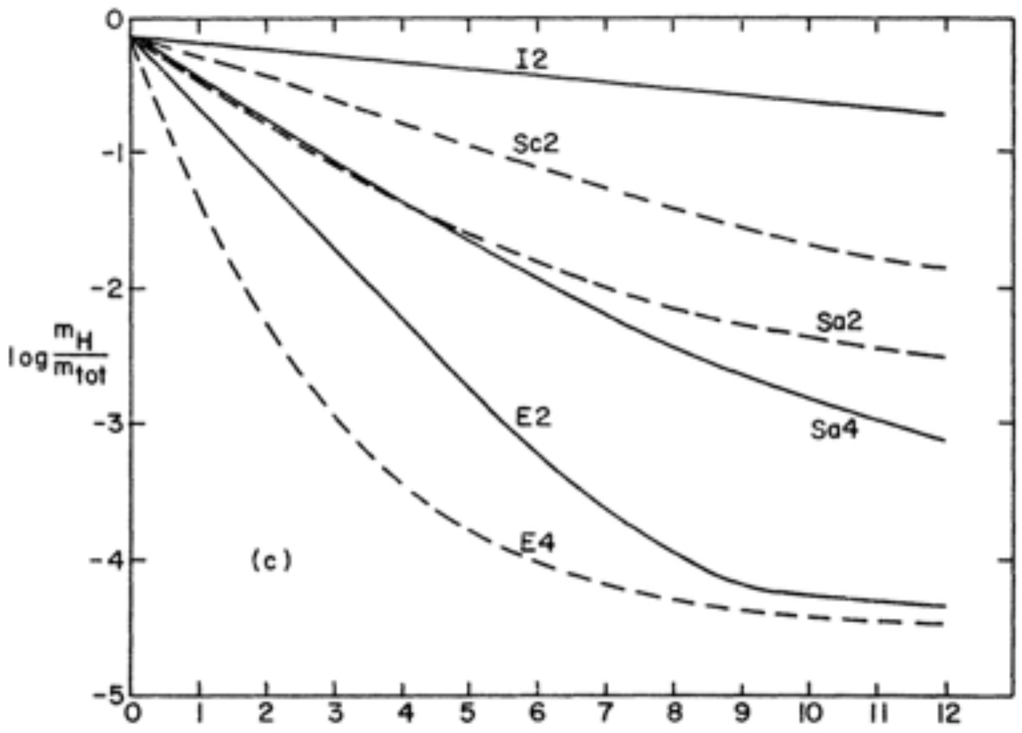
JARLE BRINCHMANN (IA CAUP PORTO/LEIDEN)

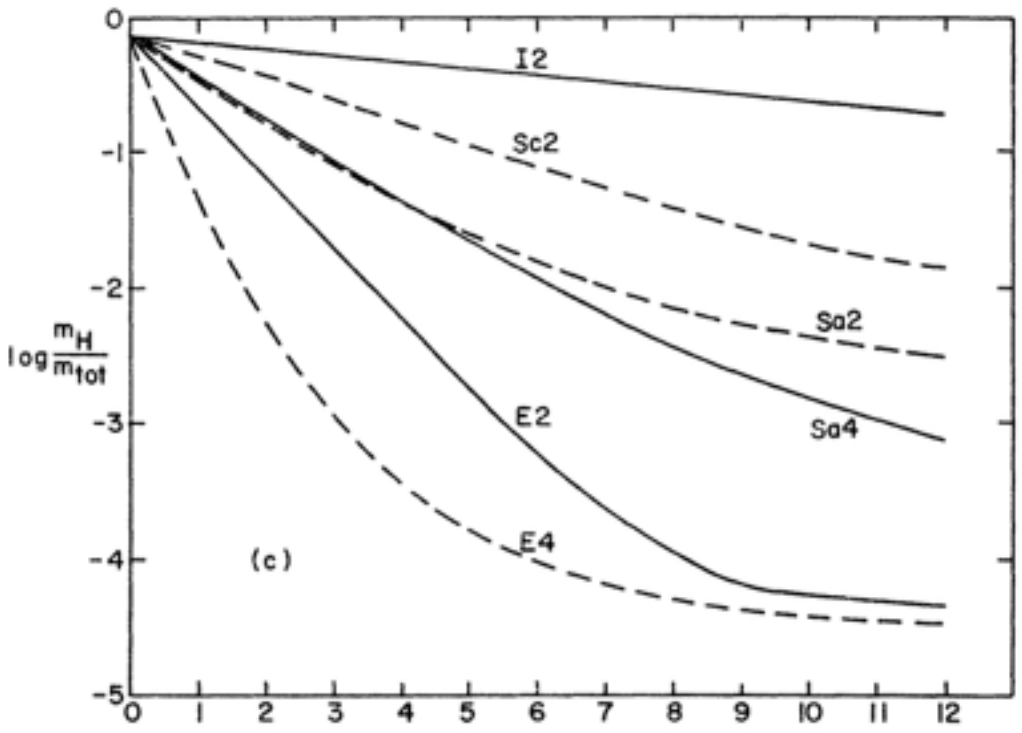
FROM AN OBSERVED SED TO PHYSICAL PARAMETERS

