muito mais altos que os outros 3 subgrupos.

Comparando-se as matrizes de correlação (Apêndice C), é possível perceber que quando as AGNs não estão presentes na amostra as correlações entre os parâmetros ficam mais fortes. Esses resultados parecem sugerir que tanto a presença de barras quanto a de AGNs tendem a enfraquecer as correlações entre a população estelar do bojo e os parâmetros estruturais da galáxia. Pela matriz de correlação para a amostra completa (figura C.1), é notado que existe pouca correlação entre os parâmetros do bojo (idade e metalicidade, além dos parâmetros de população estelar) e os parâmetros da barra, o que pode indicar que o bojo evolui com pouca influência do restante da galáxia, levando-se em consideração também o resultado anterior que

mostra que os grupos sem barra apresentam maiores correlações. O único parâmetro da barra que possui forte correlação com um parâmetro do bojo é a razão Bar/T, se relacionando com $B_{eff}(bulge)$, indicando que quanto mais importante é a barra na galáxia, mais brilhante é o próprio bojo.

Os resultados apresentados no capítulo de discussão ainda estão sendo elaborados, e é necessário um estudo teórico mais aprofundado para interpretá-los. Entretanto, já é possível notar que pseudo bojos e bojos clássicos possuem diferentes distribuições de [Fe/H] e isso se correlaciona fortemente com a razão bojo-total. A razão massa-luminosidade do disco também apresenta grande correlação com o [Fe/H] do bojo, podendo dar indícios entre uma relação evolutiva entre o bojo e o disco.

Com base nos resultados encontrados, foi visto que existe uma grande gama de correlações entre os bojos e o restante das componentes das galáxias. Foi possível perceber que diferentes tipos de bojos (clássicos ou pseudo bojos) apontam correlações diferentes com relação à sua população estelar, principalmente [Fe/H], evidenciando que devem se formar por meio de processos diferentes. As diferentes componentes das galáxias se correlacionam de forma distinta, e algumas não chegam a se correlacionar. As relações encontradas são muitas, mas no geral os resultados apontam que os bojos podem evoluir independentemente de algumas componentes mas que de fato sofre influência de outras, como as AGNs. É notável também que a idade do bojo não é influenciada pela evolução do restante da galáxia, uma vez que a correlação entre ela e o restante dos parâmetros morfológicos é muito baixa, em todos os testes aplicados.

Tendo em vista a complexidade em se estudar bojos de galáxias, o aprofundamento teórico em astrofísica extragaláctica e em estatística se fizeram essenciais para a realização desse estudo. A abordagem estatística utilizada se mostrou muito eficiente para o objetivo do projeto, e apontou interessantes resultados que poderão ser melhor examinados no futuro. Ainda que galáxias se tratem de estruturas complexas e muito pesquisadas atualmente, ainda há uma infinidade de estudos a serem aprofundados, e o trabalho aqui apresentado procura elucidar as relações entre suas estruturas e, eventualmente, servir de guia para possíveis estudos futuros.

Referências Bibliográficas

Bussab W. O., Morettin P. A., Estatística Básica, 1987

Callegari-Jacques S. M., Bioestatística: princípios e aplicações, 2003

Carollo C., Fergunson H., Wyse R., The Formation of Galactic Bulges, 2000

Carroll B. W., Ostlie D. A., An Introduction to Modern Astrophysics 2nd (international) edn, 2007

Cid Fernandes R., Mateus A., Sodré L., Stasińska G., Gomes J. M., Semi-empirical analysis of Sloan Digital Sky Survey galaxies - I. Spectral synthesis method, MNRAS, 2005, vol. 358, p. 363

Coelho P., Gadotti D. A., Bars Rejuvenating Bulges? Evidence from Stellar Population Analysis, ApJ, 2011, vol. 743, p. L13

Fisher D. B., Drory N., Demographics of Bulge Types within 11 Mpc and Implications for Galaxy Evolution, ApJ, 2011, vol. 733, p. L47

Gadotti D. A., The Evolution of Bar Pattern Speed with Time and Bulge Prominence, ArXiv e-prints, 2008

Gadotti D. A., Structural properties of pseudo-bulges, classical bulges and elliptical galaxies: a Sloan Digital Sky Survey perspective, MNRAS, 2009, vol. 393, p. 1531

