

# Como Vender Pipoca Na Guerra

William Schoenell

Orientador:

Prof. Dr. Roberto Cid Fernandes Jr.

• • •

Universidade Federal de Santa Catarina  
Centro de Ciências Físicas e Matemáticas  
Departamento de Física

Dissertação de mestrado apresentada ao Departamento de Física da UFSC em preenchimento parcial dos requisitos para obtenção do título de Mestre em Física.

Trabalho financiado pelo Instituto Nacional de Ciência e Tecnologia de Astrofísica e CAPES.

Florianópolis (SC) – 2011

*Aos meus pais e minha irmã.*

# Agradecimientos

## Resumo

Resumo RESUMO Resumo...

## **Abstract**

Abstract ABSTRACT Abstract

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	O STARLIGHT e o mega-levantamento astronômico SDSS . . . . .	1
1.2	Os dados do J-PAS e este trabalho . . . . .	3
<b>2</b>	<b>Do SDSS ao J-PAS</b>	<b>6</b>
2.1	O megalevantamento SDSS . . . . .	6
2.2	O megalevantamento J-PAS . . . . .	10
2.3	Conversão de Espectros do SDSS para o J-PAS . . . . .	13
2.4	Espectros de erros para o espectro J-PAS . . . . .	13
2.5	O J-Espectro . . . . .	16
<b>3</b>	<b>Método</b>	<b>18</b>
3.1	Introdução . . . . .	18
3.2	Definições . . . . .	21
3.3	Casamento Espectral: Figura de mérito e pesos . . . . .	21
3.4	Casamento Espectral: Exemplos . . . . .	23
3.5	Formalismo Estatístico . . . . .	28
<b>4</b>	<b>Amostras, bases e detalhes técnicos</b>	<b>34</b>
4.1	Amostra Geral . . . . .	34

---

4.2	Bases . . . . .	35
4.2.1	Base de Propriedades Físicas: Diagrama “cor-magnitude” . . . . .	36
4.2.2	Base sobre o WHAN . . . . .	38
4.3	Fator-cozinha $f_L$ . . . . .	43
4.4	Estimadores (Resumos da Função Distribuição de Probabilidades) . . . . .	46
<b>5</b>	<b>Resultados</b>	<b>47</b>
5.1	Propriedades de interesse . . . . .	47
5.2	Resultados: Propriedades Físicas . . . . .	49
5.3	Resultados: Linhas de Emissão . . . . .	50
5.4	“OUTLIERS” . . . . .	51
5.5	Resumo? . . . . .	55
<b>6</b>	<b>Conclusões e Perspectivas</b>	<b>56</b>
6.1	Resumo do trabalho e seus principais resultados . . . . .	56
6.2	Perspectivas . . . . .	58
6.3	RESTOS... . . . .	59
6.4	lixo possivelmente reciclavel... . . . .	60
	<b>Referências Bibliográficas</b>	<b>I</b>

# Lista de Figuras

1.1	Uma modo de analisar os J-Espectros. . . . .	4
2.1	Projeção equatorial das regiões cobertas pelo levantamento espectroscopico do SDSS. . . . .	8
2.2	Exemplo de um prato de fibras do SDSS. . . . .	9
2.3	Desenho conceitual da JPCAM que equipará o telescópio de 2,52m no OAJ. . . . .	11
2.4	Curvas de transmissividade utilizadas neste trabalho. . . . .	12
2.5	Exemplos de J-espectros normalizados. . . . .	17
3.1	Esquema gráfico do método proposto para abordar o problema da baixa resolução. . . . .	20
3.2	Quatro exemplos de J-espectros observados e seus melhores ajustes. . . . .	24
3.3	4 exmplos. . . . .	25
3.4	Exemplo de J-espectro observado e seus 5 melhores ajustes em $\chi^2$ . . . . .	26
3.5	Outro exemplo de J-espectro observado e seus 5 melhores ajustes em $\chi^2$ . . . . .	27
3.6	Figura retirada de Gallazzi et al. (2005) mostrando quatro exemplos de PDFs em relação à idade média. . . . .	29
3.7	Função distribuição de probabilidade para $\langle \log t_* \rangle_L$ do objeto da figura 3.4. . . . .	32
3.8	Função distribuição de probabilidade para $\log(N_{ii}/H_\alpha)$ do objeto da figura 3.4. . . . .	33



---

4.1	Rubik. . . . .	37
4.2	Diagrama cor/magnitude extraído de Baldry et al. (2004). . . . .	38
4.3	Diagrama cor/magnitude para a amostra de galáxias selecionadas como Base. . . . .	39
4.4	Diagrama WHAN e categorização para galáxias revisada. . . . .	41
4.5	Diagrama WHAN para a amostra de galáxias selecionadas como Base. . .	42
4.6	Dispersão da diferença entre Entrada versus média para diferentes fatores $f_L$ e diferentes propriedades para a base “diagrama cor-magnitude”. . . .	44
4.7	Dispersão da diferença entre Entrada versus média para diferentes fatores $f_L$ e diferentes propriedades para a base “WHAN”. . . . .	45
5.1	lalal . . . . .	52
5.2	lalal . . . . .	53
5.3	lalal . . . . .	54

# Lista de Tabelas

5.1	Propriedades analisadas, suas unidades e descrições. . . . .	48
5.2	Base CMD . . . . .	49
5.3	Base WHAN . . . . .	49
5.4	Base CMD . . . . .	50
5.5	Base WHAN . . . . .	51

# Capítulo 1

## Introdução

COMO OS SURVEYS MUDARAM O MODO DE SE FAZER ASTRONOMIA?

### 1.1 O STARLIGHT e o mega-levantamento astronômico SDSS

A tremenda expansão na taxa de produção de dados está promovendo uma revolução na astrofísica, proporcionando importantes avanços e ao mesmo tempo mudando radicalmente o próprio modo de fazer ciência nesta área. Observações provenientes de diversos telescópios, tanto na terra como no espaço, e cobrindo todo o espectro eletromagnético, estão gerando volumes de dados a uma taxa da ordem de petabytes por ano, com tendência crescente tanto em quantidade como em complexidade. O Sloan Digital Sky Survey (SDSS) York et al. (2000), considerada referência quando falamos em levantamentos modernos, em seus oito anos de funcionamento obteve imagens e espectros de um quarto do céu, mapeando tridimensionalmente mais de 930.000 galáxias e 120.000 quasares, sem contar estrelas, cometas, aglomerados de estrelas e outras observações indiretas. Sem dúvida, os mega-levantamentos vieram para mudar profundamente a maneira de se fazer astronomia em relação às técnicas tradicionais de coleta de dados via projetos de pedido de tempo de observação.

Nos últimos 5 anos, esta linha de pesquisa passou por uma forte transição, que resultou no hoje amplamente utilizado código STARLIGHT de síntese espectral de populações estelares. Originalmente desenvolvido por Cid Fernandes et al (2004) para o estudo de uma amostra de “apenas” 79 galáxias de tipo Seyfert 2, este código foi posteriormente

otimizado e adaptado para processar espectros de galáxias em geral (Cid Fernandes et al 2005, 2008). O STARLIGHT ajusta espectros observados Angstrom por Angstrom, utilizando combinações lineares de populações estelares simples, cujos espectros são previstos pela nova safra de modelos de síntese evolutiva de populações estelares inaugurada pelo trabalho seminal de Bruzual & Charlot (2003). Resumidamente, o código processa os dados levando-os de um espaço de observáveis (fluxos em cada  $\lambda$ ) a um espaço de propriedades físicas, como massa estelares, idades e metalicidades médias, profundidade óptica de poeira, dispersão de velocidades estelares e a própria história de formação estelar (SFH – Star Formation History) de galáxias são recuperados através desta análise.

Recuperar a SFH de galáxias a partir de espectros integrados tem sido uma espécie de “holy grail” da astronomia extragaláctica desde os trabalhos pioneiros de Tinsley e Faber nas décadas de 60 e 70, e de Bica e Alloin nos anos 80. A maioria dos métodos tradicionais para atacar este problema faz uso de índices espectrais (como cores e larguras equivalentes de linhas de absorção) escolhidos “a dedo” para representar um “resumo” do espectro total. Isto é verdade tanto para trabalhos “antigos” (eg, Bica 1988; Cid Fernandes et al 2001) como recentes (Kauffmann et al 2003; Gallazzi et al 2005), que se baseiam na modelagem de uns poucos índices. A pretensão de se extrair tanto a partir de tão pouco sempre foi vista com ceticismo. Ao modelar todo o espectro observado, pixel por pixel, o STARLIGHT suplantou esta limitação e abriu novas possibilidades. Maximizando o aproveitamento dos dados, maximizamos também a quantidade de informação extraída, recuperando com bom grau de detalhamento a SFH e até mesmo a evolução química das estrelas em uma galáxia a partir de seu espectro integrado


Os resultados obtidos são excelentes, e têm sido utilizados em diversos estudos, dentre os quais destacamos a série de artigos intitulada “Semi Empirical Analysis of Galaxies” (SEAGal): Cid Fernandes et al (2005); Mateus et al (2006); Stasinska et al (2006); Mateus et al (2007); Cid Fernandes et al (2007); Asari et al (2007); Stasinska et al (2008); Asari et al (2009); Cid Fernandes et al (2010). Esta série de artigos começa com uma apresentação do método e testes diversos, além de sua aplicação a uma amostra inicial de 50000 galáxias da SDSS. Já neste primeiro trabalho, em 2005, obtivemos resultados cientificamente relevantes, originais e deveras encorajadores, como ajustes espectrais de qualidade até então nunca antes obtidos, relações empíricas entre a extinção do gás e a extinção da luz estelar, e uma forte correlação entre a metalicidade do gás e das estrelas, ambas obtidas por primeira vez na literatura. Artigos seguintes exploraram temas correlatos e ampliaram a amostra a toda a SDSS-DR7 (926246 galáxias), uma empreitada que consumiu cerca de 200 processadores por mais de 2 meses e que muito

nos ensinou sobre como administrar de maneira eficiente volumes colossais de dados.

## 1.2 O mega-levantamento J-PAS e este trabalho

O projeto J-PAS (Javalambre Physics of the Accelerating Universe Astrophysical Survey <http://j-pas.org/>) é uma colaboração Hispano-Brasileira da qual o Grupo de Astrofísica da UFSC faz parte juntamente com mais de 70 outros pesquisadores e cientistas e que pretende, em 3 anos, mapear 8000 graus quadrados do céu em imagens a partir de 56 filtros de largura média de 125 Å. Espera-se que, ao final de suas atividades, este survey proporcione “espectros de baixa resolução” para cerca de 300 milhões de galáxias. Uma grande quantidade de dados que merece ser explorada na sua totalidade.

Nesta dissertação, realizaremos algumas experiências para retirada de informações sobre dados com o formato do J-PAS. Quando olhamos os dados de experimentos como o J-PAS, gostaríamos que os dados nos contassem informações sobre as populações estelares destas.

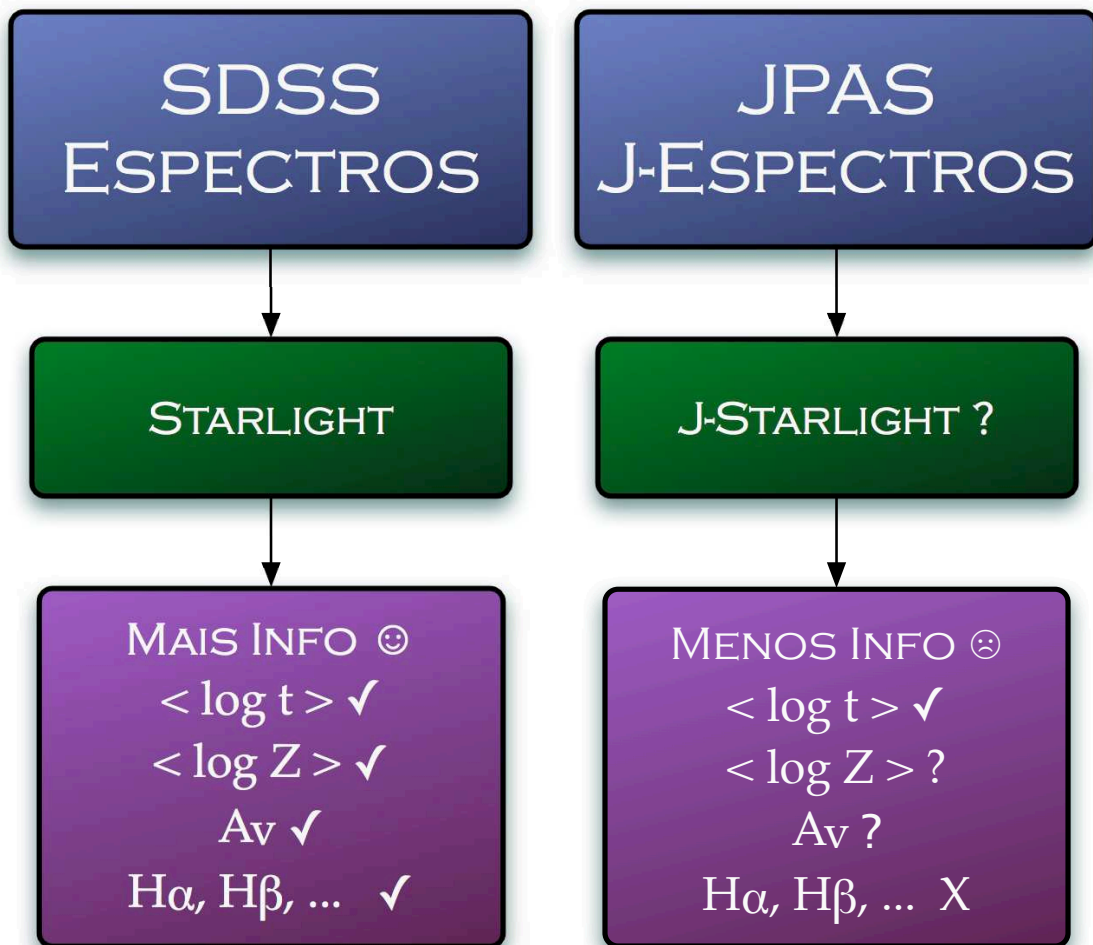
Uma das formas de abordar o problema (figura 1.1), seria realizar ajustes de populações estelares, como é feito pelo STARLIGHT: Utilizar-se de modelos de síntese evolutiva de populações estelares somados a  outros ingredientes para inferir sobre a História de Formação Estelar (SFH, em inglês) e medir as linhas de emissão no espectro residual.

Este é um método. Porém, ao realizar tal tipo de ajustes, forçamos sérias limitações proporcionais às limitações em resolução espectral do J-PAS:  $\Delta\lambda \approx 100\text{\AA}$  ou  $R = \frac{\Delta\lambda}{\lambda} \approx 50$ . Seria impossível medir razões de linhas de emissão em espectros com dada resolução espectral. Como também, separar algumas linhas de emissão como, por exemplo,  $[\text{N II}]\lambda 6584$  de  $\text{H}\alpha$ . Medidas de linhas de emissão no espectro residual do J-PAS seria praticamente impossível dadas estas limitações.

Exploramos aqui, uma alternativa que pode minimizar estas dificuldades ou, pelo menos, nos dar a mesma quantidade de informações que uma abordagem pelo ajuste de modelos similar à do STARLIGHT nos proporcionaria.

A ideia aqui seria utilizar o SDSS como uma grade referência para comparações ao invés de utilizar uma grade composta por modelos de populações estelares (como as SSPs de Bruzual & Charlot (2003) utilizadas pelo STARLIGHT).

A princípio, ao comparar o espectro de uma galáxia do J-PAS com uma do SDSS,



**Figura 1.1:** Esquema gráfico de um modo de análise para os dados J-PAS. Com uma adaptação do STARLIGHT aos J-Espectros não seria possível recuperar propriedades de galáxias com precisão comparável ao SDSS.

poderia-se inferir sobre propriedades que necessitariam de uma resolução espectral muito maior que a disponível, como a separação entre  $[\text{N II}]$  e  $\text{H}\alpha$  já mencionada, tornando inviável a classificação em SF/AGN através de diagramas de diagnóstico, medições de  $Z_{neb}$  e etc. Porém, se é possível encontrar um conjunto de espectros do SDSS com espectro do J-PAS (J-espectro) similar, então pode-se inferir sobre  $[\text{N II}]/\text{H}\alpha$ , por exemplo.

Este poderia ser um “turbinador” para a análise dos J-espectros, nos revelando informações indiretas para linhas de emissão, utilizando o J-espectro como um todo para as determinações, não somente o pixel onde a linha de emissão está centrada.

A pergunta central deste trabalho poderia ser: Até qual precisão pode-se inferir sobre as propriedades físicas de galáxias realizando uma comparação entre os dados do SDSS e do J-PAS? É o que tentaremos descobrir.

Antes de iniciar a abordagem a esta questão, tratar de uma forma mais quantitativa, resumindo aqui as quantidades que nos interessaria verificar quão precisamente poderíamos recuperar.

Assim, gostaríamos saber qual seria a precisão nas nossas medidas para:

- Idades Estelares, Absorção, Metalicidades Estelares e outras propriedades físicas;
- Larguras equivalentes de linhas de emissão (como  $\text{H}\alpha$ ,  $[\text{O II}]$ , ...), Razões entre linhas de emissão (como  $[\text{N II}]/\text{H}\alpha$ ,  $[\text{O II}]/\text{H}\beta$ , ...) outras propriedades observadas quando temos espectros de alta resolução.

E, ainda, como esta precisão varia com os seguintes parâmetros:

- escolha da amostra de comparação;
- escolha do esquema de comparação (pesos);
- redshift;
- Sinal-Ruído  $S/N$ ;
- Outras propriedades das galáxias ...

Tentaremos responder estas e algumas outras questões que surgirão ao longo de nosso estudo nos próximos capítulos.

## Capítulo 2

# Do SDSS ao J-PAS

**RESUMO:** Neste capítulo apresentamos brevemente as características dos levantamentos astronômicos SDSS e J-PAS (seções 2.1 e 2.2). Em seguida, apresentamos os cálculos necessários para converter um espectro como o do SDSS para um conjunto de filtros como os do J-PAS (seção 2.3) e os erros associados à cada ponto (seção 2.4). Por fim, na seção 2.5, mostramos como isto foi feito na prática e definimos o elemento básico que será trabalhado nos próximos capítulos: O J-Espectro.