THE NEED FOR BETTER STATISTICAL TECHNIQUES



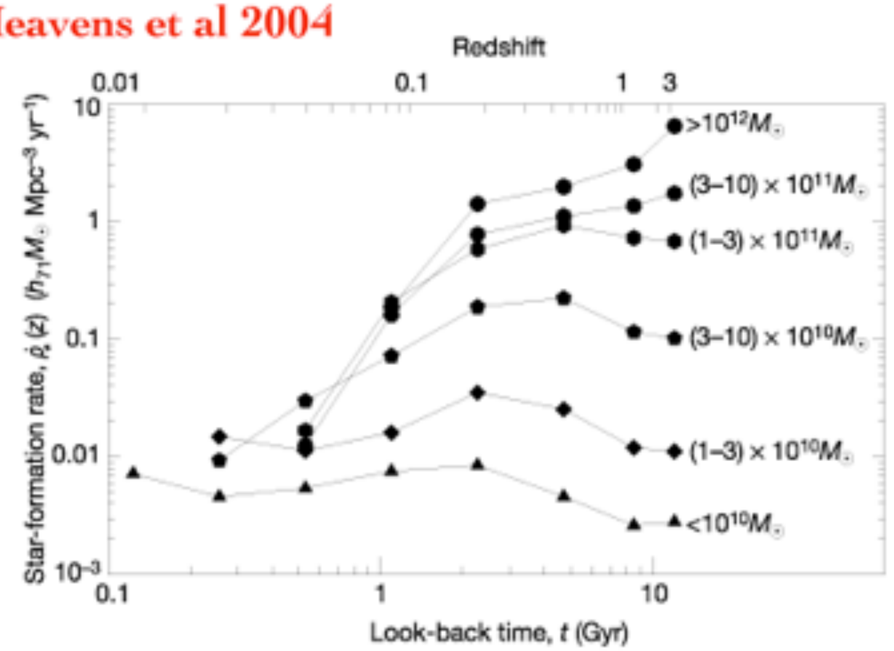
Tinsley's (1968) "star formation history" diagram.