Gavreau K ; Pagano M., Why 5%? Nutrition., 1994

Guilford J. P., Fundamental statistics in psychology and education., 1950

Hubble E. P., Extragalactic nebulae., ApJ, 1926, vol. 64, p. 321

Johnson Richard A.; Wichern D. W., Applied multivariate statistical analysis., 1988

Kant I., Universal Natural History and Theory of Heaven, 1755

Kormendy J., Bender R., Cornell M. E., Supermassive black holes do not correlate with galaxy disks or pseudobulges, Nature, 2011, vol. 469, p. 374

Kormendy J., Kennicutt Jr. R. C., Secular Evolution and the Formation of Pseudobulges in Disk Galaxies, ARA&A, 2004, vol. 42, p. 603

Magalhães M., Lima A., Noções de Probabilidade e Estatística, 2004

Messier C., , 1781 Technical report Catalogue des Nébuleuses des amas d'Étoiles (Catalog of Nebulae and Star Clusters)

Razali N., Wah Y. B., Power comparisons of ShapiroâWilk, KolmogorovâSmirnov, Lilliefors and AndersonâDarling tests, 2011

Renzini A., Stellar Population Diagnostics of Elliptical Galaxy Formation, ARA&A, 2006, vol. 44, p. 141

Sandage A., Bedke J., The Carnegie Atlas of Galaxies, 1994, vol. 1

Schneider P., Extragalactic Astronomy and Cosmology., 2006

Siegel S., Estatística não-paramétrica: para as ciências do comportamento, 1975

Sparke L., Gallagher J., Galaxies in the Universe: An Introduction, 2007

Trager S. C., Worthey G., Faber S. M., Burstein D., González J. J., Old Stellar Populations. VI. Absorption-Line Spectra of Galaxy Nuclei and Globular Clusters, ApJS, 1998, vol. 116, p. 1

Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., Evolutionary stellar population synthesis with MILES - I. The base models and a new line index system, MNRAS, 2010, vol. 404, p. 1639

Wall J. V., Jenkins C. R., Practical Statistics for Astronomers., 2012

Zar J., Biostatistical Analysis., 1999

Apêndice

Apêndice A

Tabelas da significância dos valores-p

A.1 Tabela Spearman

Tabela A.1: Valor mínimo para o coeficiente de correlação de Spearman ser significativo, dados por Leach (1979) para n < 30, e dados pela estatística z $= \rho(n-1)^{1/2}$ para n>30. Retirado de Magalhães e Lima (2004).

Nível de significância, α, num teste bilateral	
n	**0,050**
5	1,000
6	0,886
7	0,786
8	0,738
9	0,700
10	0,648
11	0,618
12	0,587
13	0,560
14	0,538
15	0,521
16	0,503
17	0,488
18	0,474
19	0,460
20	0,447
21	0,436
22	0,425
23	0,416
24	0,407
25	0,398
26	0,390
27	0,383
28	0,375
29	0,369
30	0,362
35	0,336
40	0,314
45	0,295
50	0,280
60	0,255
70	0,236
80	0,221
90	0,208
100	0.197

A.2 Tabela Pearson

Tabela A.2: Valor mínimo para o coeficiente de correlação de Pearson ser significativo, com base na estatística $t = r(n-2)^{1/2}/(1-r^2)^{1/2}$. Retirado de Magalhães e Lima (2004).

Nível de significância, α, num teste bilateral

n	0,050
5	0,878
6	0,811
7	0,754
8	0,707
9	0,666
10	0,632
11	0,602
12	0,576
13	0,553
14	0,532
15	0,514
16	0,497
17	0,482
18	0,468
19	0,456
20	0,444
21	0,433
22	0,423
23	0,413
24	0,404
25	0,396
26	0,388
27	0,381
28	0,374
29	0,367
30	0,361
35	0,334
40	0,312
45	0,294
50	0,279
60	0,254
70	0,235
80	0,220
90	0,207
100	0,197