### 2.1 O megalevantamento SDSS



O Sloan Digital Sky Survey (SDSS - York et al. (2000)) é um levantamento de dados astronômicos que, sem sombra de dúvidas, revolucionou a maneira de se fazer astronomia. Foi o primeiro grande levantamento que utilizou CCDs (Charged Couple Devices) ao invés de placas fotográficas para a obtenção de dados fotométricos e espectroscópicos. Além desta quantidade jamais vista de dados, o que leva este Survey a destacadamente ser um sucesso em comparação aos seus antecessores é a facilidade proporcionada no acesso aos seus dados através de uma interface de consulta dos dados (<http://cas.sdss.org/astrodr7/en/>). charge-coupled

O levantamento utiliza um telescópio exclusivo de 2,5m localizado no Apache Point Observatory (Novo México, EUA). Este é equipado com uma câmera CCD, dedicada ao imageamento do céu em 5 bandas fotométricas (u, g, r, i e z), e dois espectrógrafos multifibras com capacidade de realizar espectroscopia de até 640 objetos simultaneamente com uma resolução média de 3Å. No Data-Release 7 (DR7) foram observadas espectro-

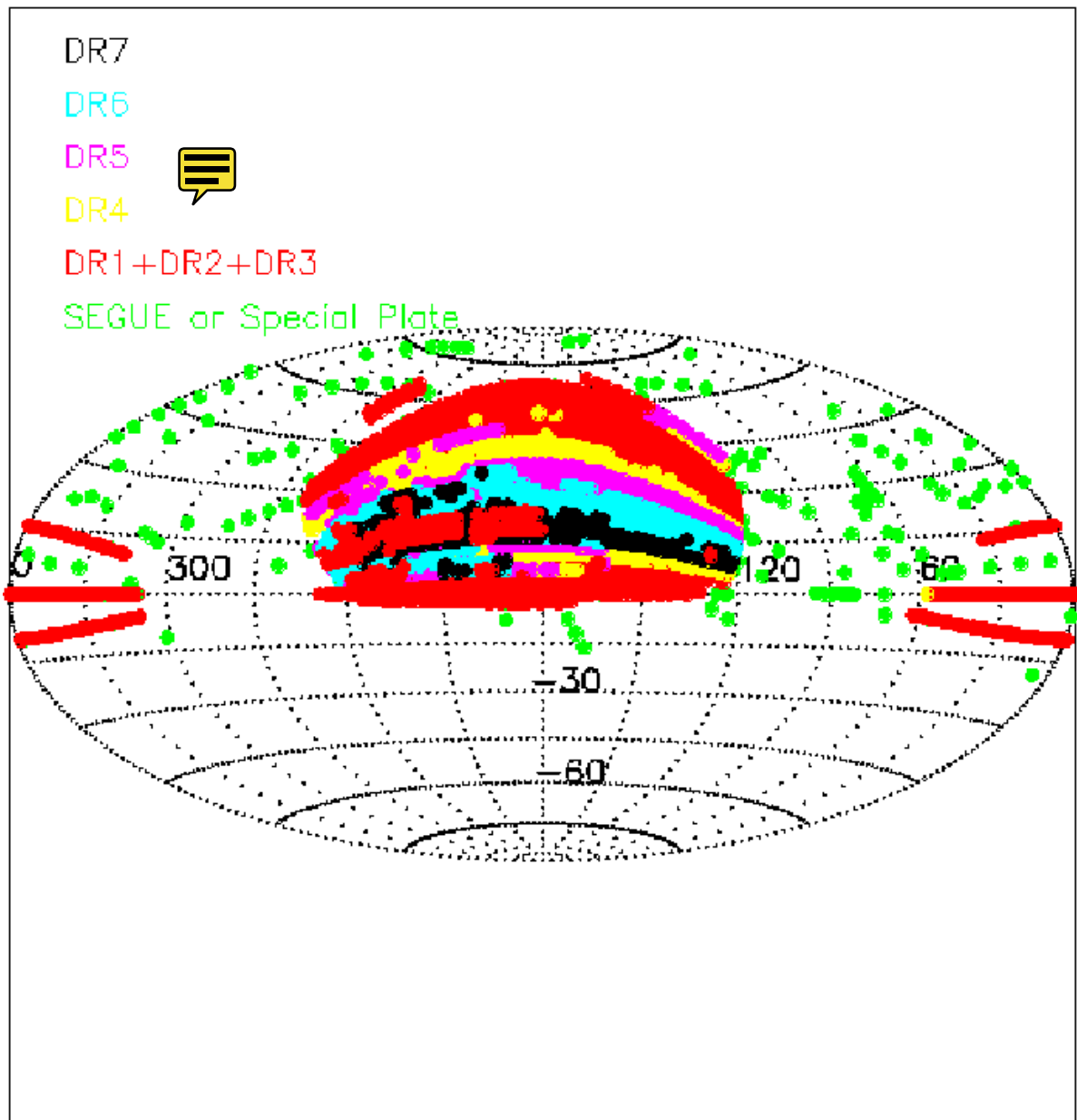


scopicamente 929555 galáxias espalhadas em uma região do céu correspondente a 9380 graus quadrados (figura 2.1) em uma faixa espectral que varia entre 3800 e 9200 Å e magnitude na banda r menor que 17.77.

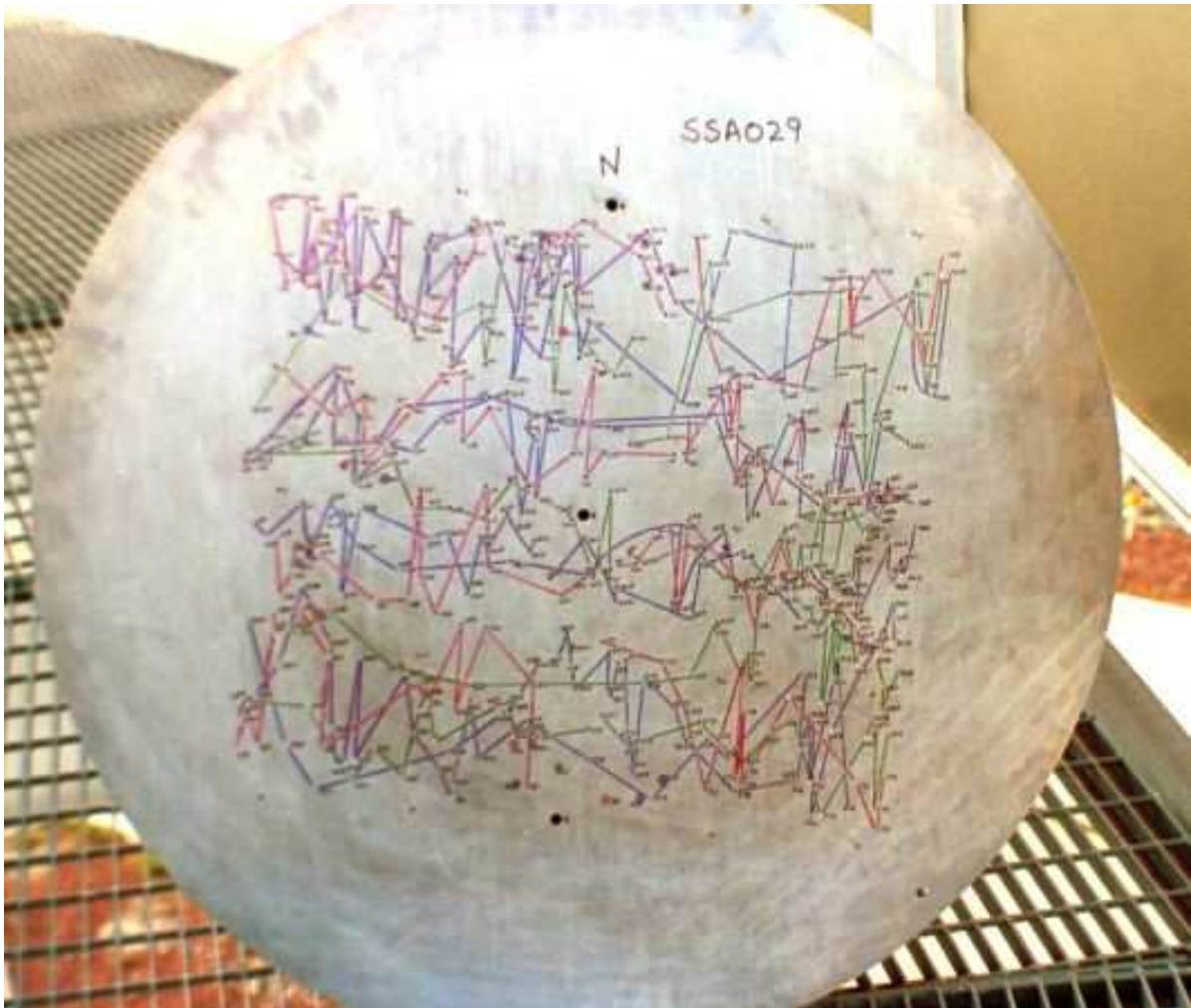
Os campos são primeiramente observados fotometricamente a fim de detectar em quais objetos será realizada a espectroscopia. Então, placas de alumínio (figura 2.2) são perfuradas com as posições exatas onde as fibras deverão estar localizadas, as fibras são conectadas ao espectrógrafo e os espectros são coletados.



Para este trabalho, utilizamos o espectro de 926246 destas galáxias reamostradas em 1Å e no referencial de repouso utilizando os referenciais de repouso determinados pela pipeline do SDSS. Do SDSS, foram extraídos os fluxos para cada  $\lambda$ , os erros associados a cada um destes fluxos e uma **flag** que indica regiões com problemas no espectro.



**Figura 2.1:** Projeção equatorial das regiões cobertas pelo levantamento espectroscópico do SDSS para seus diversos Data-Releases. Neste trabalho, os dados utilizados foram os do DR7, que engloba seus predecessores. Figura extraída de <http://www.sdss.org/>



**Figura 2.2:** Exemplo de um prato de fibras do SDSS. Após fotografar o campo com um CCD, os pratos são marcados por um software e furados a fim de localizar as 640 fibras do espectrógrafo nos lugares correspondentes aos objetos que deseje-se analisar. Figura extraída de <http://www.sdss.org/>

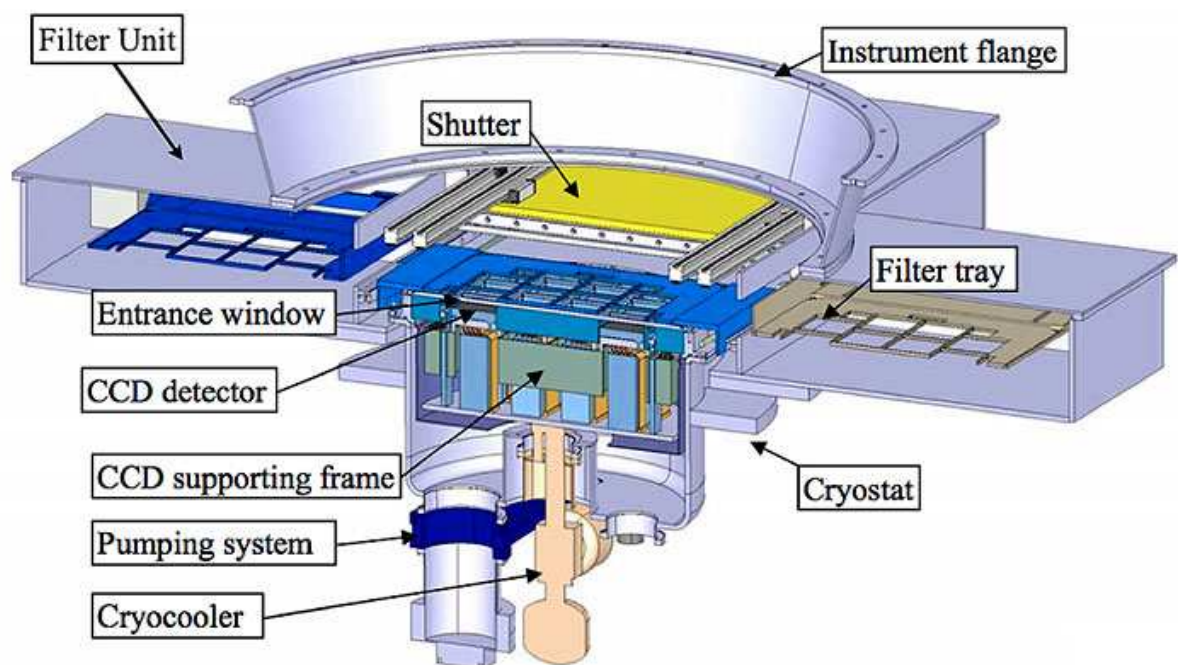
## 2.2 O megalevantamento J-PAS

O J-PAS (**Javalambre Physics of Accelerating Universe Astrophysical Survey**, <http://j-pas.org/>) é um projeto fruto de uma colaboração entre instituições Brasileiras e Espanholas que visa, utilizando-se um telescópio de 2,52m, no recém-criado Observatório Astronómico Javalambre em Teruel (OAJ) na Espanha, fotografar com um mosaico de 14 CCDs (figura 2.3) 8000 graus quadrados do céu em 56 filtros (figura 2.4). **Este levantamento, diferentemente de seus predecessores, mapeará o céu em uma quantidade superior de filtros de largura em comprimento de onda muito estreitos (da ordem de 125 Å).**

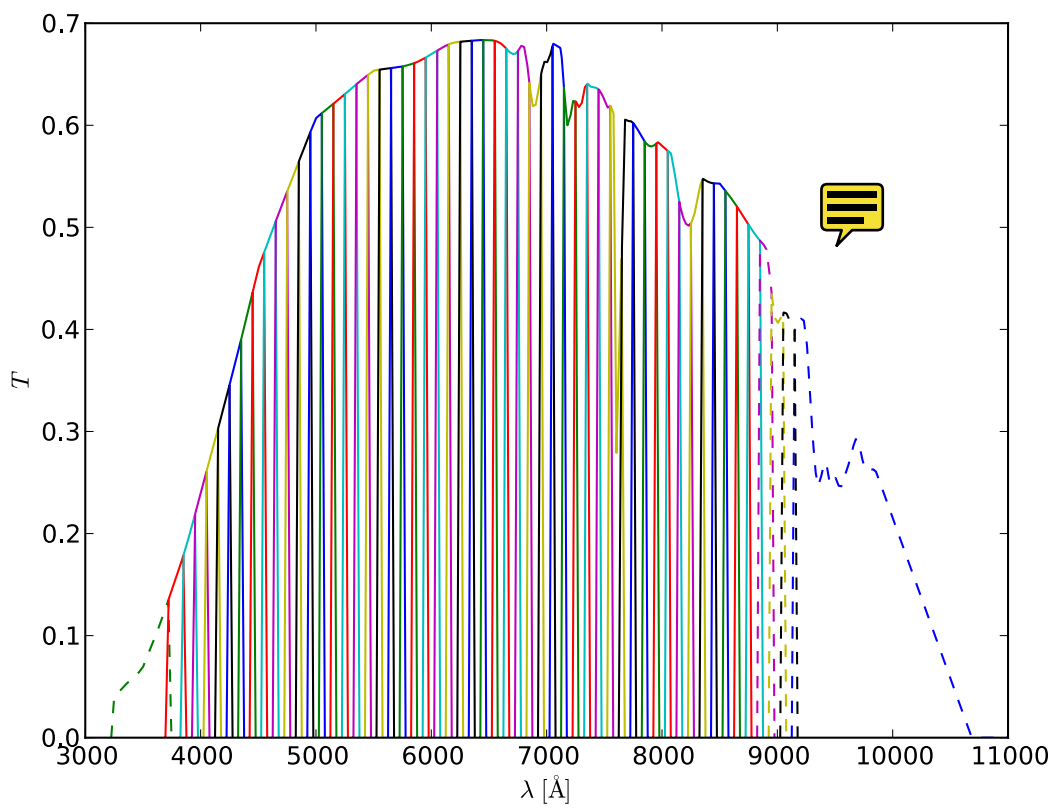
esta parte tá estranha

Como o produto das observações são imagens e não espectros, perde-se na resolução espectral, quando comparada aos 3Å do SDSS, porém ganha-se resolução espacial pois ao invés de confinar toda a luz de 2" **segundos de arco** em uma fibra, teremos a informação pixel a pixel de cada galáxia onde cada pixel corresponde a aproximadamente 0,23". Ainda, nestes 8000 graus, espera-se imagear 300 milhões de galáxias, ou seja, 300 vezes mais que no SDSS. Isto o torna um grande laboratório para estudos estatísticos das propriedades físicas e de linhas de emissão de galáxias como propomos neste trabalho.





**Figura 2.3:** Desenho conceitual da **JPCAM** que equipará o telescópio de 2,52m no OAJ. Figura extraída de <http://j-pas.org/>.





**Figura 2.4:** Curvas de transmissividade utilizadas neste trabalho. Cada curva colorida representa um dos 56 filtros de  $\approx 125 \text{ \AA}$  do J-PAS. Para este trabalho, devido à faixa de cobertura espectral do SDSS, foram excluídos os 4 filtros marcados em linha pontilhada.

## 2.3 Conversão de Espectros do SDSS para o J-PAS

Para realizarmos a conversão de um espectro do SDSS (e os erros associados a este) a dados fotométricos do J-PAS, utilizamos a seguinte relação:

$$J_l = \frac{\int O_\lambda T_{l,\lambda} d\lambda}{\int T_{l,\lambda} d\lambda} = \sum_\lambda O_\lambda t_{l,\lambda} \Delta\lambda \quad (2.1)$$

onde  $O_\lambda$  é o espectro observado no SDSS,  $T_{l,\lambda}$  é a curva de transmissividade para um determinado filtro  $l$  e  $t_{l,\lambda} \equiv \frac{T_{l,\lambda}}{\sum_\lambda T_{l,\lambda} \Delta\lambda}$ , tal que  $J_l$  é o fluxo em um dado filtro  $l$ . Definidos dessa maneira,  $J_l$  e  $O_\lambda$  têm a mesma unidade, e.g.  $\text{ergs s}^{-1} \text{cm}^{-2} \text{\AA}^{-1}$ . Exemplos desta conversão são mostrados na figura 2.5.

Para calcularmos os espectros acima, os espectros do SDSS foram trazidos para o referencial de repouso e corrigidos pelo?  para o avermelhamento Galáctico. E, para simplificação, utilizaremos todos espectros no referencial de repouso ( $z = 0$ ). Esperamos que a partir dos redshifts fotométricos calculados pelos métodos tradicionais Benítez (2000)  o ajuste deste termo seja desnecessário. Todavia, pode ser um teste interessante incluir futuramente em nosso cálculo e verificar qual a precisão obtida.

## 2.4 Espectros de erros para o espectro J-PAS

Para calcularmos o erro correspondente ao fluxo  $J_l$  em cada filtro, utilizamos o erro no espectro observado no SDSS,  $n_\lambda$ . A matemática desta transformação é a seguinte:

Temos nosso espectro observado  $O_\lambda$  como uma soma do espectro intrínseco da galáxia  $O_\lambda^0$  mais o ruído  $n_\lambda$ :

$$O_\lambda = O_\lambda^0 + n_\lambda \quad (2.2)$$

e,

$$O_\lambda^2 = (O_\lambda^0 + n_\lambda)^2 = O_\lambda^{0^2} + 2O_\lambda^0 n_\lambda + n_\lambda^2 \quad (2.3)$$

Assumindo que o valor esperado do ruído seja nulo  $E(n_\lambda) = 0$  e que a variância do ruído de  $O_\lambda$ ,  $\sigma^2(n_\lambda)$  seja conhecida, podemos calcular a variância no J-espectro a partir da relação:

cuidado com os espaços entre texto e equações

$$\sigma^2(J_l) = E(J_l^2) - E^2(J_l) \quad (2.4)$$

onde  $E(x)$  é o valor esperado de  $x$ .

Podemos, com a equação 2.1, calcular o valor esperado de  $J_l$ :

$$\begin{aligned} E(J_l) &= E\left(\sum_{\lambda} O_{\lambda} t_{l,\lambda} \Delta\lambda\right) \\ &= E\left(\sum_{\lambda} (O_{\lambda}^0 + n_{\lambda}) t_{l,\lambda} \Delta\lambda\right) \\ &= E\left(\sum_{\lambda} O_{\lambda}^0 t_{l,\lambda} \Delta\lambda\right) + E\left(\sum_{\lambda} n_{\lambda} t_{l,\lambda} \Delta\lambda\right) \\ &= J_l^0 + \sum_{\lambda} E(n_{\lambda}) t_{l,\lambda} \Delta\lambda = J_l^0 \end{aligned} \quad (2.5)$$

onde usamos o fato de que  $E(n_{\lambda}) = 0$ . De modo similar, podemos calcular ~~e~~ o valor esperado para  $J_l^2$ :

$$\begin{aligned} E(J_l^2) &= E\left(\sum_i (O_i^0 + n_i) t_{l,i} \Delta\lambda \sum_j (O_j^0 + n_j) t_{l,j} \Delta\lambda\right) \\ &= E\left(\left(\sum_i O_i^0 t_{l,i} \Delta\lambda + \sum_i n_i t_{l,i} \Delta\lambda\right) \left(\sum_j O_j^0 t_{l,j} \Delta\lambda + \sum_j n_j t_{l,j} \Delta\lambda\right)\right) \\ &= J_l^{02} + 2J_l^0 E\left(\sum_i n_i t_{l,i} \Delta\lambda\right) + E\left(\sum_i n_i t_{l,i} \Delta\lambda \sum_j n_j t_{l,j} \Delta\lambda\right) \\ &= J_l^{02} + E\left(\sum_i n_i t_{l,i} \Delta\lambda \sum_j n_j t_{l,j} \Delta\lambda\right) \end{aligned} \quad (2.6)$$

O último termo, pode ser dividido em duas somas. Uma para  $i = j$  e outra para  $i \neq j$ . Assim,



$$\begin{aligned}
E\left(\sum_i n_i t_{l,i} \Delta\lambda \sum_j n_j t_{l,j} \Delta\lambda\right) &= E\left(\sum_i n_i^2 t_{l,i}^2 \Delta\lambda^2 + \sum_i n_i t_{l,i} \Delta\lambda \sum_{j \neq i} n_j t_{l,j} \Delta\lambda\right) \\
&= \sum_i E(n_i^2) t_{l,i}^2 \Delta\lambda^2 + \sum_i \sum_{j \neq i} E(n_i n_j) t_{l,i} t_{l,j} \Delta\lambda^2 \quad (2.7)
\end{aligned}$$

Assumindo que os rúidos  $n_i$  e  $n_j$  não sejam correlacionados na janela do filtro,  $n_i$  é independente de  $n_j$ , então o segundo termo da equação 2.7 é igual a zero. Assim,

$$E(J_l^2) = J_l^{02} + \sum_i n_i^2 t_{l,i}^2 \Delta\lambda^2 \quad (2.8)$$

Substituindo as equações 2.5 e 2.6 em 2.4, com podemos encontrar uma expressão para  $\sigma_l^2$ :

espaço demais

$$\sigma^2(J_l) = E(J_l^2) - E(J_l)^2 = (J_l^{02} - \sigma^2) - J_l^2 = \sigma_l^2 \quad (2.9)$$

$$\begin{aligned}
\sigma_l^2 &= N_{\lambda,l} \langle \sigma(n_\lambda)^2 t_{l,\lambda}^2 \Delta\lambda^2 \rangle \\
&= N_{\lambda,l} \Delta\lambda \langle \sigma(n_\lambda)^2 t_{l,\lambda}^2 \rangle \Delta\lambda \\
&= \Lambda_l \langle \sigma(n_\lambda)^2 t_{l,\lambda}^2 \rangle \Delta\lambda \quad (2.10)
\end{aligned}$$

onde  $\Lambda_l \equiv N_{\lambda,l} \Delta\lambda$  é o tamanho da janela espectral do filtro  $l$ .