THE NEED FOR BETTER STATISTICAL TECHNIQUES



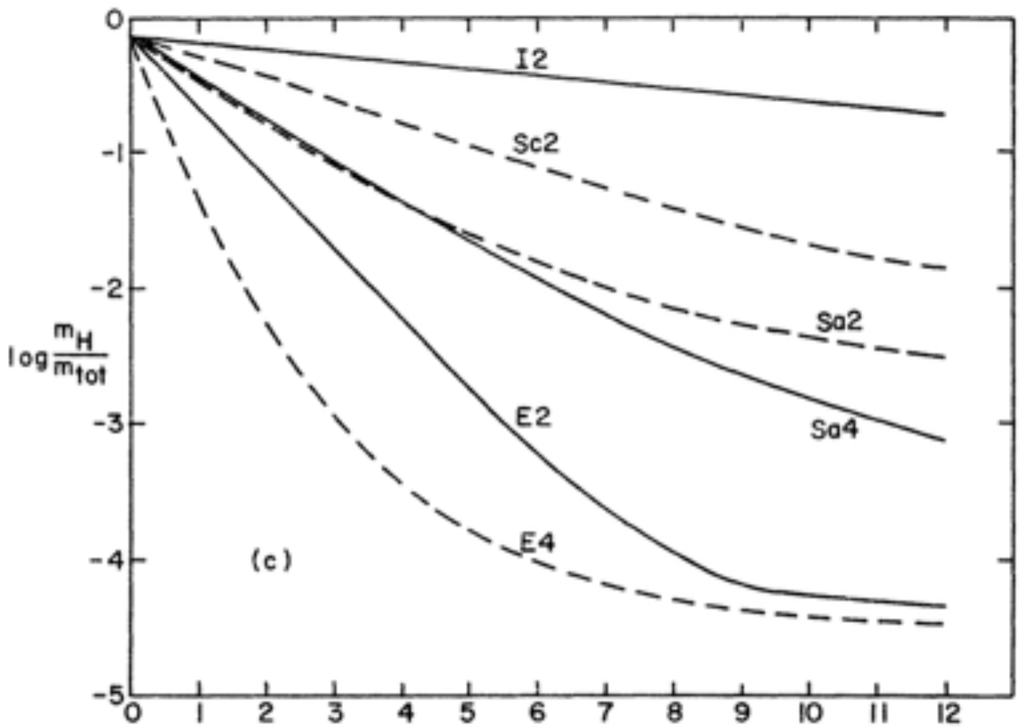
Tinsley's (1968) "star formation history" diagram.

Heavens et al 2004

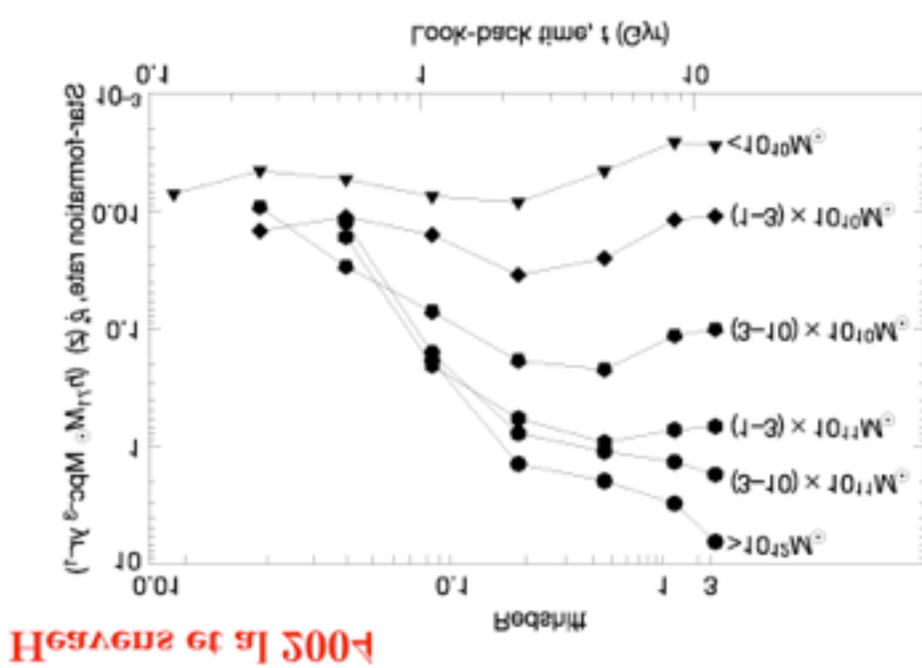


Heavens et al's (2004) paper on the star formation history of nearby galaxies from SDSS.

THE NEED FOR BETTER STATISTICAL TECHNIQUES

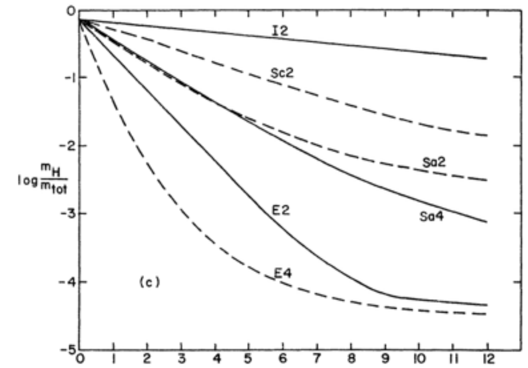


Tinsley's (1968) "star formation history" diagram.