Apêndice B

Script para matrizes de correlação em R

B.1 Script para matrizes

```r
# lê a tabela
tabela = read.table(file="arquivo.csv",header=T,sep=";")

#calcula as matrizes de correlação pearson e spearman
pearson = cor(tabela, method = "pearson", use = "complete.obs")
spearman = cor(tabela, method = "spearman", use = "complete.obs")

#calcula os p-valores
library(Hmisc)
pearsonpvalue = rcorr(as.matrix(tabela), type = "pearson")
spearmanpvalue = rcorr(as.matrix(tabela), type = "spearman")

#carrega a função 'formataMatriz.R' para formatar a matriz
source('formataMatriz.R')

#formata as matrizes de correlacao junto com os p-value
library(Hmisc)
escreverP = formataMatriz(pearsonpvalue$r, pearsonpvalue$P)
escreverS = formataMatriz(spearmanpvalue$r, spearmanpvalue$P)

#exporta as matrizes formatadas, com seus respectivos p-valores
write.table(escreverP, "c:/Users/Diana/Desktop/pearsonPVALUES.txt", sep="\t")
write.table(escreverS, "c:/Users/Diana/Desktop/spearmanPVALUES.txt", sep="\t")

#plota os correlogramas. Em algumas tabelas a função "hclust" foi retirada.
library(corrplot)
corrplot(pearson, order="hclust", method='color', diag = FALSE)
corrplot(spearman, order="hclust", method='color',  diag=FALSE)

#plota os chart
library(PerformanceAnalytics)
chart.Correlation(melior, histogram=TRUE, pch=19) #para pearson
chart.Correlation(melior, histogram=TRUE, pch=19, method = "spearman") #para spearman
```

B.2 Função para formatar as matrizes

```r
# ++++++++++++++++++++++++++++
# Formata Matriz de correlação
# ++++++++++++++++++++++++++++
# cormat : matriz dos coeficientes de correlação
# pmat : matriz dos p-valores
formataMatriz <- function(cormat, pmat) {
  ut <- upper.tri(cormat)
  data.frame(
    row = rownames(cormat)[row(cormat)[ut]],
    column = rownames(cormat)[col(cormat)[ut]],
    cor  =(cormat)[ut],
    p = pmat[ut]
    )
}
```