Para termos uma ideia aproximada de como o erro em  $J_l$  depende do erro em  $O_\lambda$  podemos considerar um conjunto de filtros retangulares, para os quais a função de transmissividade é simplesmente  $t_{l,\lambda} = \frac{1}{\Lambda_l}$ . Neste caso temos:

$$\begin{aligned}
\sigma_l^2 &= \Lambda_l \Delta\lambda \langle \sigma^2(n_\lambda) t_{l,\lambda}^2 \rangle \\
&= \frac{\Delta\lambda}{\Lambda_l} \langle \sigma^2(n_\lambda) \rangle
\end{aligned}$$

Substituindo os valores típicos  $\Lambda_l^{\text{JPAS}} \approx 100 \text{ \AA}$  e  $\Delta\lambda^{\text{SDSS}} \approx 3 \text{ \AA}$ , temos:

$$\sigma_l = \sqrt{\frac{3}{100} \langle \sigma^2(n_\lambda) \rangle}$$

Esperamos, então, que nossos erros em cada filtro do JPAS sejam da ordem de  $\sqrt{\frac{100}{3}} \approx 6$  vezes menor que em cada pixel do SDSS na mesma faixa espectral.

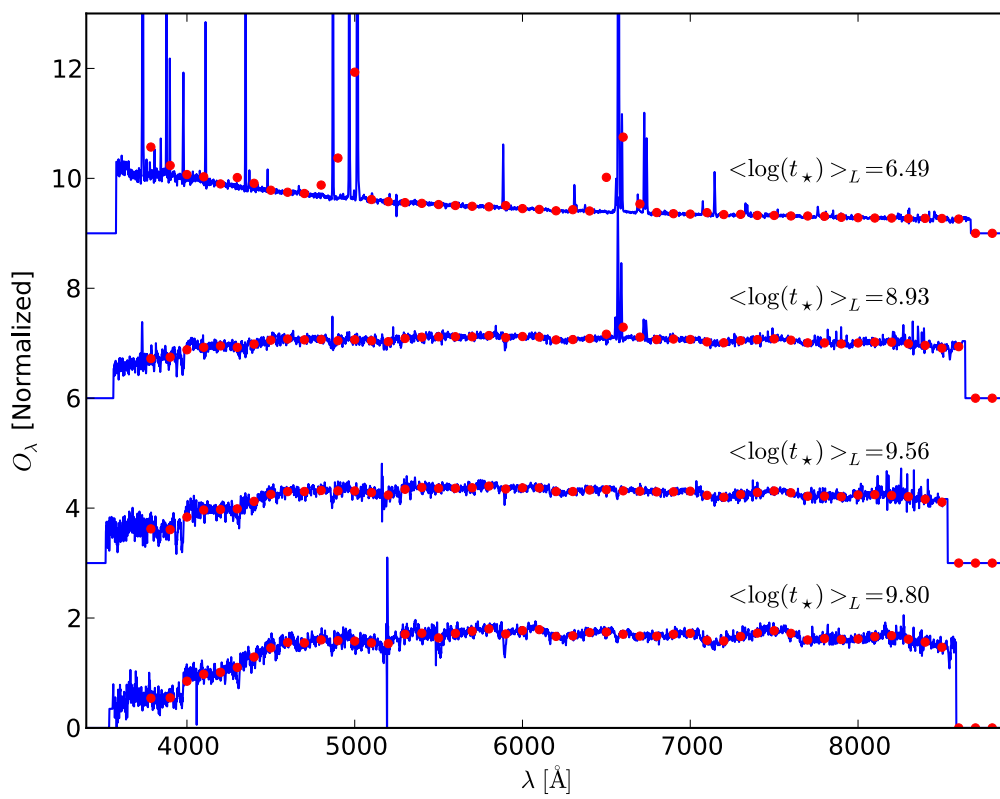
## 2.5 O J-Espectro

Partindo das equações 2.1 e 2.10, podemos degradar os espectros do SDSS para simular os J-espectros. Porém, para converter estes espectros, temos que resolver a situação de quando não possuímos um ou mais valores para  $O_\lambda$ . Neste caso, adotamos a seguinte receita em cada filtro do JPAS:

- Caso o número de pontos com problemas seja superior a 50% do número de  $\lambda$ s no filtro, este J-pixel é descartado nas análises.
- Caso o número de pontos com problemas seja inferior ou igual a 50% do número de  $\lambda$ s no filtro, usamos o fluxo observado do SDSS para os pontos sem problemas e o fluxo sintético do STARLIGHT para os pontos com problemas. No cálculo do J-erro, pontos com problemas são negligenciados na equação 2.10.

ignorados?

Assim, transformamos todo os espectros do SDSS-DR7 para J-espectros e utilizamos como base para este trabalho.



**Figura 2.5:** Exemplos de J-espectros normalizados distribuídos em  $\langle \log t_\star \rangle_L$ . Em azul o espectro observado pelo SDSS e em vermelho o J-espectro correspondente.

# Capítulo 3

## Método

**RESUMO:** *Este capítulo apresenta o método de casamento espectral adotado para extrair informação dos J-espectros de forma indireta, utilizando resultados obtidos através dos SDSS pela aplicação do STARLIGHT. Começamos com uma descrição da filosofia geral do método (seção 3.1) e em seguida definimos a notação usada para os ingredientes básicos envolvidos no casamento. Na seção 3.3 apresentamos a figura de mérito utilizada para o casamento espectral bem como a receita adotada para dar pesos aos píxeis de cada J-espectro. Seguido disso, apresentamos alguns exemplos de casamentos espectrais na seção 3.4 e finalizamos na seção 3.5 apresentando a abordagem probabilística utilizada para inferir propriedades físicas e observacionais a partir dos resultados dos casamentos espectrais.*

### 3.1 Introdução

O SDSS sem dúvida foi divisor de águas na astronomia como um todo tanto pela quantidade de dados quanto à forma com a qual estes foram reduzidos, organizados e disponibilizados. Nosso grupo, com a aplicação do STARLIGHT para 926246 galáxias, pode ao longo dos últimos anos construir (e disponibilizar, <http://starlight.ufsc.br/>) um banco de dados para uma infinidade de propriedades físicas (massa estelar, extinção, idade média e outros) e medidas de fluxos e larguras equivalentes de linhas de emissão com grande precisão (Asari et al. (2007)).

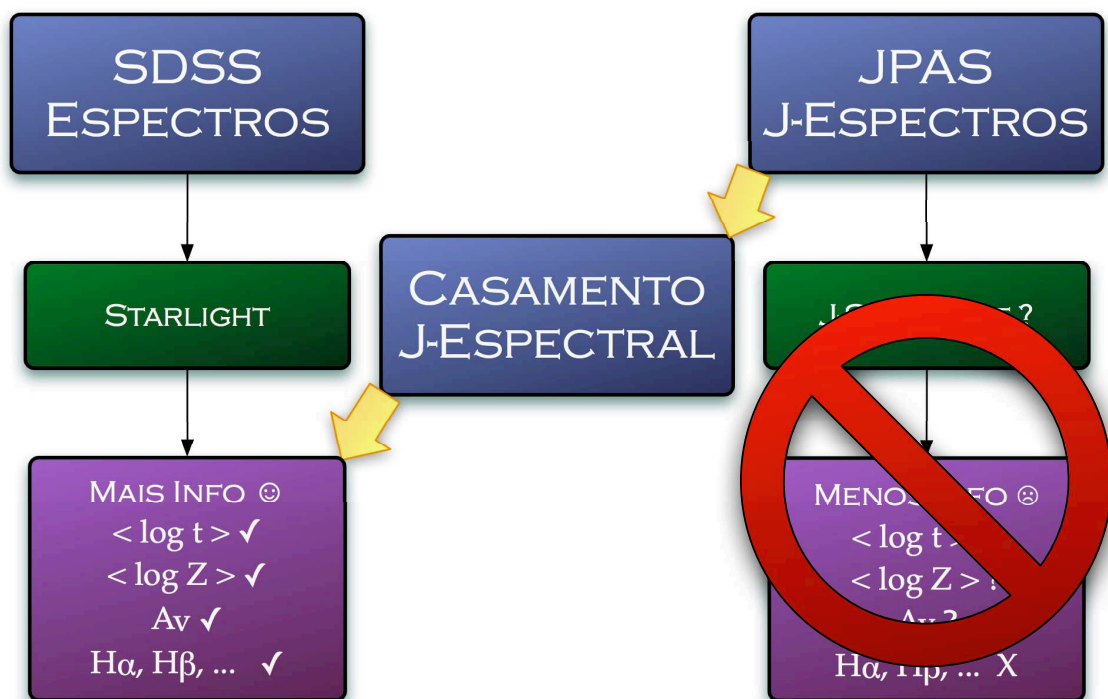
No capítulo 1 vimos quão bem este maquinário funciona e, na figura 1.1, especulamos que talvez seja possível para os dados do J-PAS seguir uma rota similar à aplicada por nós

ao SDSS. Porém, devido à baixa resolução espectral dos dados, esta abordagem direta aos dados nos traria muito menos informação. Por exemplo é intuitivamente óbvio que não poderemos extrair histórias de formação estelar de J-espectros com o mesmo grau de detalhe que fazemos no SDSS. E, muito menos, medições diretas de linhas de emissão. Será impossível, por exemplo, separar as linhas de [N II] e H $\alpha$  em J-espectros, o que acarreta sérias limitações em termos de muitas aplicações astrofísicas.

Assim, a ideia chave deste trabalho é tomar um atalho que nos remeta às propriedades recuperadas do SDSS partindo dos dados do J-PAS a partir do casamento espectral, como esquematizado na figura 3.1.

Assumimos que se duas galáxias possuem fluxos em 56 filtros estreitos iguais ou muito parecidos elas devem ter seus espectros em média-alta resolução também muito parecidos e extrapolamos isto para que, se os espectros são parecidos, suas propriedades físicas e linhas de emissão também serão.

Esta ideia é uma variação de uma metodologia muito utilizada hoje em dia, como por exemplo, por vários autores que desenvolveram códigos que comparam alguns índices espectrais a conjuntos de modelos gerados a partir de cálculos de síntese evolutiva (e.g. Bruzual & Charlot (2003)) e, também como o que fazem Gallazzi et al. (2005). No nosso caso, não compararemos a modelos mas sim a espectros de fato observados o que nos traz uma série de vantagens como a possibilidade de da medida indireta de linhas de emissão que veremos nos próximos capítulos.



**Figura 3.1:** Esquema gráfico do método proposto para abordar o problema da baixa resolução. Ao invés de adaptar o STARLIGHT para os dados do J-PAS, propomos o casamento espectral para obter propriedades observadas somente em uma resolução maior.

## 3.2 Definições

Os ingredientes básicos desta análise são um J-espectro observado  $O$  e seus erros associados, uma base  $B$  de J-espectros de referência para o casamento espectral e um conjunto  $P$  de propriedades a serem estimadas.

~~Assim que~~, a fim de facilitar a exposição do método nas próximas seções, é didático definirmos uma notação para estes ingredientes da seguinte forma:

- $O_{l,i}$ : É o fluxo no filtro  $l$  de um Objeto  $i$  para uma amostra de objetos. Estes simularão galáxias observadas pelo J-PAS seguindo a transformação da seção 2.3.
- $B_{l,j}$ : É o fluxo no filtro  $l$  de um elemento  $j$  de uma Base de espectros para referência. Seguem a mesma transformação.
- $\sigma(B_{l,j})$  e  $\sigma(O_{l,i})$ : Erros associados aos J-fluxos em cada elemento da Base e dos Objetos conforme definidos na seção 2.4
- $M_{l,i,j}$ : É o fluxo no filtro  $l$  de um Modelo para Objeto  $i$  construído a partir do elemento  $j$  da Base. O modelo é dado por  $M_{l,i,j} = a_{i,j}B_{l,j}$ , onde  $a_{i,j}$  é um fator de escala a ser otimizado.

## 3.3 Casamento Espectral: Figura de mérito e pesos

Para o casamento entre os pares de J-espectros calculamos, em conjunto, o  $\chi^2_{i,j}$  na sua forma tradicional:

$$\chi^2_{i,j} = \sum_l (O_{l,i} - a_{i,j}B_{l,j})^2 w_{l,i,j}^2 \quad (3.1)$$

onde fator de escala  $a_{i,j}$  é obtido através da minimização do  $\chi^2$  pela expressão:

$$\frac{\partial \chi^2}{\partial a_{i,j}} = \frac{\partial \sum_l (O_{l,i} - a_{i,j}B_{l,j})^2 w_{l,i,j}^2}{\partial a_{i,j}} = 0$$

o que nos dá:

$$a_{i,j} = \frac{\sum_l O_{l,i} B_{l,j} w_{l,i,j}^2}{\sum_l B_{l,j}^2 w_{l,i,j}^2} \quad (3.2)$$

Há várias maneiras para a escolha dos pesos  $w_{l,i,j}$ , dependendo de quais características gostaríamos de dar mais importância no casamento. A mais tradicional seria,

$$w_{l,i,j}^2 = \frac{1}{\sigma^2(O_{l,i}) + a_{i,j}^2 \sigma^2(B_{l,i})} \quad (3.3)$$

onde as incertezas tanto em  $O_{l,i}$  quanto em  $M_{l,i}$  são consideradas.

Todavia, preferimos uma versão levemente diferente, dada por:

$$w_{l,i,j}^2 = \left( \frac{\langle O_{l,i}^2 \rangle}{O_{l,i}^2} \right) \left( \frac{1}{\sigma^2(O_{l,i}) + a_{i,j}^2 \sigma^2(B_{l,i})} \right) \quad (3.4)$$

onde o segundo termo é dado pela equação 3.3 e o denominador do primeiro termo é tal que dá um peso igual às regiões com maior e menor fluxos. O numerador deste foi inserido a fim de manter a adimensionalidade do peso.

Para entender melhor a motivação para a introdução deste termo **ad-hoc**, note que ele faz com que a diferença quadrática entre  $O_{l,i}$  e o modelo  $M_{l,j}$  em um filtro  $l$  seja proporcional a  $[(O_{l,i} - M_{l,j})/O_{l,i}]^2$ . Desta maneira, uma diferença no pico de uma linha de emissão (alto  $O_{l,i}$ ) e no fundo de uma linha de absorção (baixo  $O_{l,i}$ ) contribuem igualmente ao  $\chi^2$  (afora, é claro, os outros fatores envolvidos no segundo termo em 3.4). Julgamos esta uma estratégia mais adequada, mas cabe notar que esta decisão no fundo não é tão relevante, uma vez que resultados muito parecidos são obtidos seguindo a receita padrão da equação 3.3.

O leitor mais atento percebe que o erro dado pela equação 3.4 depende de  $a_{i,j}$  que é dado pela equação 3.2 que, por sua vez, é dependente de  $w_{i,j}$ . Para resolver este fato, no cálculo do  $\chi^2$  foram realizadas iterações no cálculo de  $a_{i,j}$  e  $w_{l,i,j}$  até houvesse convergência. Em geral, não foram necessárias mais do que 3 ou 4 iterações para que o



critério de convergência por nós estabelecido ( $\Delta a_{i,j} < 10^{-5}$ ) fosse satisfeito.

Ao fim deste processo temos, para cada Objeto  $O_i$ , um vetor  $\chi_{i,j}^2$ , um vetor  $N_{i,j}$  dado pelo número de píxeis considerados bons no J-espectro do Objeto  $O_i$  e do  $j$ -ésimo elemento de Base  $B_j$  simultaneamente e um vetor com os fatores de escala  $a_{i,j}$  que melhor ajusta o Objeto  $O_i$  ao elemento de Base  $B_j$ . Com isso, calculamos o  $\chi_{i,j,\text{RED}}^2$ :

$$\chi_{i,j,\text{RED}}^2 = \frac{1}{N_{i,j}} \sum_l (O_{l,i} - a_{i,j} B_{l,j})^2 w_{l,i,j}^2 \quad (3.5)$$

que é chamado de  $\chi^2$  reduzido.

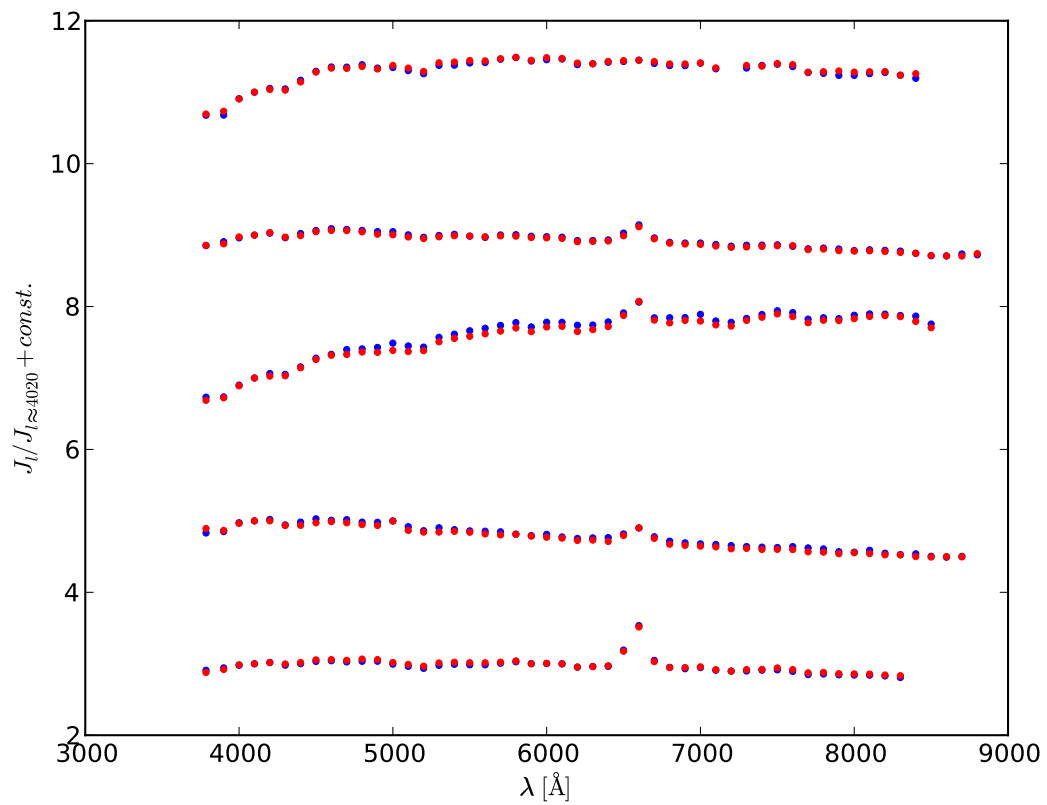
### 3.4 Casamento Espectral: Exemplos

A versão mais simples do casamento espectral é o **pareamento** de galáxias onde pegamos o elemento  $j$  da Base que tem o menor  $\chi^2$  em relação ao Objeto  $i$  e, então, supõe-se que a galáxia  $j$  é idêntica à  $i$  e que, portanto, podemos dizer em uma aproximação grosseira que todas as propriedades de  $i$  são idênticas às propriedades de  $j$  (figuras 3.2 e 3.3).