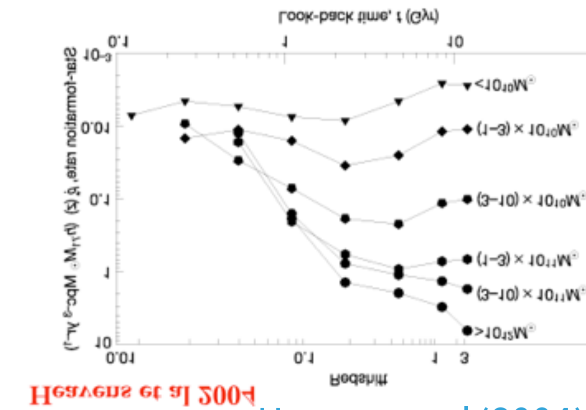


Heavens et al's (2004) paper on the star formation history of nearby galaxies from SDSS.

THE NEED FOR BETTER STATISTICAL TECHNIQUES



Tinsley (1968)



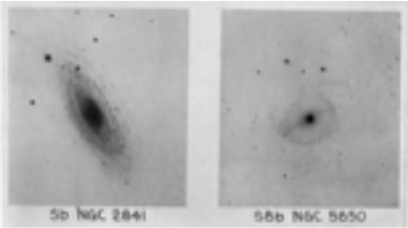
Heavens et al (2004)

Some parameters are robustly constrained - most methods should give similar results, when using the same assumptions.

However we do not necessarily know how well we do.

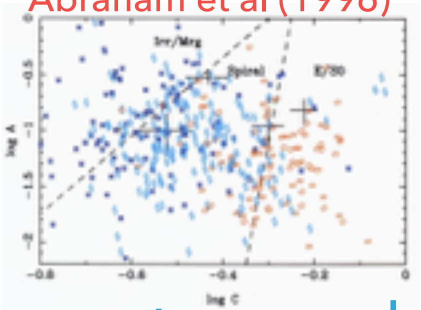
Morphology:

Hubble (1926)



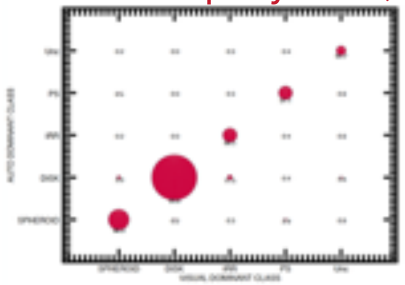
Visual classification

Abraham et al (1996)



Computer morphology

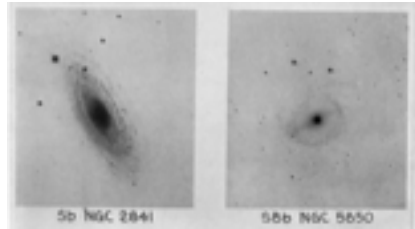
Huertas-Company et al (2015)



Deep learning/SVMs/
Neural networks

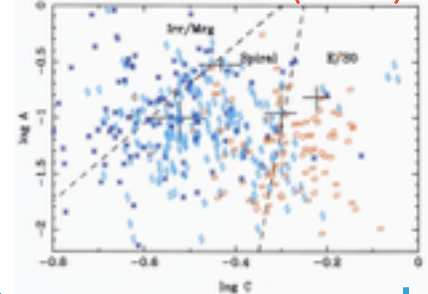
Morphology:

Hubble (1926)



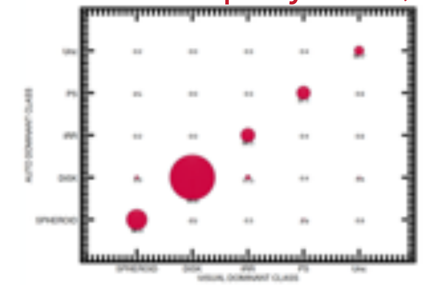
Visual classification

Abraham et al (1996)