Apêndice C

Matrizes de correlação Spearman, organizadas por componentes

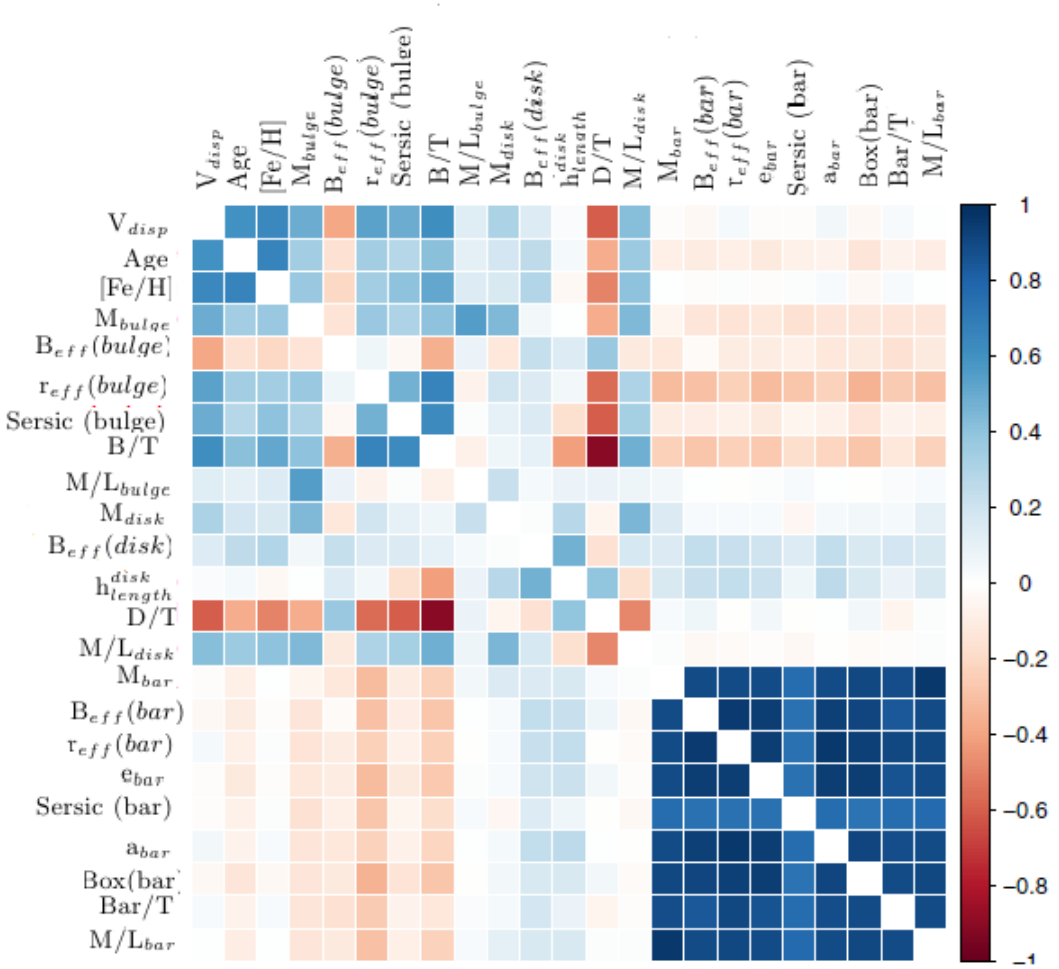

Figura C.1: Matriz de correlação de Spearman para toda a amostra.

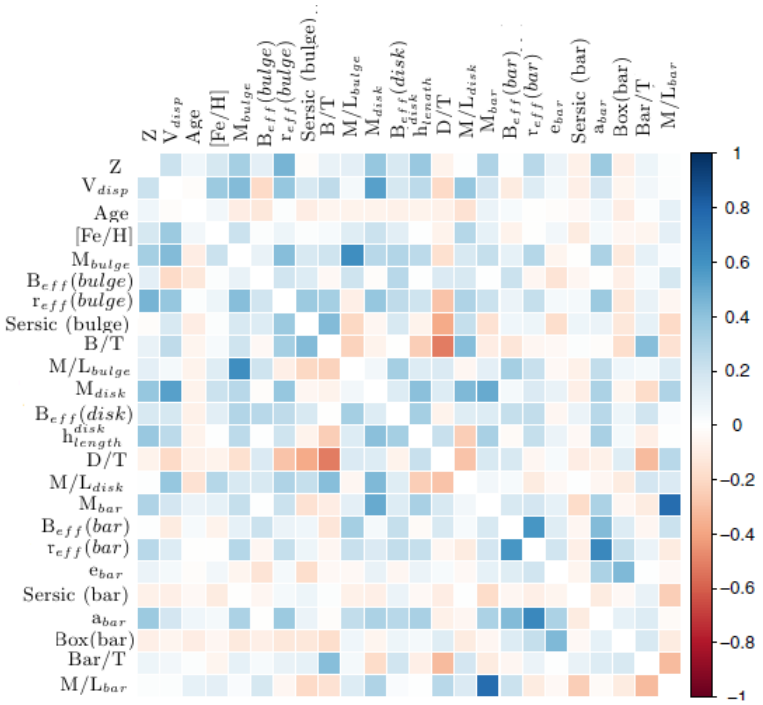

Figura C.2: Matriz de correlação de Spearman para bojos com barra e com AGN.

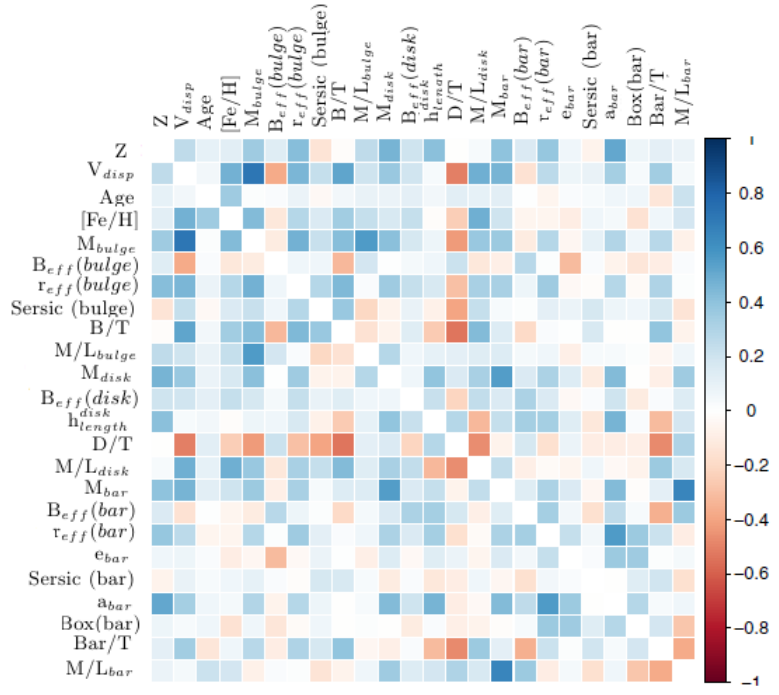

Figura C.3: Matriz de correlação de Spearman para bojos com barra e sem AGN.