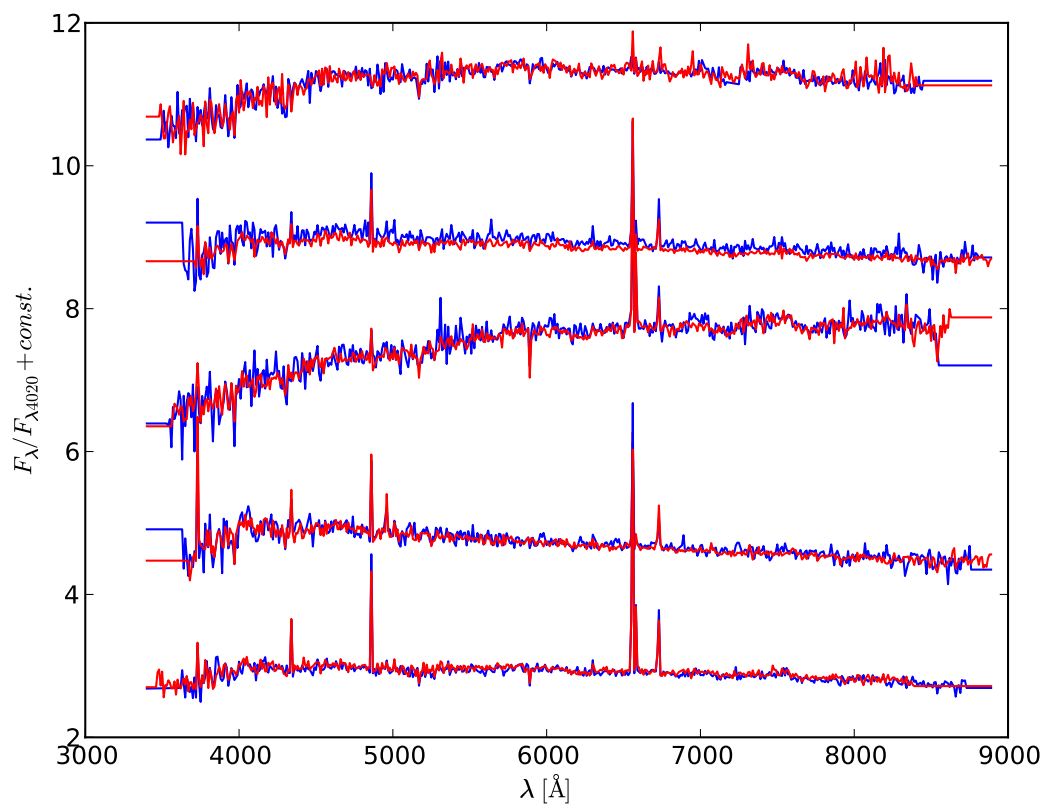
Todavia, quando olhamos além do menor  $\chi^2$ , observamos rapidamente que o pareamento não é o melhor método já que, quando ordenamos por  $\chi^2$  e observamos (figuras 3.4 e 3.5), constatamos que na própria base existem galáxias muito semelhantes e com propriedades muitas vezes diferentes entre si.

Ao observarmos tamanha semelhança entre os espectros que **compõe** os melhores ajustes, nos veio a ideia de, ao invés de escolher um único par Objeto-Base para definir as propriedades de determinado Objeto, dar um enfoque estatístico onde todas os elementos da Base são levados em conta proporcionalmente à semelhança com o J-espectro Observado. Decidimos então a dar um enfoque análogo ao adotado por Kauffmann et al. (2003), Brinchmann et al. (2004) e Gallazzi et al. (2005) exceto que, no caso deles, a comparação é feita a partir de modelos teóricos enquanto a nossa é eminentemente empírica.

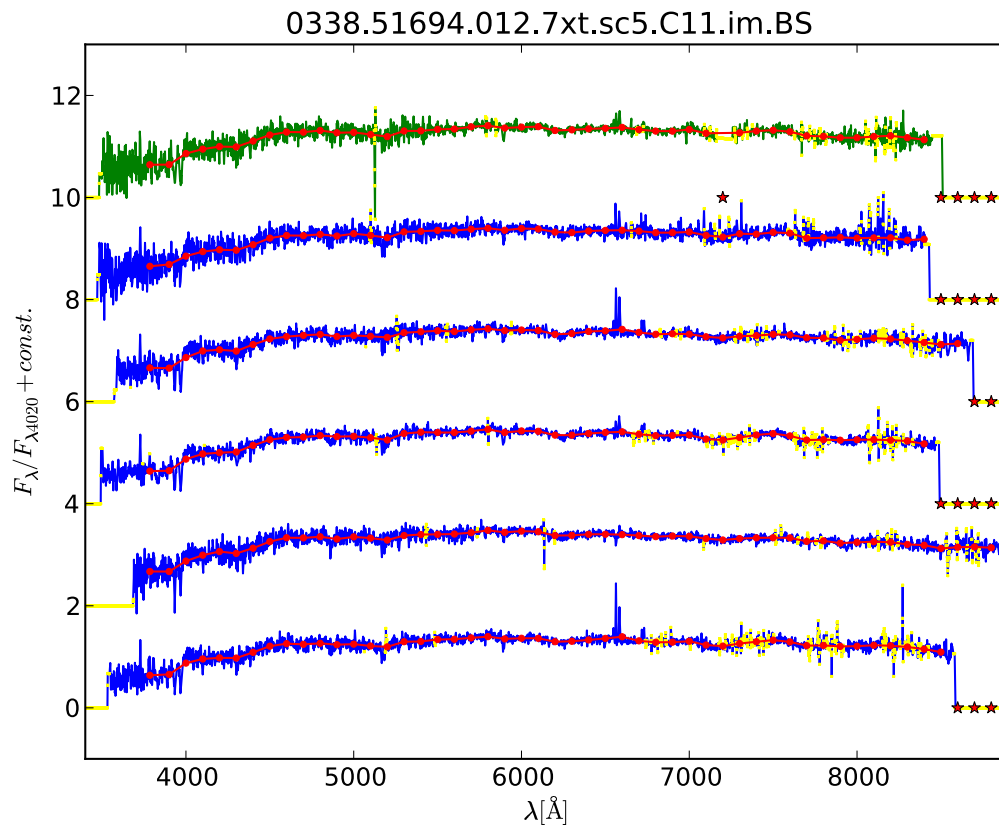
Esta técnica nos permite medir muita coisa, em particular, coisas que não conseguiríamos medir de forma direta no J-espectro, como separar [N II] e H $\alpha$ , o que nos permite inferir sobre medidas de larguras equivalentes e razões de linhas de emissão



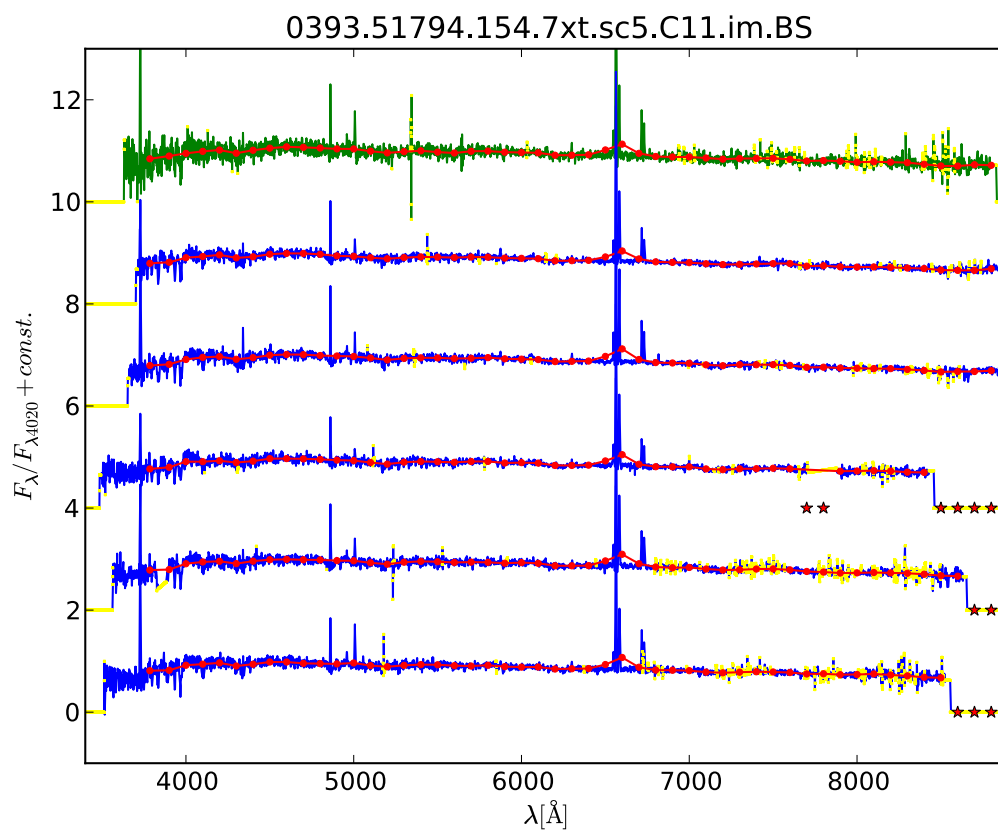
**Figura 3.2:** Quatro exemplos de J-espectros observados (em azul) e seus primeiros melhores ajustes (em vermelho).



**Figura 3.3:** Espectros suavizados em  $10 \text{ \AA}$  dos exemplos da figura 3.2. Nesta figura pode-se evidenciar qualitativamente a precisão do ajuste, principalmente em relação às linhas de emissão.



**Figura 3.4:** Exemplo de J-espectro observado de uma galáxia de  $\langle \log t \rangle_L = 9.31$  e seus 5 melhores ajustes em  $\chi^2$ . De cima para baixo, temos em verde o objeto e, em azul, os 5 melhores ajustes a este J-espectro. Em amarelo estão marcados os pontos que não foram levados em conta por estarem mascarados pela pipeline do SDSS, devido a problemas. Em vermelho, o J-espectro correspondente. As estrelas indicam pontos do J-espectro não foi considerado por problemas no espectro do SDSS.



**Figura 3.5:** Outro exemplo de galáxia e seus 5 melhores pares encontrados, mas desta vez para um objeto mais jovem, com  $\langle \log t \rangle_L = 8.50$ .

que são importantes para, por exemplo, a classificação de galáxias usando diagramas de diagnóstico.

A comparação realizada nesta seção foi somente uma demonstração de como podemos encontrar pares (e até conjuntos) de galáxias semelhantes, o que nos encoraja a seguir para a próxima seção, onde desenvolvemos um formalismo mais rigoroso que nos permite fazer inferências sobre propriedades físicas e linhas de emissão.


## 3.5 Formalismo Estatístico

Dado um determinado Objeto e uma determinada Base, gostaríamos de inferir sobre os valores das propriedades deste Objeto como, por exemplo, as larguras equivalentes de linhas de emissão, idade média das populações estelares, razões de medidas de linhas de emissão e outros.

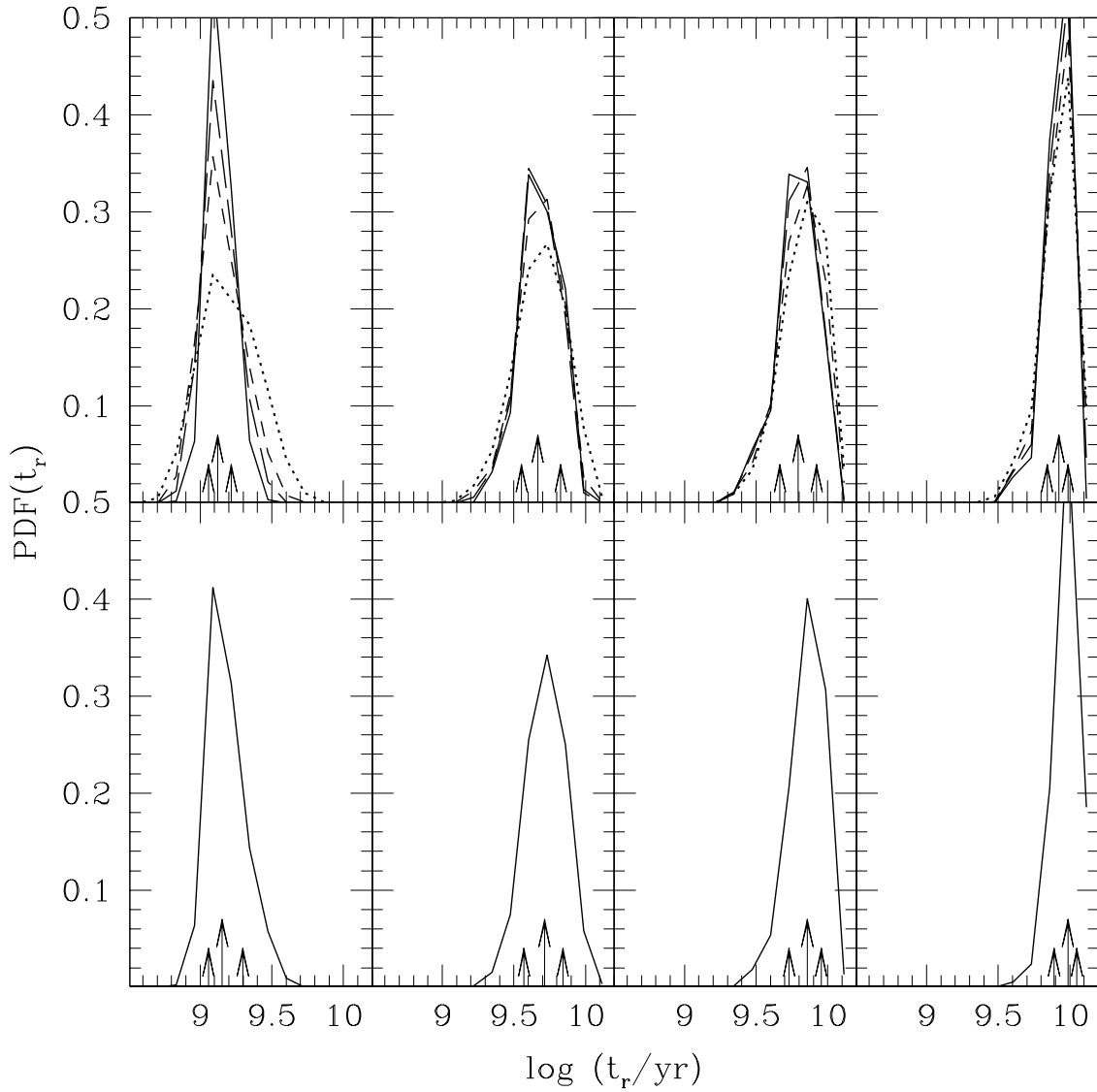
Com o valor de  $\chi_{i,j}^2$  para todos os elementos da Base e cada Objeto, podemos calcular uma função do tipo **likelihood**  $l_{i,j} = e^{-\frac{1}{2}\chi_{i,j}^2}$ , definindo uma função distribuição de probabilidades  $P(p)$ .

Como em Gallazzi et al. (2005), “A **PDF** para um determinado parâmetro é simplesmente dada pela distribuição naquele parâmetro das likelihoods de todos os modelos da biblioteca”. Em nosso caso, estamos chamando os parâmetros de propriedades e bibliotecas de Bases.

Na figura 3.6, retirada de Gallazzi et al. (2005), temos um exemplo de aplicação da estatística bayesiana para a recuperação da idade média de quatro galáxias do SDSS com duas bibliotecas diferentes. Neste caso, foram comparados alguns índices espectrais como comparação aos modelos e não o espectro.

A diferença  é que em Gallazzi et al. (2005) foram utilizados modelos de populações estelares como **Base** e aqui estamos utilizando galáxias reais (do SDSS), o que dá ao nosso método maior robustez uma vez que estamos buscando por galáxias através da comparação com outras galáxias e não modelos. Modelos, além de não possuir importantes informações como as medidas das linhas de emissão, muitas vezes podem conter erros nas suas formulações fazendo com que a biblioteca de modelos talvez não seja uma boa representação do universo de possibilidades.

De uma maneira formal, temos que a probabilidade para uma determinada pro-



**Figura 3.6:** Figura retirada de Gallazzi et al. (2005). Função de distribuição de probabilidade para idade média ponderada pela luz na banda para quatro galáxias com  $S/N > 30$  e diferentes idades. As linhas sólidas do painel inferior foram obtidas com uma Base que leva em consideração somente índices sensíveis à idade e, no painel superior, com uma base que leva em consideração índices que dependem também dos metais. As setas indicam o 16º percentil, a mediana e o 84º percentil, respectivamente.

priedade  $p$  para um dado um objeto  $O_i$  dada por:

$$P(p|O_i) = \sum_j P(p|h_j)P(h_j|O_i) \quad (3.6)$$

onde,  $h_j$  (hipótese) é um conjunto de parâmetros observados em  $O_i$  e  $B_j$  e o segundo termo desta equação é dado pelo Teorema de Bayes:

$$P(h_j|O_i) = \frac{P(O_i|h_j)P(h_j)}{P(O_i)} \quad (3.7)$$

Cada um dos termos da equação 3.7 possui um nome formal. O termo à direita na equação,  $P(h_j)$ —é chamado de prior, que mede nosso estado de conhecimento (ou ignorância) sobre a hipótese antes da análise dos dados. O termo à esquerda da equação,  $P(O_i|h_j)$  é chamado de **likelihood** (verossimilhança), e mede a probabilidade dos dados dado o modelo ( $B$ , no nosso caso).

Assumindo erros gaussianos e independentes em cada filtro, a **likelihood** assume a seguinte forma:

$$P(O_i|h_j) = \mathcal{L}_{i,j} = e^{-\frac{1}{2}f_L\chi_{i,j}^2} \quad (3.8)$$

Por fim, a evidência  $P(O_i)$  é uma constante que desaparecerá na normalização.