Computer morphology

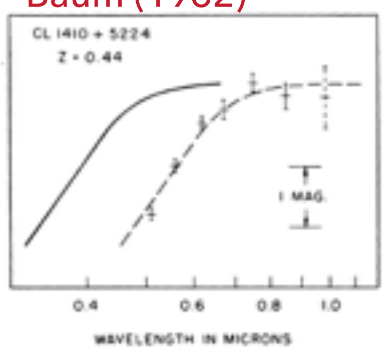
Huertas-Company et al (2015)



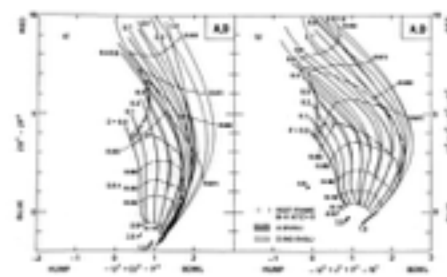
Deep learning/SVMs/
Neural networks

Photo-zs:

Baum (1962)



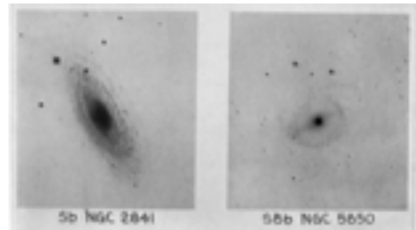
Koo (1985)



Template fits (Loh & Spillar 1986)
Linear regression (Connolly et al 1995)
Neural networks (Collister & Lahav 2004)
Random forests, Gaussian process regression, Support Vector Machines, +++

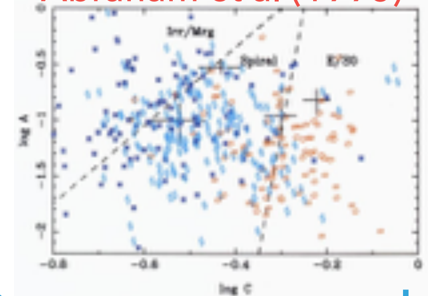
Morphology:

Hubble (1926)



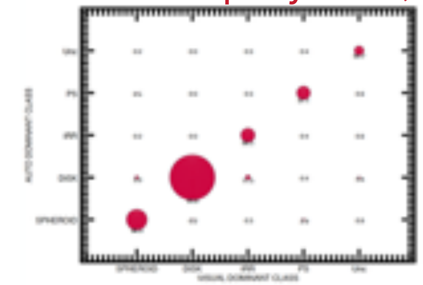
Visual classification

Abraham et al (1996)



Computer morphology

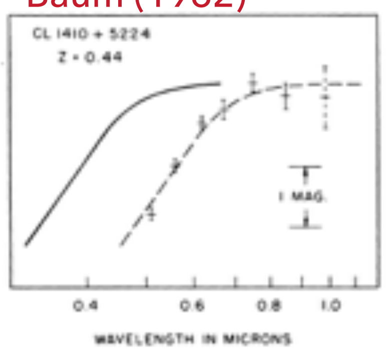
Huertas-Company et al (2015)



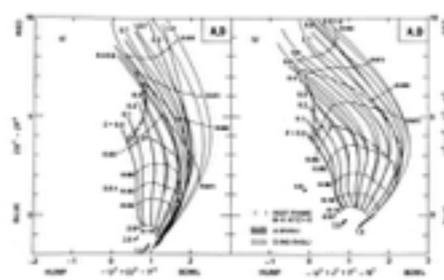
Deep learning/SVMs/
Neural networks

Photo-zs:

Baum (1962)



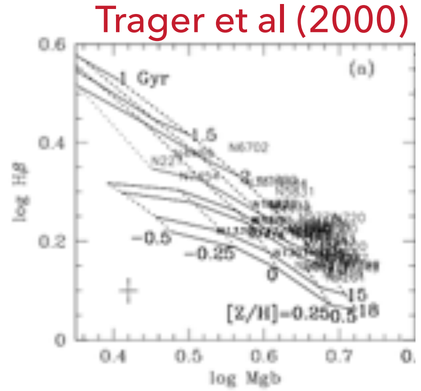
Koo (1985)