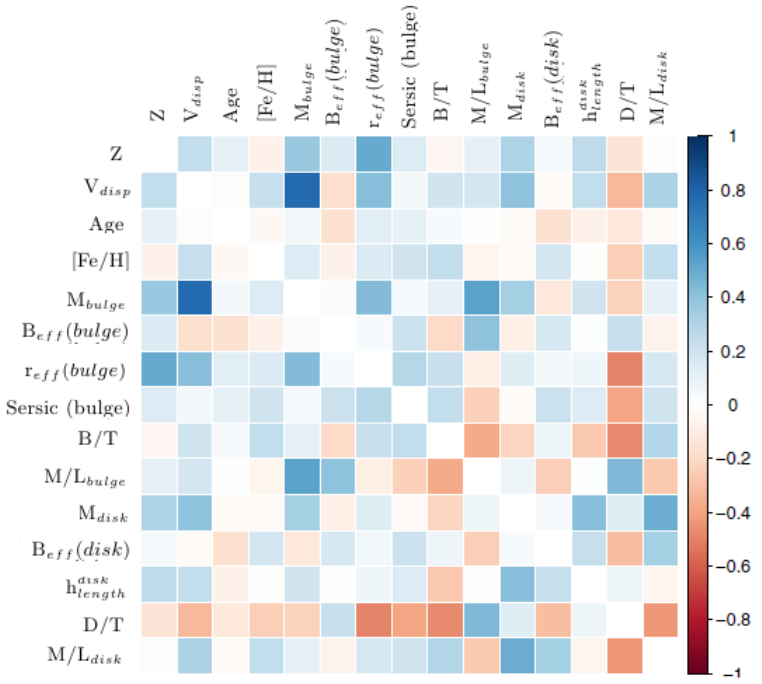

Figura C.4: Matriz de correlação de Spearman para bojos sem barra mas com AGN.

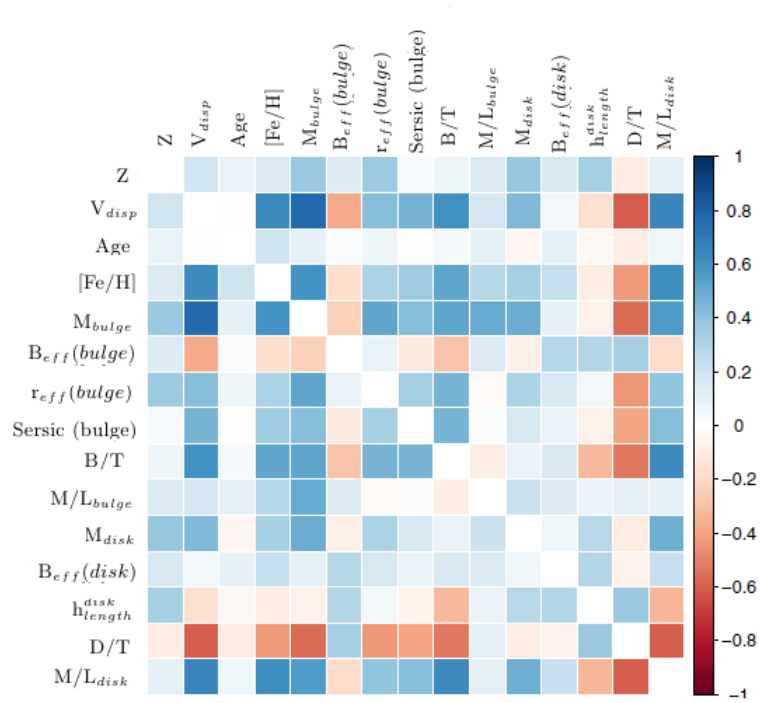

Figura C.5: Matriz de correlação de Spearman para bojos sem barra e sem AGN.