Podemos então reescrever a equação 3.6 como:

$$P(p|O_i) \propto \sum_j P(p|h_j)e^{-\frac{1}{2}f_L\chi_{i,j}^2}P(h_j) \quad (3.9)$$

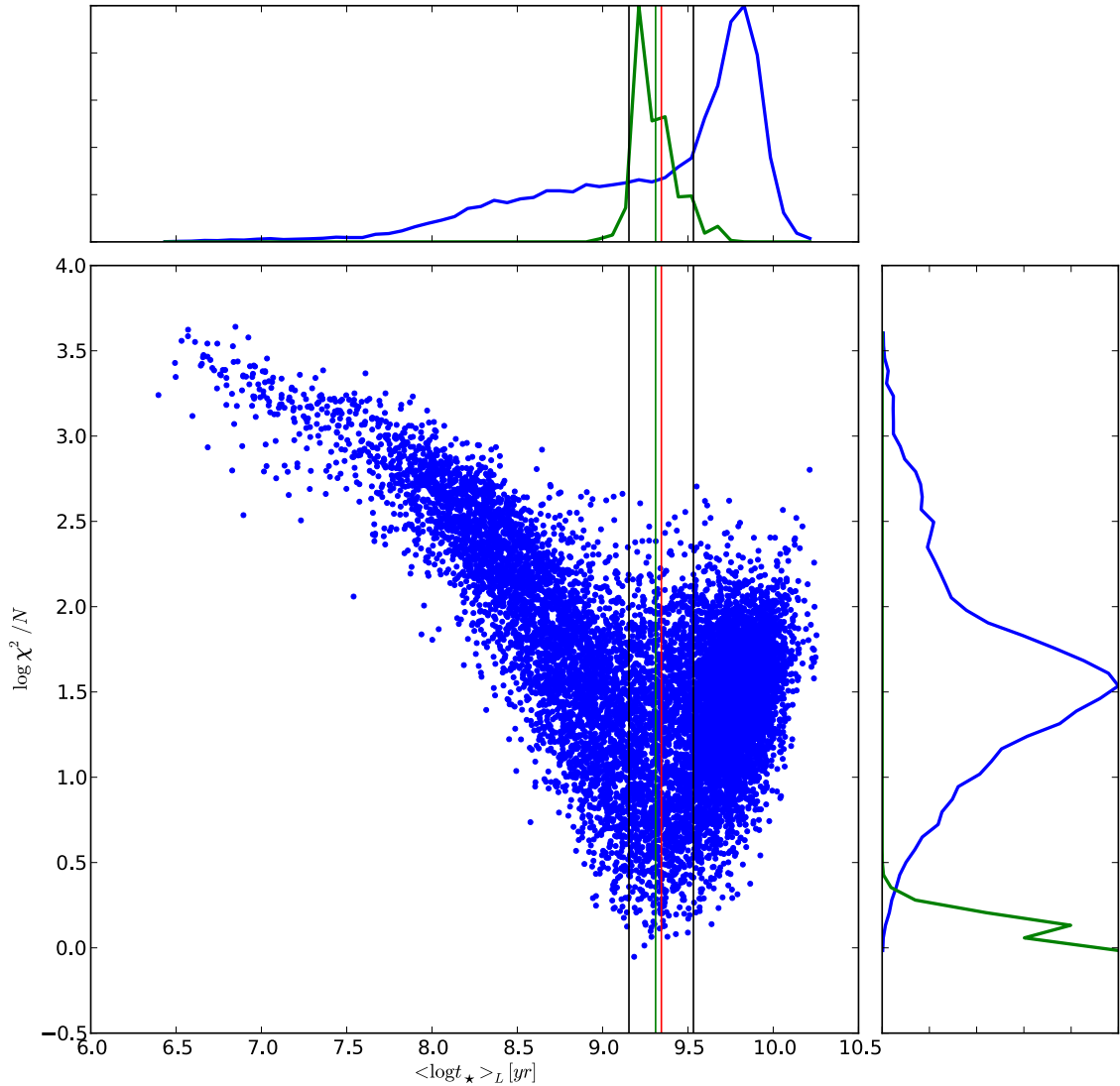
Devido ao fato de estarmos utilizando dados observados, passíveis de erros, tanto no espectro observado como nos espectros da Base, precisamos de alguma forma excluir da conta os píxeis ~~que não estão bons~~ <sup>ruins</sup>. Para isto, substituímos na equação anterior o  $\chi^2$  pelo  $\chi_{RED}^2$  (equação 3.5), ficando a função distribuição de probabilidades aqui utilizada na forma:



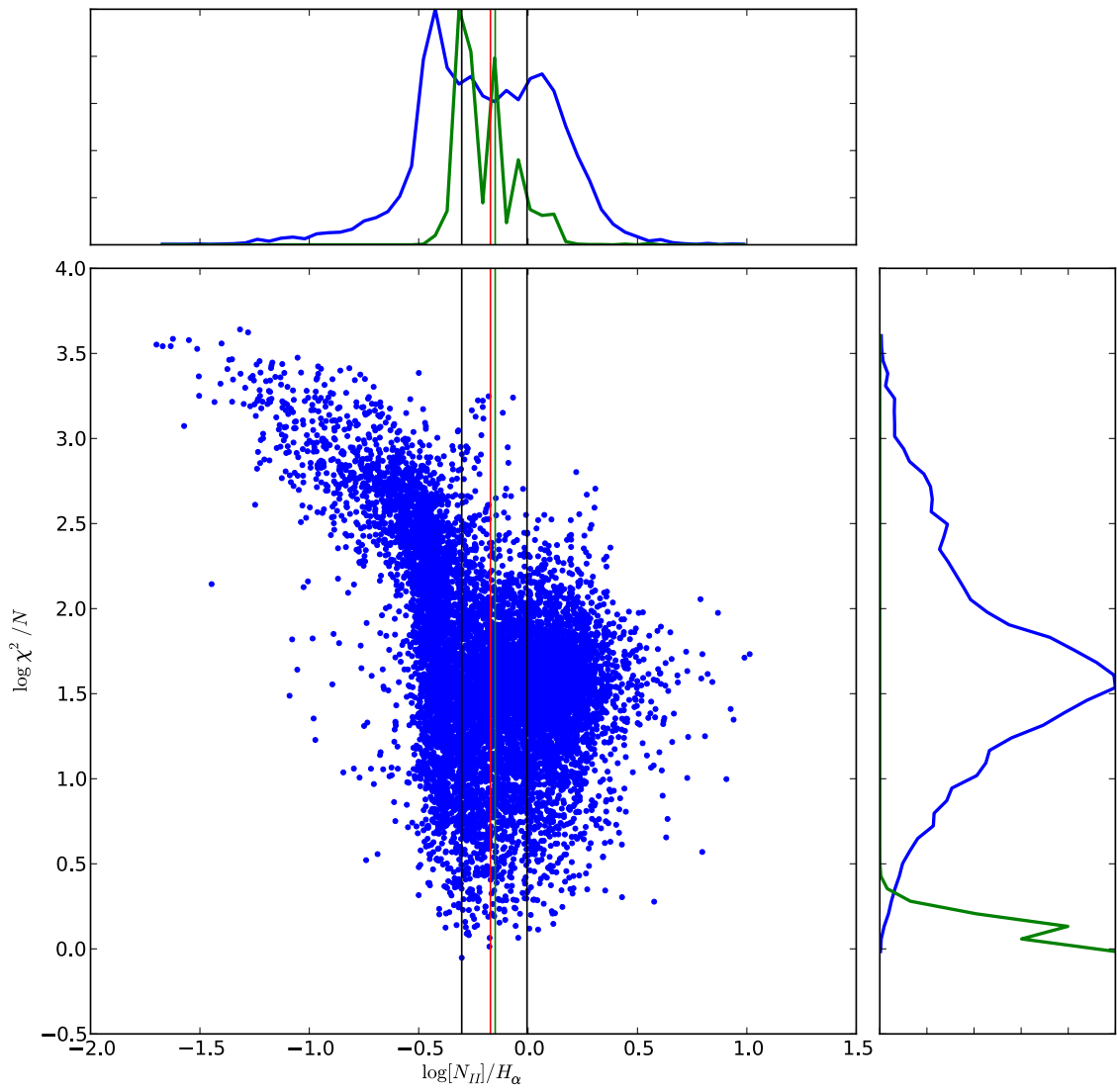
$$P(p|O_i) \propto \sum_j P(p|h_j) e^{-\frac{1}{2}f_L \chi_{i,j,\text{RED}}^2} P(h_j) \quad (3.10)$$

Em geral, o fator  $f_L$  deve ser igual ao número de píxeis bons, mas iremos aqui utilizá-lo como mais um ingrediente que servirá para alargar ou estreitar a função distribuição de probabilidade a fim de aumentar ou diminuir a quantidade de objetos da Base que contribuirão significativamente para a definição de determinada propriedade. Este fator será explorado posteriormente na seção 4.3.

A seguir, nas figuras 3.7 e 3.8, mostramos para uma galáxia a determinação da idade média  $\langle \log t_\star \rangle_L$  e  $[\log(\text{N II}/\text{H}\alpha)]$  a partir de uma Base qualquer usando o formalismo apresentado, onde a função distribuição de probabilidade descrita na equação 3.9 está representada pela linha sólida verde no painel superior.



**Figura 3.7:** Função distribuição de probabilidade para  $\langle \log t_{\star} \rangle_L$  do objeto da figura 3.4. Cada um dos pontos representa um elemento da Base. No eixo  $\langle \log t_{\star} \rangle_L$ , temos a idade média ponderada pelo fluxo extraída da síntese do STARLIGHT e no eixo  $\log \chi^2 / N$  temos a idade média ponderada pelo fluxo extraída da síntese do STARLIGHT e o  $\chi^2_{RED}$ . Nos histogramas lateral e superior temos em azul a distribuição para a Base e em verde a mesma distribuição ponderada pela verossimilhança. As linhas verticais em preto representam os percentis de 16% e 84%, a verde é o valor da idade média que a síntese do STARLIGHT retornou para este objeto e a vermelha é a média das idades médias ponderada pela verossimilhança.



**Figura 3.8:** Mesma figura anterior, porém para a razão de linhas de emissão  $[\log(N_{II}/H\alpha)]$ .

## Capítulo 4

# Amostras, bases e detalhes técnicos

**RESUMO:** Neste capítulo descrevemos a amostra geral de dados do SDSS utilizada tanto para gerar os Objetos como para a construir Bases de referência na seção 4.1. Além disso, na seção 4.2, apresentamos duas bases específicas que usamos para testar o método de casamento espectral e avaliar sua precisão. Por fim, discutimos o efeito do fator  $f_L$  e, na seção 4.3, apresentamos experimentos numéricos de calibração de valores  $f_L$  otimizados para distintas finalidades. Esses desenvolvimentos completam a apresentação do método proposto neste trabalho, e abrem caminho para a apresentação dos resultados de nossas simulações, apresentados no capítulo seguinte.

### 4.1 Amostra Geral

Apesar de já termos apresentado vários exemplos reais no capítulos anteriores, até agora não havíamos delimitado a amostra de dados trabalhada nesta dissertação. Esta é a função desta seção.

Nossas amostras são totalmente oriundas do SDSS. Inicialmente construímos uma amostra geral aplicando um conjunto de critérios semelhantes aos utilizados nos últimos artigos do grupo SEAGal (Cid Fernandes et al. (2010) e Cid Fernandes et al. (2011)). Especificamente, partimos da amostra de 926426 galáxias do DR7 do SDSS analisadas com o STARLIGHT e que estão disponíveis para consulta em nosso banco de dados no sítio [www.stralight.ufsc.br](http://www.stralight.ufsc.br), aplicando a seguinte seleção:

- As galáxias têm de ser do “Main Galaxy Sample” do SDSS;

- Uma caixa de  $31 \text{ \AA}$  e centrada nas linhas de emissão  $H\alpha$ ,  $H\beta$ ,  $[\text{N II}]\lambda 6584$ ,  $[\text{O II}]\lambda 3727$  e  $[\text{O III}]\lambda 5007$  não deve possuir nenhum píxel ruim;
- O redshift das galáxias deve estar entre 0,01 e 0,11

Estes critérios de seleção nos forneceram uma amostra geral de 299253 galáxias. O último critério deve-se ao fato de querermos que os dados a faixa espectral do SDSS caiba completamente na região dos 52 filtros deste trabalho (figura 2.4). **O segundo é imposto para que, se não há medida de linha de emissão, é porque esta é inexistente ou muito fraca e não devido a algum problema na observação do SDSS.** Esta amostra reúne todas os tipos espectrais de galáxias do universo local (com formações estelares de todos os tipos, altas e baixas metalicidades, diferentes extinções, núcleos ativos de linhas de emissão estreitas, etc), com a única exceção são AGNs de linhas largas (Seyferts).

Esta amostra geral foi então dividida entre galáxias cujos espectros possuem razão **Sinal/Ruído  $S/N$**  (medida em um intervalo entre  $4730 \text{ \AA}$  e  $4780 \text{ \AA}$ ) maior ou menor que 20. As 113821 galáxias com  $S/N \geq 20$  formam o grupo das quais serão utilizadas como Base de referência principal e as outras 185432 com  $S/N < 20$  formam o conjunto das quais serão utilizadas como Objetos, cujas propriedades deverão ser estimadas a partir do casamento espectral em relação a uma sub-amostra das galáxias de referência.

Como todas as informações das galáxias vêm do nosso bancos de dados, temos medidas diretas e padronizadas de todas as propriedades que gostaríamos inferir de uma galáxia usando os dados do J-PAS e isto que usaremos para aferir a precisão do método que propomos. Em outras palavras, compararemos os parâmetros de entrada conhecidos (por exemplo, a largura equivalente de  $H\alpha$  em um objeto) com as estimativas que obtemos a partir de nosso método.

## 4.2 Bases

A seleção de uma Base de Referência é um dos principais ingredientes do método de casamento espectral. Assim como em métodos que comparam medidas observadas com bibliotecas de modelos (e.g. Gallazzi et al. (2005)), a composição de uma Base que representará um universo de parâmetros assume o papel implícito de um “prior” (equação 3.7).

Em geral, análises bayesianas devem ser relativamente imunes à escolha do prior,

no sentido de que o resultado não deve depender criticamente do prior. Neste caso, no entanto, é evidente que não devemos construir uma Base composta somente de galáxias sem linhas de emissão ou, ao contrário, somente com galáxias com linhas de emissão fortes. Devemos ter em mente ao definir critérios de seleção para uma base que ela seja o mais genérica possível para que cubra todo o espectro de tipos de galáxias possível.

Tendo isto em mente, surge o problema de como implementar este “princípio filosófico” de forma objetiva. Nesta seção apresentamos duas propostas de Bases que satisfazem este princípio geral de conter todos os tipos espectrais relevantes, mas são construídas de formas completamente independentes. Em ambos casos, as bases foram construídas a partir de um diagrama 2D, pixelizado em pequenas caixas nos eixos X e Y e populando cada pixel (X,Y) de acordo com uma terceira dimensão Z (ver figura 4.1). Uma das bases (seção 4.2.1) usa no plano (X,Y) um análogo “físico” do diagrama cor-magnitude para galáxias e a outra (seção 4.2.2) utiliza o diagrama de diagnóstico WHAN introduzido por Cid Fernandes et al. (2011). No caso da Base física, o terceiro eixo (Z) foi, quando possível, a extinção  $A_V$  e, no caso da Base WHAN, o eixo Z escolhido, quando possível, foi a razão das linhas de emissão  $\log[\text{O III}]/\text{H}\beta$ . Para escolha dos elementos, ordenamos os elementos da caixa (X,Y) em ordem crescente de Z e, então, escolhemos 10% + 1 dos elementos, ou seja, se determinada caixa (X,Y) tem 50 elementos, os elementos ordenados de menor para maior Z a serem escolhidos serão os elementos 1, 5, 10, 15, ..., 45 e 50.

### 4.2.1 Base de Propriedades Físicas: Diagrama “cor-magnitude”

Em astrofísica, o diagrama cor-magnitude (Hertzsprung–Russell) é a base para estudos da evolução estelar. As estrelas passam sua vida ao longo deste diagrama, passando a maior parte de sua vida em uma linha bem definida chamada sequência principal, movendo-se à região das estrelas gigantes e, logo, explodindo em forma de super novas ou buracos negros. Em galáxias (Bell et al. (2004), Baldry et al. (2004)), temos uma distribuição análoga à estelar, com uma distribuição bimodal de galáxias divididas em dois grupos principais: jovens (blue cloud) e velhas (red sequence), com uma parte de transição intermediária (green valley). Na figura 4.2, podemos ver um exemplo de um diagrama destes para 66.856 galáxias do SDSS.

Na figura 4.3, mostramos nossa versão do diagrama “CMD-físico” onde, de forma análoga, substituímos a magnitude pela idade média e a cor pela massa por serem propriedades físicas que é o foco principal deste trabalho. Pode-se observar uma ev-

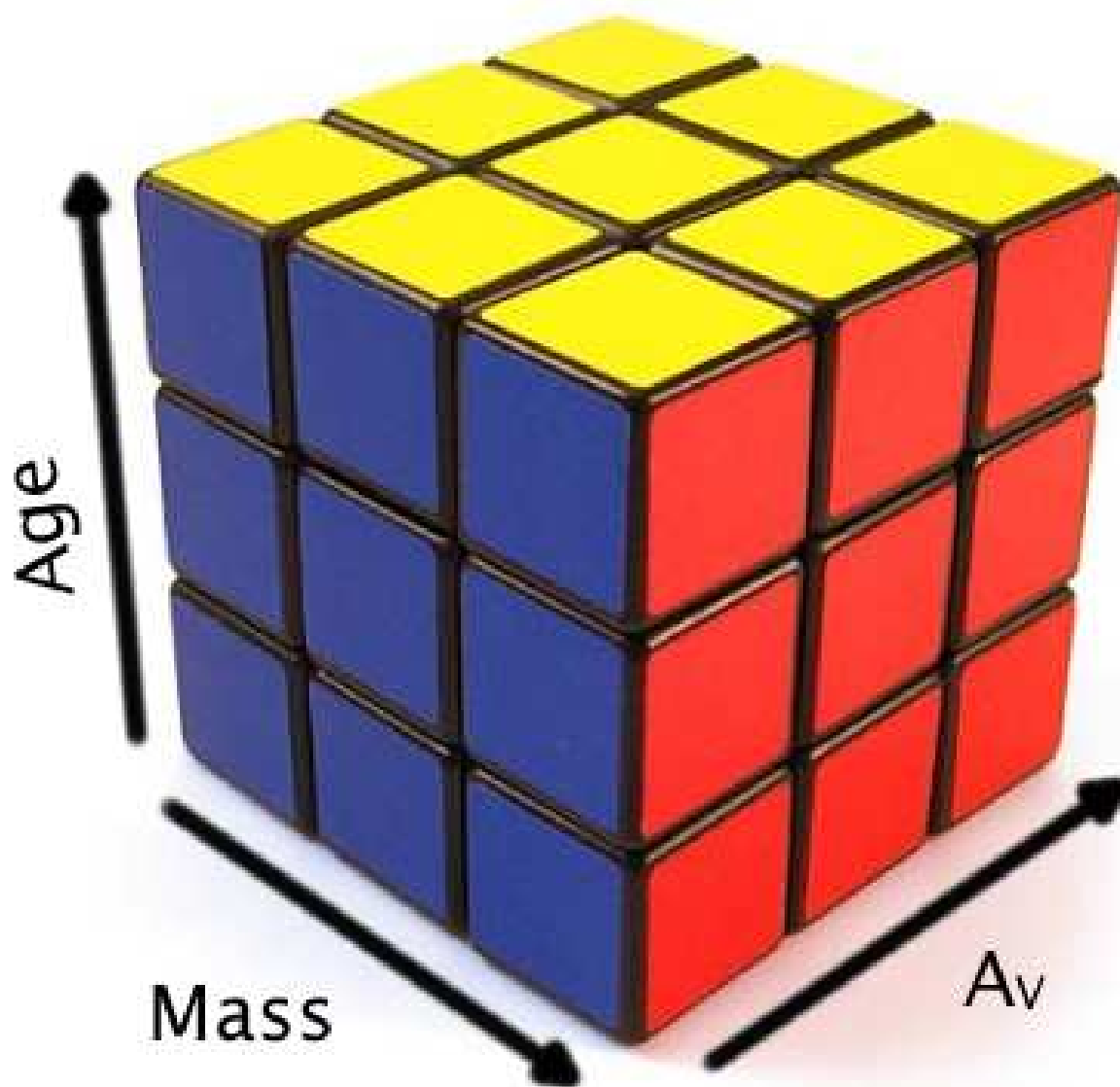
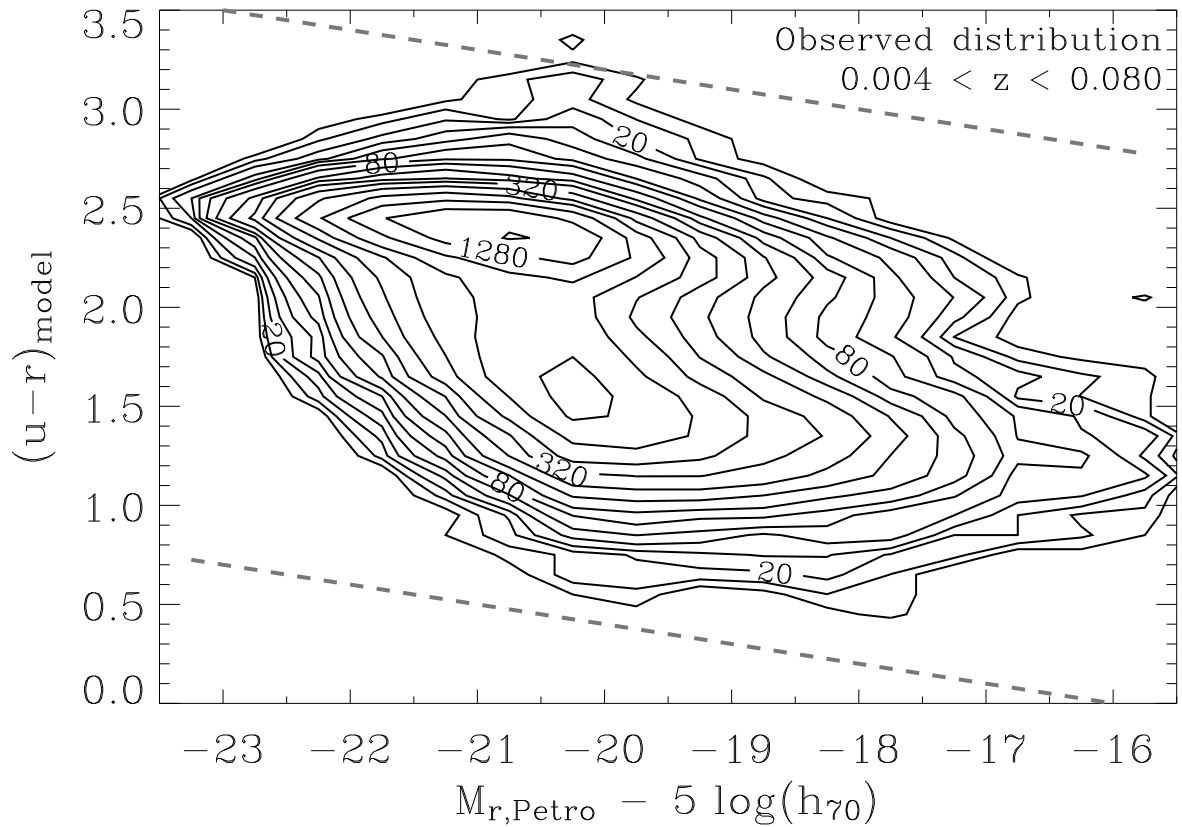


Figura 4.1: Rubik.