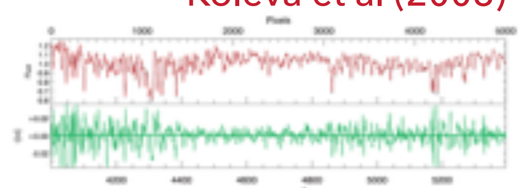
Template fits (Loh & Spillar 1986)
Linear regression (Connolly et al 1995)
Neural networks (Collister & Lahav 2004)
Random forests, Gaussian process regression, Support Vector Machines, +++

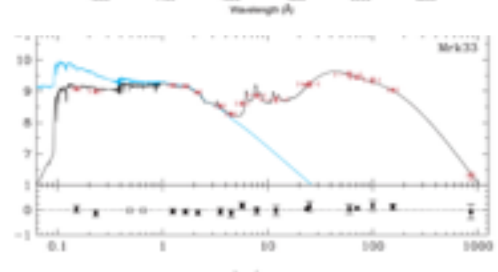
SED fitting:

Trager et al (2000)



Koleva et al (2008)





da Cunha et al (2008)

Min χ^2

Grid-based Bayesian

MCMC Bayesian

THE NEED FOR BETTER STATISTICAL APPROACHES

How well do we know what we know?

How can we fully exploit large datasets?

But: We need the right ingredients!

THE NEED FOR BETTER STATISTICAL APPROACHES

How well do we know what we know?

An estimate is insufficient - we need reliable confidence/credible intervals

Needs:

- Rigorous analysis exploring parameter space properly.
- A deep understanding of our photon (typically) gathering process.

How can we fully exploit large datasets?

But: We need the right ingredients!

THE NEED FOR BETTER STATISTICAL APPROACHES

How well do we know what we know?

An estimate is insufficient - we need reliable confidence/credible intervals

Needs:

- Rigorous analysis exploring parameter space properly.
- A deep understanding of our photon (typically) gathering process.

How can we fully exploit large datasets?

Lots of data ≠ easily lots of (good) science

Needs:

- Fast algorithms, low (or calibratable) bias

But: We need the right ingredients!

THE IMPORTANCE OF THE RIGHT INGREDIENTS

Advanced statistical tools are great - **but**:

Any chain (of reasoning) is only as good as its weakest link

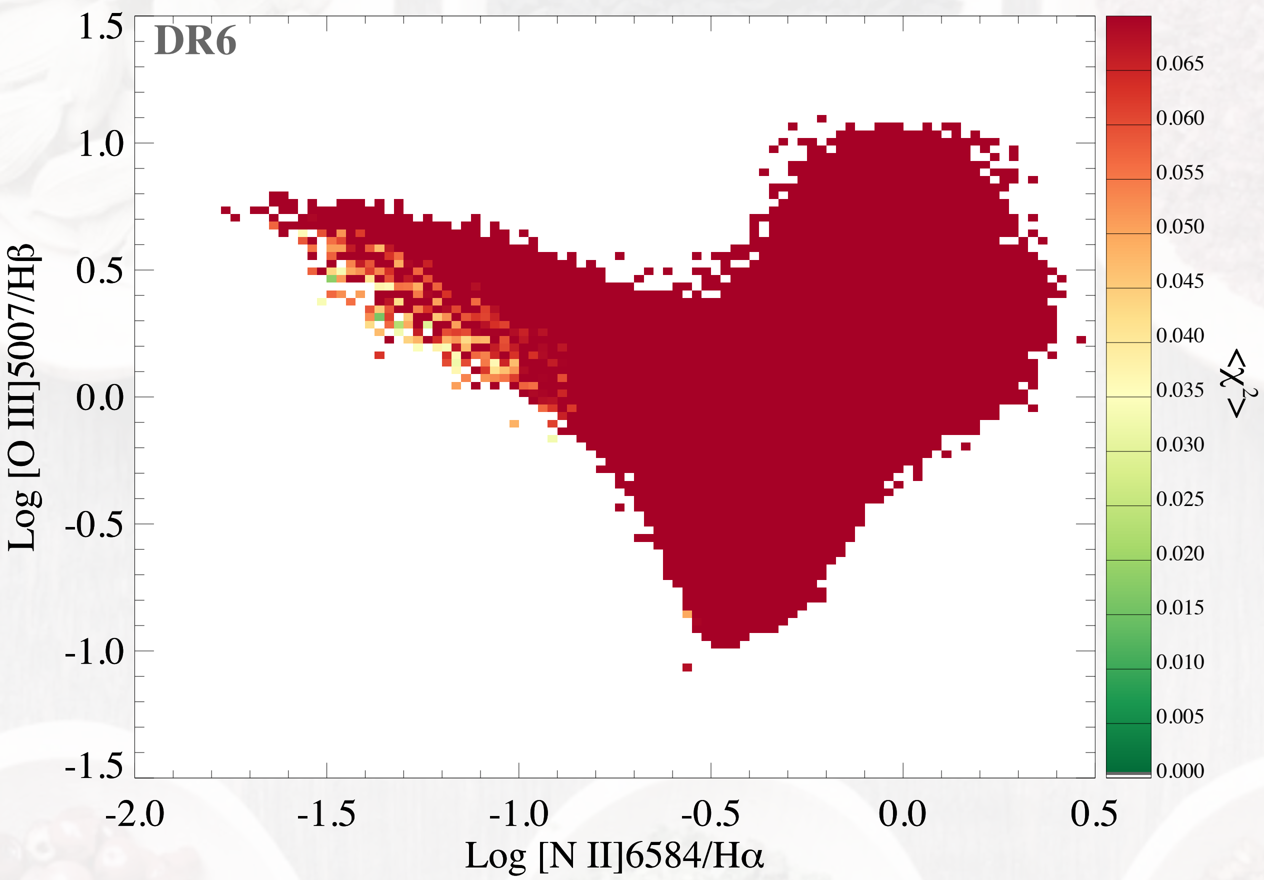
The good news:

Lot of excellent work in the last 10-15 years: Massive progress

BUT: This remains a key uncertainty in much of our work:

Binaries, stellar rotation, α-enhancement, IMF, far UV spectra, turbulence, …

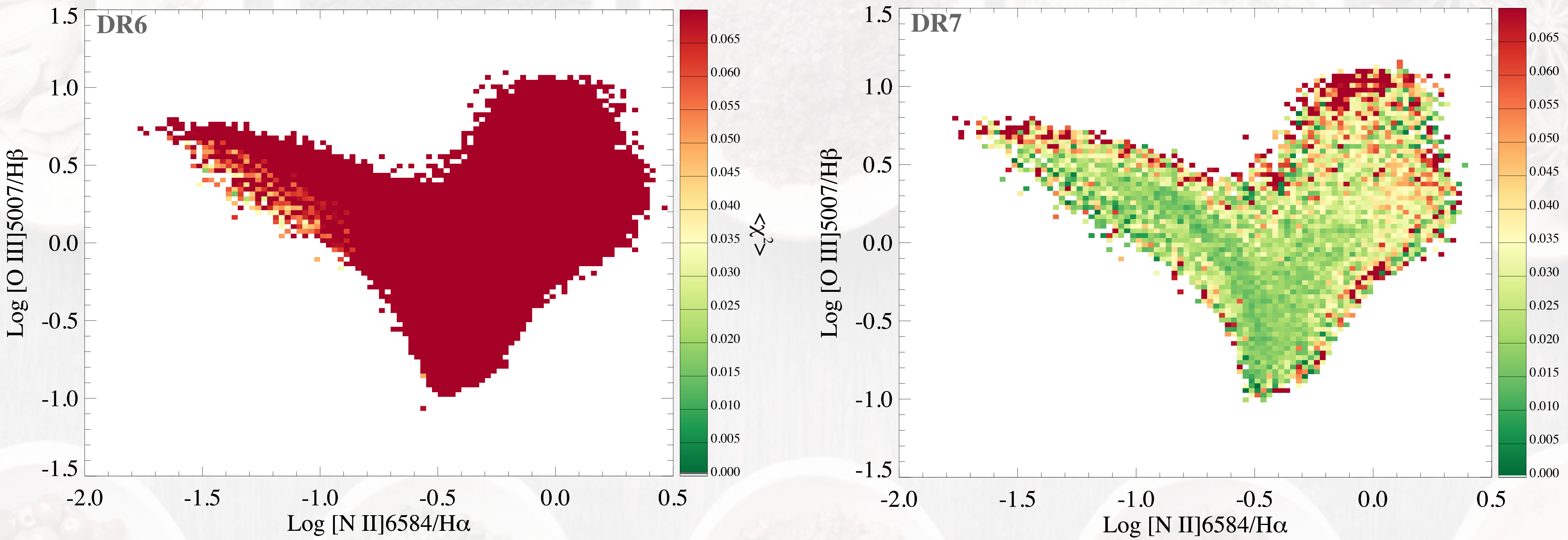
THE IMPORTANCE OF THE RIGHT INGREDIENTS



The quality of fit to the continuum in ~200,000 SDSS galaxies.

Moving to a better basis set for the continuum fit (in this case MILES spectra), made much more of a difference than changes in the algorithm used for fitting.