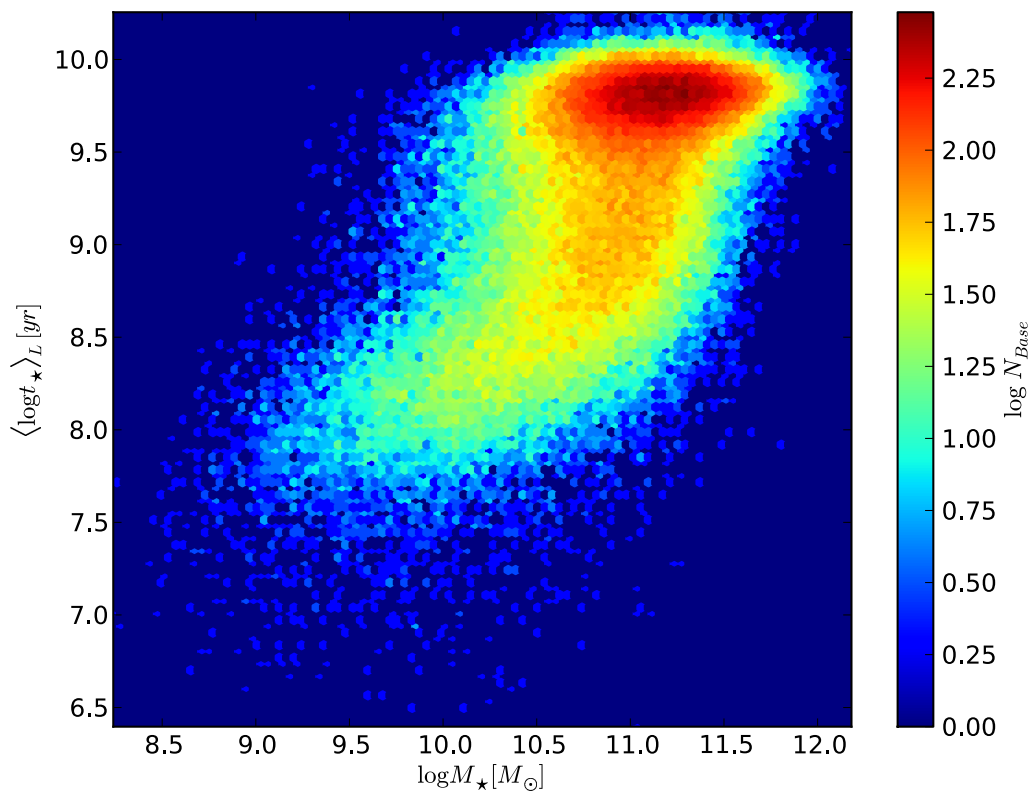
**Figura 4.2:** Diagrama cor/magnitude extraído de Baldry et al. (2004). Os contornos são determinados pela quantidade de galáxias contidas em caixas de 0.1 dex em cor por 0.5 em magnitude de forma logarítmica, iniciando em 10 e dobrando a cada dois contornos, com um total de 66.856 galáxias.

idente correlação entre os dois diagramas. Como massa e idade de galáxias são duas propriedades que se correlacionam (Baldry et al. 2004) muito com as outras que desejamos medir, este critério de seleção de base se justifica.

### 4.2.2 Base sobre o WHAN

O diagrama cor-magnitude físico apresentado na seção anterior nos remete a uma distribuição homogênea em propriedades estelares do espectro de uma galáxia. Porém, uma outra importante componente que faz parte do espectro de uma galáxia são as linhas de emissão que estão relacionadas ao gás no meio interestelar. Apesar de existirem correlações entre propriedades estelares e as linhas de emissão, é interessante definir uma Base que levará a um prior baseado nas linhas de emissão, como já discutimos na seção





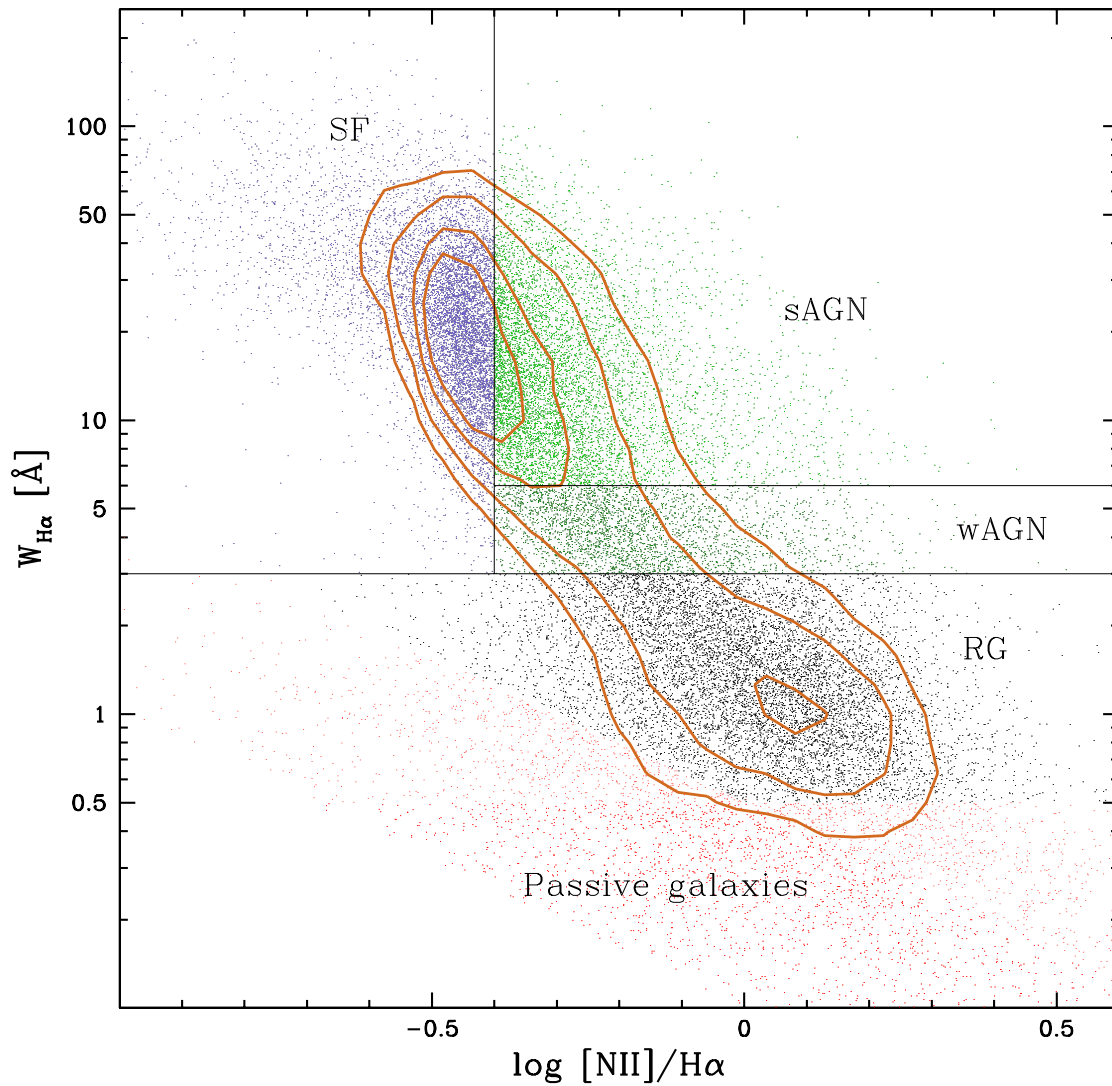
**Figura 4.3:** Diagrama cor/magnitude para a amostra de galáxias selecionadas como Base. Esta figura foi fatiada em caixas de .1 dex e, para cada caixa, foram retirados 10% dos objetos **com**, variando ao longo de  $A_V$  quando possível. Nesta base, encontramos 11952 elementos.

## 4.2.

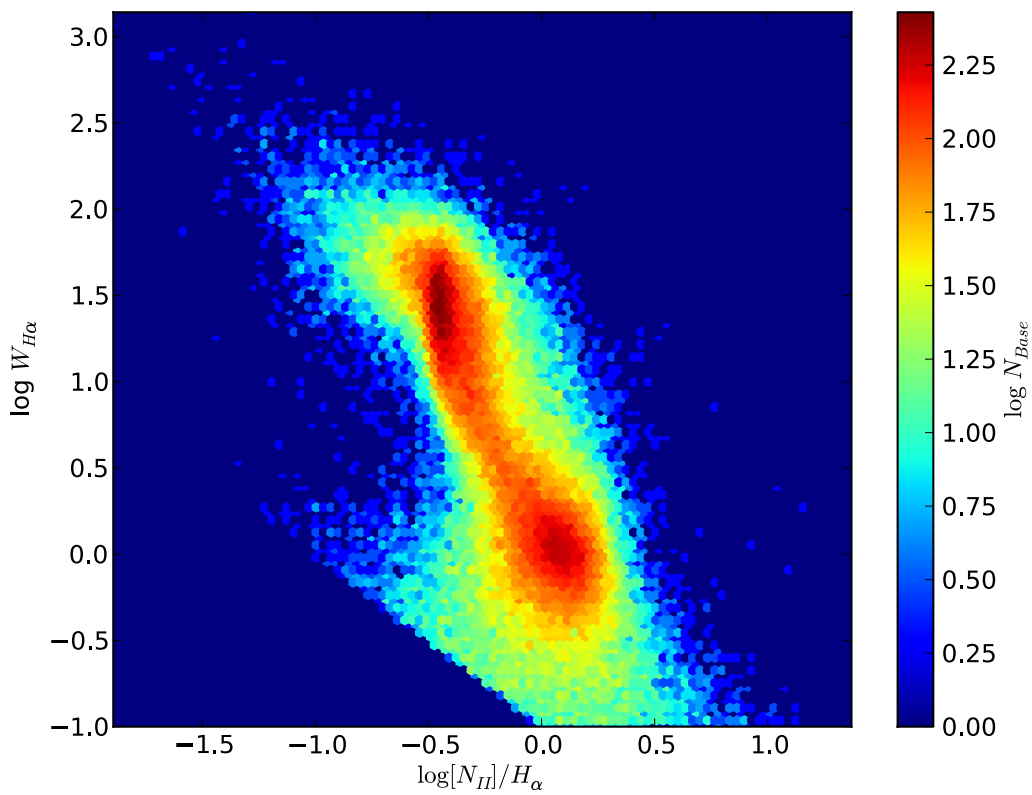
Neste trabalho, visando uma aplicação direta da medida das linhas de emissão, temos o interesse em classificar galáxias a partir de diagramas de diagnóstico. Assim que, estabelecemos nossa segunda base sobre o diagrama WHAN (Cid Fernandes et al. (2011)). A escolha do WHAN como diagrama diagnóstico base se deve ao fato de que este diagrama estende o tradicional diagrama BPT (Baldwin et al. (1981)) para uma amostra mais completa. O diagrama WHAN, em uma amostra qualquer, abrange um número maior de galáxias do que o BPT e principalmente devido ao fato deste não depender da medida direta da linha de emissão [O III], geralmente mais fraca e muitas vezes inexistente, efeito que já ocorre em espectros de média-alta resolução como os do SDSS.

Além disto, o diagrama WHAN (figura 4.4), reclassifica as galáxias em cinco categorias: galáxias com formação estelar (SF), com núcleo ativo forte (sAGN), com núcleo ativo fraco (wAGN), aposentadas (RG) e passivas (PG). Como sumarizado em Cid Fernandes et al. (2011), esta nova classificação possui três virtudes principais:

- Identifica os três principais agentes ionizantes em galáxias: estrelas jovens, núcleos ativos e HOLMES (do inglês: **Hot low-mass evolved stars**)
- É baseada no mais barato e, conseqüentemente, mais inclusivo diagrama de diagnóstico, o único capaz de classificar a grande população de galáxias com linhas fracas (Cid Fernandes et al. (2010)) até agora negligenciadas. De fato, ela até nos permite visualizar a região ocupada pelas galáxias passivas, característica que não é possível através de outro método tradicional (vale ressaltar que galáxias com nenhuma linha de emissão ainda não aparecem neste diagrama).
- As fronteiras que delimitam galáxias SF de AGN e sAGN de wAGN são baseadas em transições otimizadas de linhas baseadas em diagramas de diagnóstico tradicionais, onde as classificações funcionam bem, fazendo que em classificações com o diagrama WHAN estas sejam preservadas e, ao mesmo tempo, resolve o problema das galáxias “esquecidas” por causa de suas linhas de emissão fracas e da incapacidade dos outros métodos de distinguir núcleos ativos verdadeiros de falsos.



**Figura 4.4:** Diagrama WHAN e categorização para galáxias revisada dividida em: galáxias com formação estelar (SF), com núcleo ativo forte (sAGN), com núcleo ativo fraco (wAGN), aposentadas (RG) e passivas (PG). Retirado de Cid Fernandes et al. (2011).



**Figura 4.5:** Diagrama WHAN para a amostra de galáxias selecionadas como Base. Esta figura foi fatiada em caixas de .01 dex e, para cada caixa, foram retirados 10% dos objetos, variando ao longo de  $\log[\text{O III}]/\text{H}\beta$  quando possível. Nesta base encontramos 27536 elementos.

### 4.3 Fator-cozinha $f_L$

Antes de partir para as aplicações no próximo capítulo, temos que considerar um último ingrediente do método de casamento: o fator  $f_L$ . Apesar de ser menos relevante que a escolha da Base no ponto de vista filosófico, este fator tem um forte impacto prático na estimativa de propriedades através do método aqui proposto.

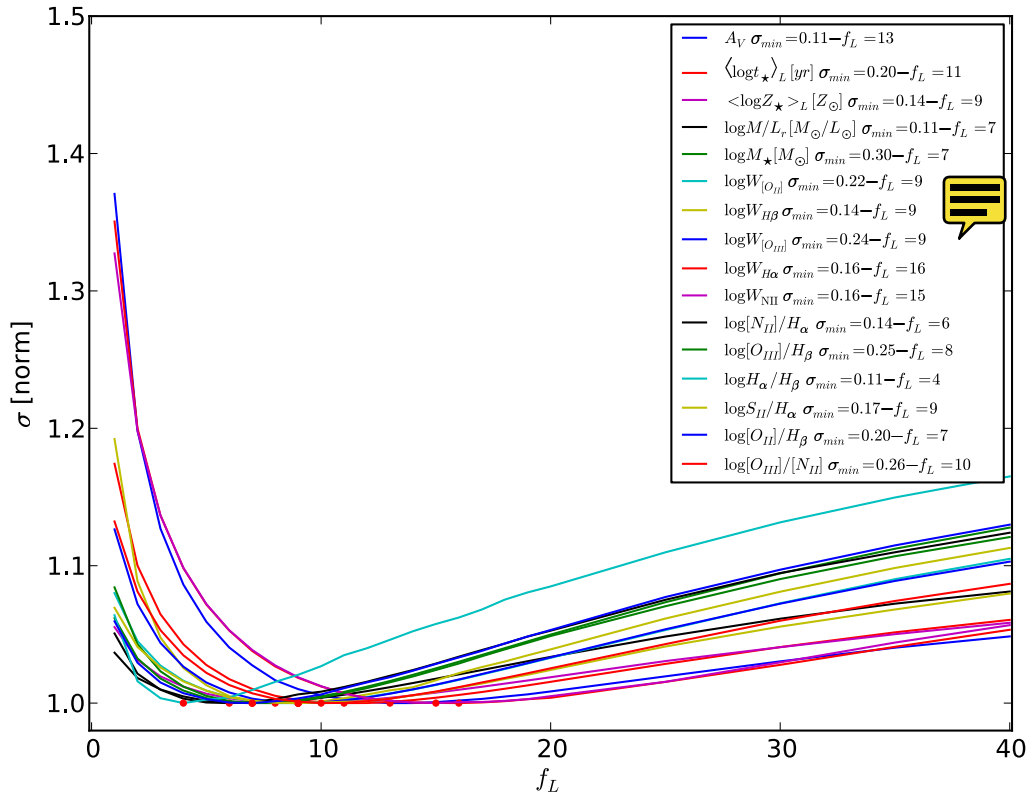
Introduzido na equação 3.9, este fator “Ad-hoc” deveria, em princípio ser igual  $! \odot j \odot !$  ao número de pontos bons nos espectros observado e da base  $! \odot j \odot !$ , mas na prática podemos experimentar a variação deste número para outros valores que levem à uma maior eficiência do método.

O papel prático do  $f_L$  pode ser entendido da seguinte maneira: quando  $f_L$  tende a zero, todas as galáxias tendem a contribuir igualmente nas funções distribuição de probabilidade da equação 3.10. Isto representa uma função likelihood que diz muito pouco sobre a comparação e o posterior será quase igual ao prior. No outro extremo, com  $f_L$  tendendo ao infinito, apenas a galáxia com o menor  $\chi^2$  contribuirá significativamente para o posterior. Neste limite, o resultado do método será igual ao do pareamento discutido na seção 3.4. Valores intermediários de  $f_L$  definirão se mais ou menos galáxias contribuirão efetivamente na conta dos posteriores.

Para estabelecer o valor ideal para  $f_L$ , adotaremos uma abordagem heurística, determinando o  $f_L$  como simplesmente o que nos remete aos melhores resultados. Por “melhor resultado” queremos dizer aquele que minimize (em um sentido estatístico) a diferença entre os valores de entrada e os de saída (i.e., os estimados bayesianamente) para uma dada propriedade dentre as que desejamos recuperar. Isto implica na possibilidade de termos um  $f_L$  ótimo para cada propriedade, ou seja,  $f_L = f_L(p)$ , onde  $p$  é uma propriedade qualquer.

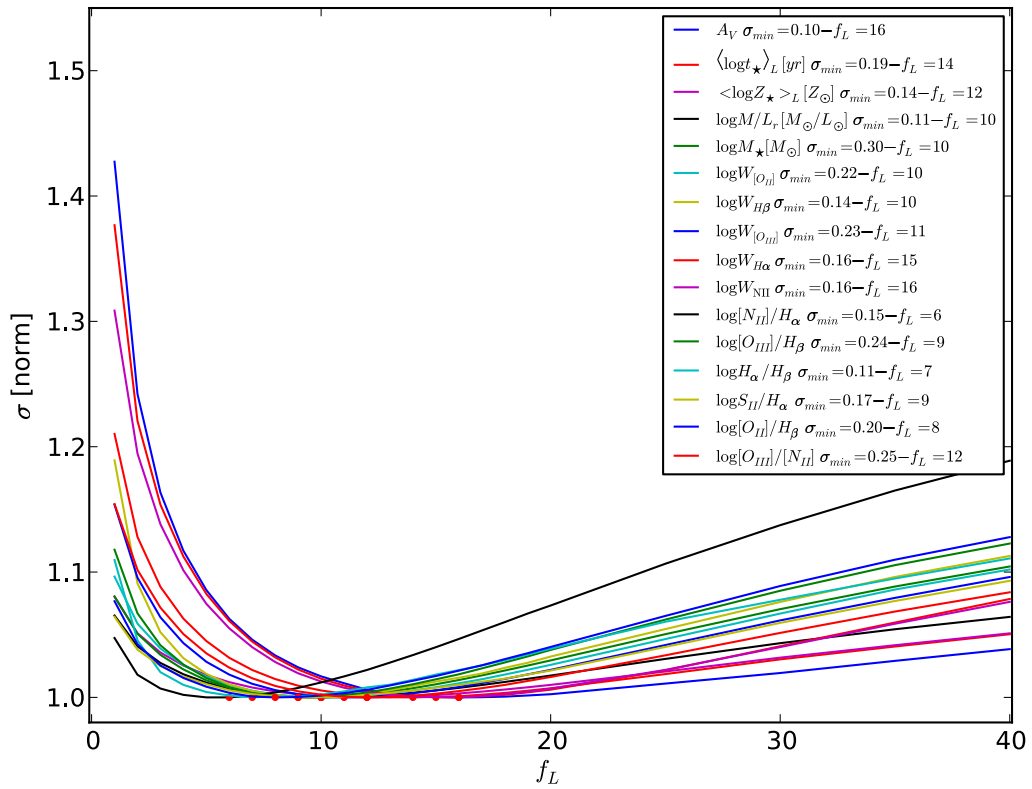
Para determinar o  $f_L$  ideal, foram calculados as dispersões das diferenças entre os parâmetros de entrada (provenientes do STARLIGHT) e a média de determinado parâmetro ponderada pela função distribuição de probabilidades (descrita mais formalmente na próxima seção, equação 4.2). Assim, o fator  $f_L$  a ser escolhido para cada parâmetro  $P$  será o que minimiza a relação abaixo:

$$\sigma(P_{f_L}) = \sigma(P_{f_L, \text{saida}} - P_{f_L, \text{entrada}}) \quad (4.1)$$



**Figura 4.6:** Dispersão da diferença entre Entrada versus Saída para diferentes fatores-cozinha  $f_L$  e diferentes propriedades para a base “diagrama cor-magnitude”. Para este gráfico, foram calculadas estatísticas para  $f_L = 1$  a  $f_L = 20$  com  $\Delta f_L = 1$  e para  $f_L = 25, 30, 35$  e  $40$ . Na legenda, estão representados os valores mínimos para a dispersão de cada valor e o  $f_L$  que representa estes mínimos.

Isto foi realizado para ambas Bases, conforme nas figuras a seguir.



Mesmo gráfico ...

Figura 4.7: Gráfico anterior, para a base “WHAN”.

## 4.4 Estimadores (Resumos da Função Distribuição de Probabilidades)

Mesmo com toda a elegância que análises probabilísticas proporcionam, no sentido que a função distribuição de probabilidade é formal e filosoficamente a maneira mais correta de expressar os resultados, para **realizamos** comparações concretas entrada versus saída necessitamos um número que quantifique determinada propriedade. Na prática, com uma função distribuição de probabilidades não podemos responder perguntas do tipo: “Qual a largura equivalente de  $H\alpha$  do Objeto número tal do J-PAS?”. Assim ~~que~~ se torna necessário resumir esta função em um número que expresse a quantidade que desejamos medir mais um outro que represente o erro (ou a confiança) nesta medida.

Há várias maneiras de se fazer isto. Neste trabalho, utilizaremos a média e o desvio padrão como representações destes valores. Assim ~~que~~, nas comparações entrada versus saída no capítulo seguinte, os valores que adotaremos como saída de nosso método serão, na realidade, a média da propriedade dada a função distribuição de probabilidades dada pela equação 4.2. E, quando falamos nos erros de saída, estamos nos referindo ao desvio padrão desta função dado pela equação 4.3:

$$\bar{P}_i = \frac{\sum_j P \mathcal{L}_{i,j}}{\sum_j \mathcal{L}_{i,j}} \quad (4.2)$$

$$\sigma(P_i) = \frac{\sqrt{\sum_j (P_i - \bar{P}_i)^2 * \mathcal{L}_{i,j}}}{\sum_j \mathcal{L}_{i,j}} \quad (4.3)$$

lembrando que  $\mathcal{L}_{i,j} = e^{-\frac{1}{2}fL\chi_{i,j}^2}$ , como definido na equação 3.8.

!⊙j⊙! *lll* Colocar um gráfico de exemplo???



# Capítulo 5

## Resultados

**RESUMO:** *Nos capítulos anteriores apresentamos os princípios gerais do nosso método, seus ingredientes, o formalismo e os detalhes técnicos. Neste capítulo, aplicamos a metodologia desenvolvida com o objetivo de testar e quantificar quão precisamente podemos recuperar uma série de propriedades de galáxias.*


### 5.1 Propriedades de interesse

O foco deste trabalho é testar se o casamento J-espectral é capaz de nos fornecer informações úteis sobre os nossos Objetos alvo. Com a metodologia apresentada nos capítulos anteriores, neste capítulo verificaremos se estas propriedades podem ser recuperadas e com qual precisão. A precisão será aferida comparando-se a saída de nosso método com os valores conhecidos, provenientes da saída do STARLIGHT, filosofia típica utilizada em qualquer simulação: comparar a entrada com a saída.

Estas comparações se centrarão em um conjunto de propriedades físicas e observacionais que são de interesse e que acreditamos poder recuperar. Dentre as muitas propriedades que dispomos em nosso banco de dados, investigaremos as listadas na tabela 5.1.

**Tabela 5.1:** Propriedades analisadas, suas unidades e descrições.

Propriedade	Descrição	Unidade
$A_V$	Extinção estelar	mag
$\langle \log t_\star \rangle_L$	Idade estelar média ponderada pela luminosidade	anos
$\langle \log Z_\star \rangle_L$	Metalicidade estelar média ponderada pela luminosidade	$Z_\odot$
$\log M/L_r$	Razão massa-luminosidade na banda r do SDSS	$M_\odot/L_\odot$
$\log M_\star$	Massa total	$M_\odot$
$\log W_{[OII]}$	Largura equivalente de $[OII]$	$\text{\AA}$
$\log W_{H\beta}$	Largura equivalente de $H_\beta$	$\text{\AA}$
$\log W_{[OIII]}$	Largura equivalente de $[OIII]$	$\text{\AA}$
$\log W_{H\alpha}$	Largura equivalente de $H_\alpha$	$\text{\AA}$
$\log W_{[NII]}$	Largura equivalente de $[NII]$	$\text{\AA}$
$\log[NII]/H_\alpha$	Razão de fluxo de linhas de emissão $[NII]$ por $H_\alpha$	-
$\log[OIII]/H_\beta$	Razão de fluxo de linhas de emissão $[OIII]$ por $H_\beta$	-
$\log H_\alpha/H_\beta$	Razão de fluxo de linhas de emissão $H_\alpha$ por $H_\beta$	-
$\log S_{II}/H_\alpha$	Razão de fluxo de linhas de emissão $S_{II}$ por $H_\alpha$	-
$\log[OII]/H_\beta$	Razão de fluxo de linhas de emissão $[OII]$ por $H_\beta$	-
$\log[OIII]/[NII]$	Razão de fluxo de linhas de emissão $[OIII]$ por $[NII]$	-

Esta lista contempla um conjunto de propriedades que diagnosticam desde as propriedades estelares dos objetos (como idade média das populações, metalicidade e poeira) como as propriedades do gás (através das medidas das linhas de emissão e razões de seus fluxos). Vale a pena, mais uma vez, ressaltar a importância da medida de linhas de emissão pelo nosso método, uma vez que propriedades como extinção, idade média podem ser facilmente medidos utilizando-se outros métodos.  Preferimos aqui converter para log ao invés dos valores normais para as propriedades para que os gráficos ficassem melhor colocados e para que as dispersões nas análises de entrada versus saída fossem sempre em dex.  $A_V$  é a única exceção, pois já é uma unidade com ordens de grandeza logarítmicas.

## 5.2 Resultados: Propriedades Físicas

**Tabela 5.2:** Base CMD

Propriedade	$f_L$	$\langle \Delta p \rangle$	$\tilde{\Delta p}$	$\sigma(\Delta p)$	Mínimo	Máximo
$A_V$	13	0.023	0.022	0.106	-0.804	2.927
$\langle \log t_\star \rangle_L$	11	-0.018	-0.017	0.199	6.000	10.255
$\langle \log Z_\star \rangle_L$	9	-0.021	-0.006	0.144	-2.292	0.398
$\log M/L_r$	7	-0.045	-0.034	0.114	-1.716	0.682
$\log M_\star$	7	-0.184	-0.165	0.301	7.146	12.125

**Tabela 5.3:** Base WHAN

Propriedade	$f_L$	$\langle \Delta p \rangle$	$\tilde{\Delta p}$	$\sigma(\Delta p)$	Mínimo	Máximo
$A_V$	16	0.023	0.020	0.101	-0.804	2.927
$\langle \log t_\star \rangle_L$	14	-0.013	-0.012	0.192	6.000	10.255
$\langle \log Z_\star \rangle_L$	12	-0.020	-0.006	0.141	-2.292	0.398
$\log M/L_r$	10	-0.040	-0.029	0.110	-1.716	0.807
$\log M_\star$	10	-0.182	-0.164	0.296	7.146	12.125

NA FIG ?? MOSTRAMOS HISTOGRAMAS DE XOUT - XIN PARA AS PROPRIEDADES FISICAS NA TABELA S??. NESTA FIG A BASE DE REFERENCIA EH A BASE CMD, E OS OBJETOS ANALISADOS SAO AS 12345? GALAXIAS DESCRITAS EM ??. ??? FAZER COM MEDIA E COM P50 ....COMENTAR FIGURA??

A FIG ?? MOSTRA UMA MANEIRA MAIS ADEQUADA DE COMPARAR AS COISAS. NELA AS ABACISSAS SAO A DIFERENCA ENTRE SAIDA E ENTRADA EM UNIDADES DA INCERTEZA. (ISSO EH MAIS ADEQUADO PQ UMA GRANDE DIFERENCA ENTRE ENTRADA E SAIDA NAO EH REALMENTE UM PROBLEMA QDO A INCERTEZA ASSOCIADA EH POPORCIONALMENTE GRANDE).

!o que tenho em mente aqui eh algo parecido com a fig 2 da Gallazzi

et al 2005 - tem uma fig parecida em algum outr paper onde o bruzual tah metido...)

BLABLABLA ... ATABELA ?? RESUME OS RESULTADOS DESTA ANALISE. AO CONTRARIO DAS FIGURAS ?? E ??, AMBAS BASEADAS NA BASE CMD, ESTA TABELA TB MOSTRA RESULTADOS OBTIDOS COM A BASE WHAN. VEMOS QUE A INCERTEZA TIPICA NISSO E NAQUILOO EH DA ORDEM DE .... VEMOS TB QUE NAO FAZ LAH GRANDE DIFERENCA USAR A BASE CMD OU WHAN (???EH VERDADE??)...

### 5.3 Resultados: Linhas de Emissão

Tabela 5.4: Base CMD

Propriedade	$f_L$	$\langle \Delta p \rangle$	$\tilde{\Delta p}$	$\sigma(\Delta p)$	Mínimo	Máximo
$\log W_{[OII]}$	9	0.051	0.065	0.223	-0.971	2.232
$\log W_{H\beta}$	9	0.024	0.020	0.145	-1.000	2.119
$\log W_{[OIII]}$	9	0.046	0.048	0.245	-1.000	2.924
$\log W_{H\alpha}$	16	0.010	0.008	0.160	-1.000	2.881
$\log W_{[NII]}$	15	-0.028	-0.024	0.159	-1.000	2.282
$\log[NII]/H\alpha$	6	-0.045	-0.042	0.141	-2.099	1.140
$\log[OIII]/H\beta$	8	0.026	0.027	0.250	-1.603	1.350
$\log H\alpha/H\beta$	4	-0.011	-0.013	0.107	-0.990	2.268
$\log SII/H\alpha$	9	-0.006	0.019	0.172	-2.381	1.199
$\log[OII]/H\beta$	7	0.036	0.049	0.202	-1.150	1.833
$\log[OIII]/[NII]$	10	0.075	0.063	0.265	-1.944	2.098

AS FIGS ?? SAO ANALOGAS AS ?? E ??, MAS AGORA PARA AS LINHAS DE EMISSAO... VEMOS ISSO E AQULO. DISCUTIR SE BASE WHAN EH REALEM-NTE MELHOR QUE A BASE CMD PARA ESTIMAR LINHAS DE EMISSAO (ACHO QUE A DIFERENCA VAI SER POUCA).

!ojo! meter tabelas aqui...

Tabela 5.5: Base WHAN

Propriedade	$f_L$	$\langle \Delta p \rangle$	$\tilde{\Delta p}$	$\sigma(\Delta p)$	Mínimo	Máximo
$\log W_{[OII]}$	10	0.040	0.052	0.218	-0.971	2.232
$\log W_{H\beta}$	10	0.030	0.027	0.143	-1.000	2.119
$\log W_{[OIII]}$	11	0.029	0.034	0.232	-1.000	2.924
$\log W_{H\alpha}$	15	0.022	0.015	0.157	-1.000	2.881
$\log W_{[NII]}$	16	-0.010	-0.013	0.156	-1.000	2.282
$\log[NII]/H\alpha$	6	-0.044	-0.041	0.146	-2.099	1.140
$\log[OIII]/H\beta$	9	-0.000	0.004	0.238	-1.603	1.344
$\log H\alpha/H\beta$	7	-0.010	-0.012	0.114	-1.089	2.268
$\log SII/H\alpha$	9	-0.016	0.011	0.174	-2.381	1.227
$\log[OII]/H\beta$	8	0.016	0.028	0.198	-1.158	1.833
$\log[OIII]/[NII]$	12	0.044	0.037	0.252	-1.944	2.098

## 5.4 “OUTLIERS”

APESAR DOS RESULTADOS SATISFATORIOS OBTIDOS, UMA INSPECAO CUIDADOSA MOSTRA QUE TEM UNS FDP CUJAS PROPIEDADES RECUPERADAS SAO GROTESCAMENTE DIFERENTES DAS ORIGINAIS. ?? DAR UM EXEMPLO NUMERICO!

INVESTIGAMOS ESSES CASOS PROBLEMATIOS E CONCLUIMOS QUE ELES SAO PROBLEMATICOS PQ OS PROPRIOS DADOS SAO PROBLEMATICOS!! AS FIGS BLABLABLA MOSTRAM 2 CASOS/EXEMPLOS....

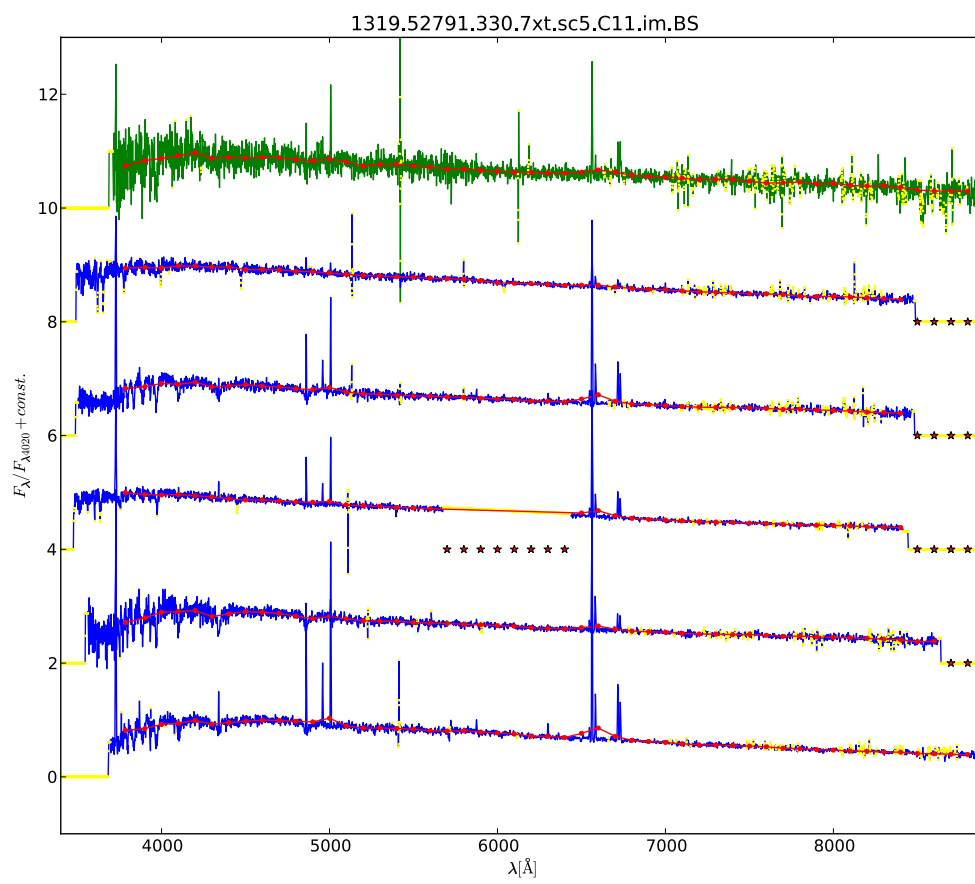


Figura 5.1: alala

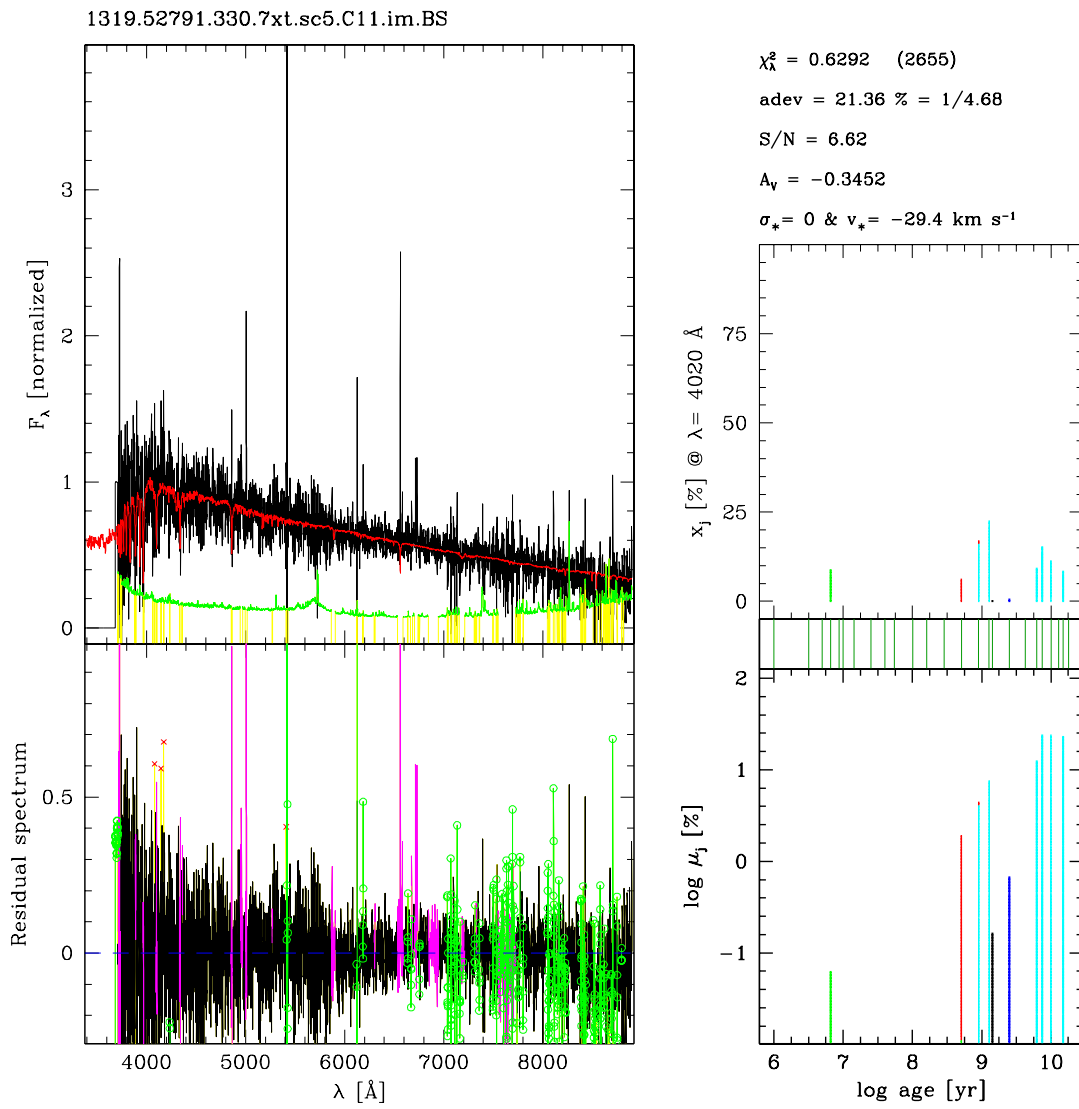


Figura 5.2: Objeto

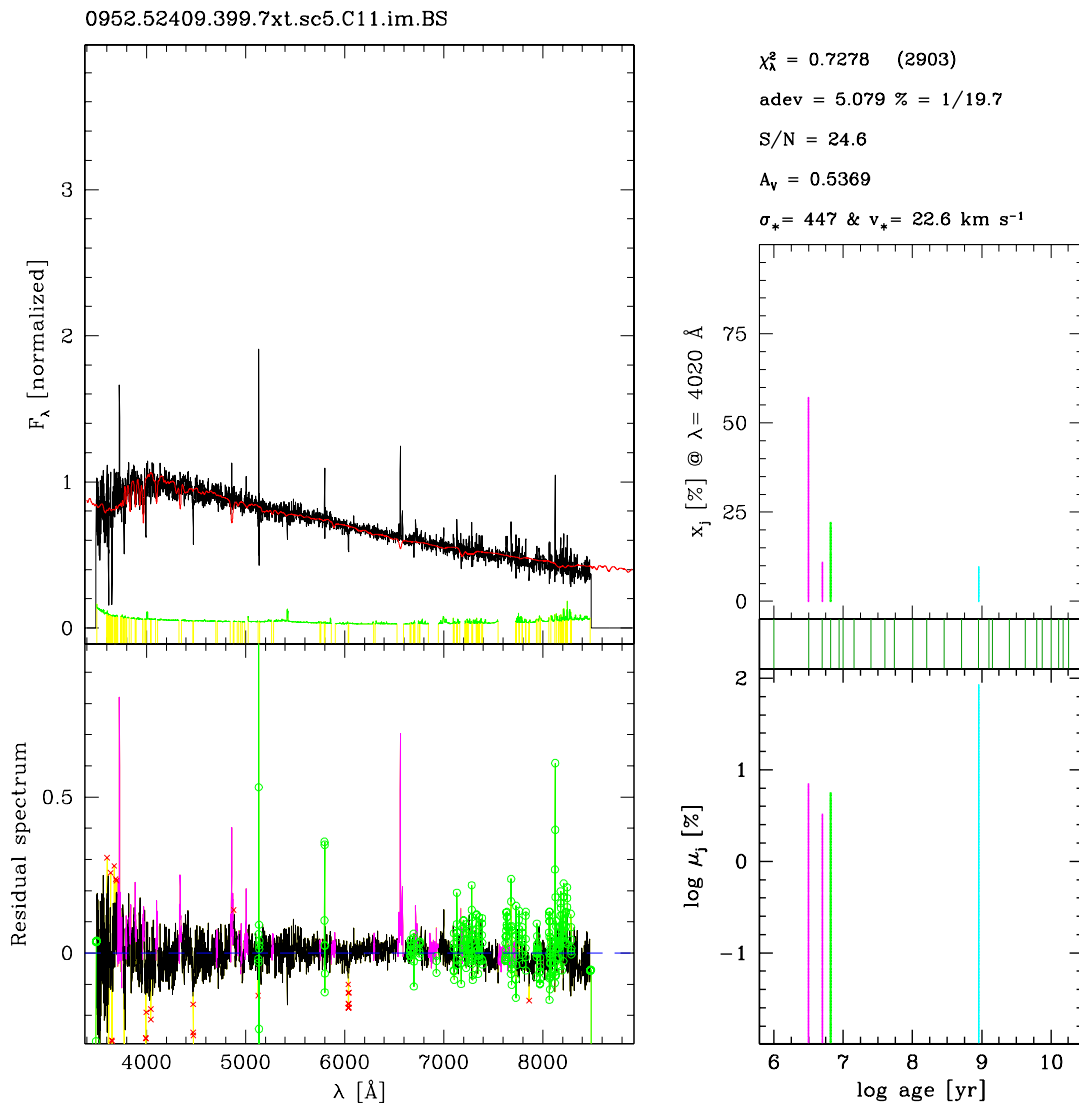


Figure 5.3: Base



## 5.5 Resumo?

resumir precisoes com que estimamos isso e aquilo!

# Capítulo 6

## Conclusões e Perspectivas

**RESUMO:** *Neste capítulo apresentamos um resumo do trabalho desenvolvido, seus resultados principais e enumeramos os próximos passos para o seguimento do trabalho.*

### 6.1 Resumo do trabalho e seus principais resultados

O SDSS foi até hoje o maior levantamento de dados espectroscópicos até a atualidade do qual extraímos uma grande quantidade de informação sobre a composição química das galáxias. No entanto, levantamentos espectroscópicos são caros e, pelo menos para a próxima década, não é realista se esperar a realização de outro projeto de mesmo porte para galáxias mais fracas (distantes) nem para espectroscopia espacialmente resolvida de galáxias. Apesar de haverem planos nestes sentidos, levantamentos fotométricos de bandas estreitas serão em um futuro próximo muito mais abrangentes e abundantes. Boa parte da motivação para a realização destes levantamentos é cosmológica e têm seus esforços direcionados à medida precisa de redshifts de milhares de galáxias para estimar Oscilações Acústicas Bariônicas e propriedades da matéria escura (Benítez et al. (2009)). Neste trabalho, propomos uma forma de aproveitar esta massa gigantesca de dados em prol de outra causa: medir propriedades das populações estelares e linhas de emissão nas galáxias.

Neste trabalho, desenvolvemos e testamos uma técnica para extrair este tipo de informação de dados de fotometria para 56 filtros estreitos iguais aos planejados para o projeto JPAS. O método consiste em comparar os fluxos nestes filtros (chamados aqui de J-Espectros) observados com aqueles medidos em amostras de referência obtidas

com o SDSS e para as quais dispomos de um grande número de propriedades muito bem medidas em espectros com resolução muito maior do que será obtida com o JPAS. A premissa principal é que se os J-Espectros de duas galáxias são parecidos, os seus espectros em maior resolução também serão e, conseqüentemente, suas propriedades.

Assim, no capítulo 2 apresentamos os dois levantamentos (SDSS e JPAS) e, em seguida, apresentamos como foram convertidos os espectros do SDSS para os do JPAS bem como o espectro de erros e no capítulo 3 introduzimos o método de casamento espectral e apresentamos o formalismo estatístico adotado neste seguido de alguns exemplos.

Um ingrediente essencial neste método é a Base de comparação. Formalmente, ela entra como um “prior” implícito na análise. Em princípio, diferentes bases podem ser construídas, dependendo do que se quer inferir sobre os J-Espectros. No capítulo 4 apresentamos duas bases com as quais trabalhamos: uma baseada no diagrama massa-idade (análogo ao amplamente usado diagrama cor-magnitude para galáxias) e outra baseada no diagrama WHAN, que usa as linhas de emissão de  $H\alpha$  e  $[N II]$  para diagnosticar que tipo de agente ionizante domina a produção de emissão nebular em galáxias. Este capítulo também define quais amostras do SDSS foram utilizadas como Objeto e Base em nossas simulações do método a calibração do fator  $f_L$  que regula o peso dado a cada “posterior”.

No capítulo 5 apresentamos os resultados da aplicação do método e com qual precisão com que se recupera diferentes propriedades. Identificamos também alguns casos onde o método falha absurdamente e percebemos que estes correspondem a espectros com problemas, o que nos diz que nosso método é confiável para a medida destas propriedades, mas que devemos tomar cuidado identificando dados ruins e retirando estes das nossas análises.

A principal questão a ser respondida neste trabalho é com que precisão podemos recuperar uma dada propriedade de um Objeto. Os resultados apresentados no capítulo 5 podem ser resumidos desta forma:

- PROPRIEDADES FISICAS: ...
- LINHAS DE EMISSAO: ... AQUI ENFATIZAR QUE ESTE PARECE SER O MELHOR METODO PRA “MEDIR” LINHAS DE EMISSAO. METODOS PRA MEDIR PROPRIEDADES FISICAS EXISTEM E ESTAO SENDO APRIMORADOS (EG GALAZZI, ADAPTACOES DO STARLIGHT, ETC), MAS NAO EX-

ISTE NADA PARECIDO PRA EM LINES (ATEH SE PODEM CRIAR BASES DE MODELOS INCLUINDO EMISSAO NEBULAR, MAS DAH PRA SABER DE ANTEMAO QUE ELAS NAO SERIAM NE, REMOTAMENTE TAO CONFI-AVEIS/UTEIS COMO BASES DE MODELO DE STELLAR POPULATIONS).

## 6.2 Perspectivas

Do ponto de vista metodológico,

- Realizar mais testes com outras bases construídas de outras formas e aferir que tipo de base é mais adequada a cada aplicação;
- Substituir as funções deltas de dirac que representam cada propriedade  $p$  de cada galáxia da base por distribuições de probabilidade associadas à esta propriedade. Ou seja, temos que substituir a função  $\text{PDF}(p) = \delta(p - p_j)$  por uma função distribuição de probabilidades realista, como, por exemplo, a função gaussiana  $\text{PDF}(p) = (2\pi\sigma(p_j)\text{exp}-(p - p_j)^2/2\sigma^2(p_j))$ . Não esperamos que este aprimoramento modifique de modo significativo nossos resultados, porém representa um requinte formal fácil de ser implementado (ainda que computacionalmente dispendioso);
- Não foram incluídos Quasares nem Seyferts do tipo 1 em nossa amostra geral e, portanto, tampouco as incluímos em nossas Bases de referência e amostras de Objetos. Em um estudo mais completo, estes tipos de galáxias devem ser incluídos;
- O passo mais importante (e também o mais trabalhoso) será incluir o redshift na análise. Por simplicidade, e para focar nosso trabalho no método em si, optamos por trabalhar com espectros sempre no referencial de repouso ( $z = 0$ ), mas na prática as galáxias serão observadas entre  $z = 0$  e  $\sim 1$  (no caso do JPAS). Pode-se imaginar que  $z$  será dado pelos outros membros da colaboração que medirão com grande precisão para seus fins cosmológicos ou então incorporar  $z$  como uma propriedade a mais a ser determinada em nossas análises. A maior dificuldade, no entanto, será outra: para  $z > 1$  mais de 30% da faixa espectral de repouso a ser coberta pelo JPAS estará fora daquela coberta pelo SDSS e isto provavelmente comprometerá a qualidade dos resultados obtidos. Simulações detalhadas neste sentido devem ser realizadas.

O trabalho teórico acima descrito é interessante e importante, mas a prova real do método seria a aplicação à dados reais. Apesar de termos centrado nossas simulações neste trabalho ao JPAS, podemos adaptar a metodologia para qualquer sistema fotométrico. Assim que, do ponto de vista prático podemos enumerar os próximos passos como a aplicação aos dados do ALHAMBRA que já tem fotometria em 20 filtros de largura média de 300 Å disponível e, indo mais além, aplicar o método com os dados fotométricos do próprio SDSS (apenas 5 filtros de 1000 Å). Naturalmente, para o SDSS, esperamos um resultado muito pior que para o ALHAMBRA, mas esta aplicação pode ser vista como um exercício que pode ser útil para nos dizer sobre o que pode ser recuperado com outros surveys similares que estão por vir (com destaque para o LSST).

## 6.3 RESTOS...

Abdalla, Sodre, Lahav?


!☺j☺! PRA APROVEITAR NALGUM LUGAR ....

UMA CRITICA QUE SE PODE FAZER DESTE METODO EH QUE ELE NAO VAI NOS ENSINAR NADA QUE O SDSS JAH NAO TENHA ENSINADO. APE-SAR DE PERTINENTE, ESTA CRITICA NAO INVALIDA NOSSO TRABALHO. P EX, EXISTE UM MONTAO DE CASOS ONDE NAO QUEREMOS APRESNDER NADA NOVO SOBRE AS PROPRIEDADES ESPECTROSCOPICAS DE GALAX-IAS, MAS SIMPLEMENTE IDENTIFICAR SEUS TIPOS ESPECTRAIS E COM ISSO FAZER CIENCIA. UM EXEMPLO SERIA A CLASSIFICACAO DE GALAXIAS EM SF/AGN/... FAZENDO A CLASSIFICACAO COM NOSSO METODO, E UTI-LIZANDO A ESTATISTICA DO JPAS, PODEMOS CONSTRUIR FUNCOES DE LU-MINOSIDADE DE DIFERENTES TIPOS DE GALAXIAS EM DIFERENTES RED-HIFTS E DIFERENTES AMBIENTES. OUTRO EXEMPLO SERIA CORRELACIONAR A CLASSIFICACAO ESPECTRAL COM A INFORMACAO MORFOLOGICA QUE AS IMAGENS APORTARAO. COM ISSO PODERIAMOS DEFINIR A FREQUEN-CIA DE AGNS EM MERGERS, NO CAMPO, EM GRUPOS E AGLOMERADOS, entre OUTRAS APLICACOES CIENTIFICAENTE RELEVANTES.

TB CABE OBSERVAR QUE A CRITICA DE QUE “NAO APRENDEREMOS NADA DE NOVO” NAO EH INTEIRAMENTE CORRETA. AO CONSTRUIR UMA BASE SE SUPOE QUE ELA MAPEIA TODO O UNIVERSO DE POSSIVEIS ES-

PECTROS DE GALAXIES, O QUE LIMITA O POTENCIAL DE DESCOBERTAS. ISSO EH VERDADE, MAS TB EH CERTO QUE O SPECTRAL MATCHING SEMPRE VAI NOS DAR UMA FIGURA DE MERITO. UMA GALAXIA SUBSTANCIALMENTE DIFFERENTE DAS QUE TAO NA BASE PODERAH SER RECONHECIDA COMO TAL POR ESTA FIGURA DE MERITO. OBJETOS QUE NAO SAO BEM AJUSTADOS PELA BASE PODEM ENTAO SER SEPARADOS E ESTUDADOS DE OUTRAS MANEIRAS, SEJA ATRAVES DE BASES ESPECIAIS, DADOS EM OUTRAS FAIXAS ESPECTRAIS OU FOLLOW-UP OBSERVATIONS.

## 6.4 lixo possivelmente reciclavel...

PARA VER COMO FICA OS MATCHES COM ESTAS 2 BASES DE REFERENCIA, ESCOLHEMOS ALGUMAS GALAXIAS DE TODO TIPO (SF/AGN/ RG/ PASSIVA) E BUSACMOS OS MELHORES PARES NAS BASES CMD E WHAN. AS FIGS MOSTRAM OS RESULTADOS. A FIG ?? MOSTRA OS J-SPECTROS ORIGINAL (EM PRETO, COLUNA DA ESQUERDA), E OS MELHORES PARES NAS BASES CMD (AZUL, AO MEIO) E WHAN (VERMELHO, A DIREITA). NA FIGURA SEGUINTE MOSTRAMOS OS ESPECTROS SDSS CORRESPONDENTES.  **TEM QUE FAZER!!!!**

DAH PRA VER QUE ???TEM-Q-FAZER PRA VER O QUE SAI!!!

COMO DITO ACIMA, ESPERA-SE QUE A BASE CMD PRODUZA MELHORES AJUSTES DE PROPRIEDADES FISICAS, E QUA A BASE WHAN PROPORCIONE MELHORES AJUSTES DE PROPRIEDADES DE LINHAS DE EMISSAO (PARTIULARMENTE HA E [NII]). ESTA EXPECTATIVA SERAH MELHOR ANALISADA MAIS ADIANTE, MAS PARA TER UMA IDEIA RAPIDA FAZEMOS O SEGUINTE EXERCICIO: PEGAMOS UNS TANTOS OBJETOS E CALCULAMOS A DIFERENCA ENTRE OS VALORES DE ENTRADA E O DO BEST MATCH PARA:

- IDADE MEDIA, METALICIDADE MEDIA E A EXTINCAO ESTELAR  $A_V^*$  (PARA REPRESENTAR O CONJUNTO DE PROPRIEDADES FISICAS)
- WHA, N2HA, O3HB e HAHB (REPRESENTANDO LINHAS DE EMISSAO).

FIZEMOS ISSO PARA AS BASES CMD E WAHN. AS FIGS ?? E ?? MOSTRAM OS RESULTADOS DESTE EXPERIMENTO RAPIDO. NELAS COMPARAMOS EN-

TRADA E SAIDA PARA 10000 OBJETOS. SE VE QUE ???? ...C/UM POUCO DE  
SORTE A WHAN DARAH WHAN E N2HA MELHORES QUE A CMD,  
E A BASE CMD DARA MELHORES LOGT E Z...

# Referências Bibliográficas

- Asari, N. V., Cid Fernandes, R., Stasińska, G., Torres-Papaqui, J. P., Mateus, A., Sodré, L., Schoenell, W., & Gomes, J. M. 2007, *MNRAS*, 381, 263
- Baldry, I. K., Glazebrook, K., Brinkmann, J., Ivezić, Ž., Lupton, R. H., Nichol, R. C., & Szalay, A. S. 2004, *ApJ*, 600, 681
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
- Bell, E. F., Wolf, C., Meisenheimer, K., Rix, H.-W., Borch, A., Dye, S., Kleinheinrich, M., Wisotzki, L. et al. 2004, *ApJ*, 608, 752
- Benítez, N. 2000, *ApJ*, 536, 571
- Benítez, N., Gaztañaga, E., Miquel, R., Castander, F., Moles, M., Croce, M., Fernández-Soto, A., Fosalba, P. et al. 2009, *ApJ*, 691, 241
- Brinchmann, J., Charlot, S., White, S. D. M., Tremonti, C., Kauffmann, G., Heckman, T., & Brinkmann, J. 2004, *MNRAS*, 351, 1151
- Bruzual, G. & Charlot, S. 2003, *MNRAS*, 344, 1000
- Cid Fernandes, R., Stasińska, G., Mateus, A., & Vale Asari, N. 2011, *MNRAS*, 413, 1687
- Cid Fernandes, R., Stasińska, G., Schlickmann, M. S., Mateus, A., Vale Asari, N., Schoenell, W., & Sodré, L. 2010, *MNRAS*, 403, 1036
- Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, *MNRAS*, 362, 41
- Kauffmann, G., Heckman, T. M., White, S. D. M., Charlot, S., Tremonti, C., Brinchmann, J., Bruzual, G., Peng, E. W. et al. 2003, *MNRAS*, 341, 33
- York, D. G., Adelman, J., Anderson, J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R. et al. 2000, *AJ*, 120, 1579