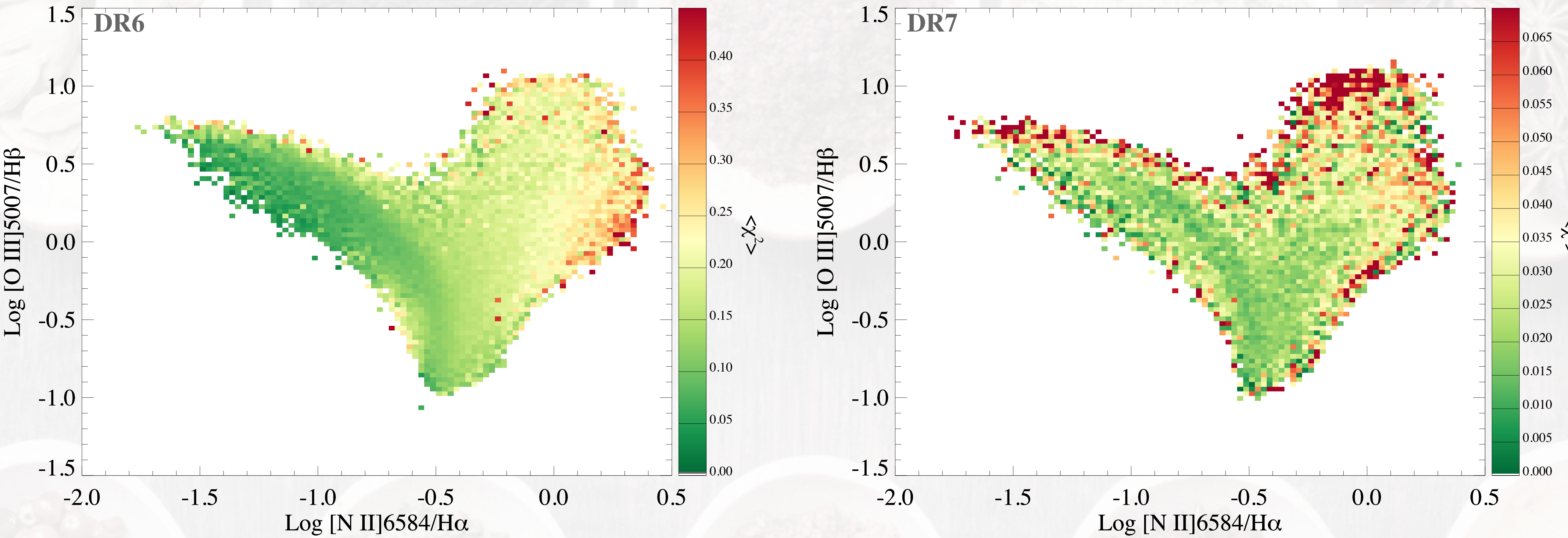
THE IMPORTANCE OF THE RIGHT INGREDIENTS



The quality of fit to the continuum in ~200,000 SDSS galaxies.

Moving to a better basis set for the continuum fit (in this case MILES spectra), made much more of a difference than changes in the algorithm used for fitting.

THE IMPORTANCE OF THE RIGHT INGREDIENTS



The quality of fit to the continuum in ~200,000 SDSS galaxies.

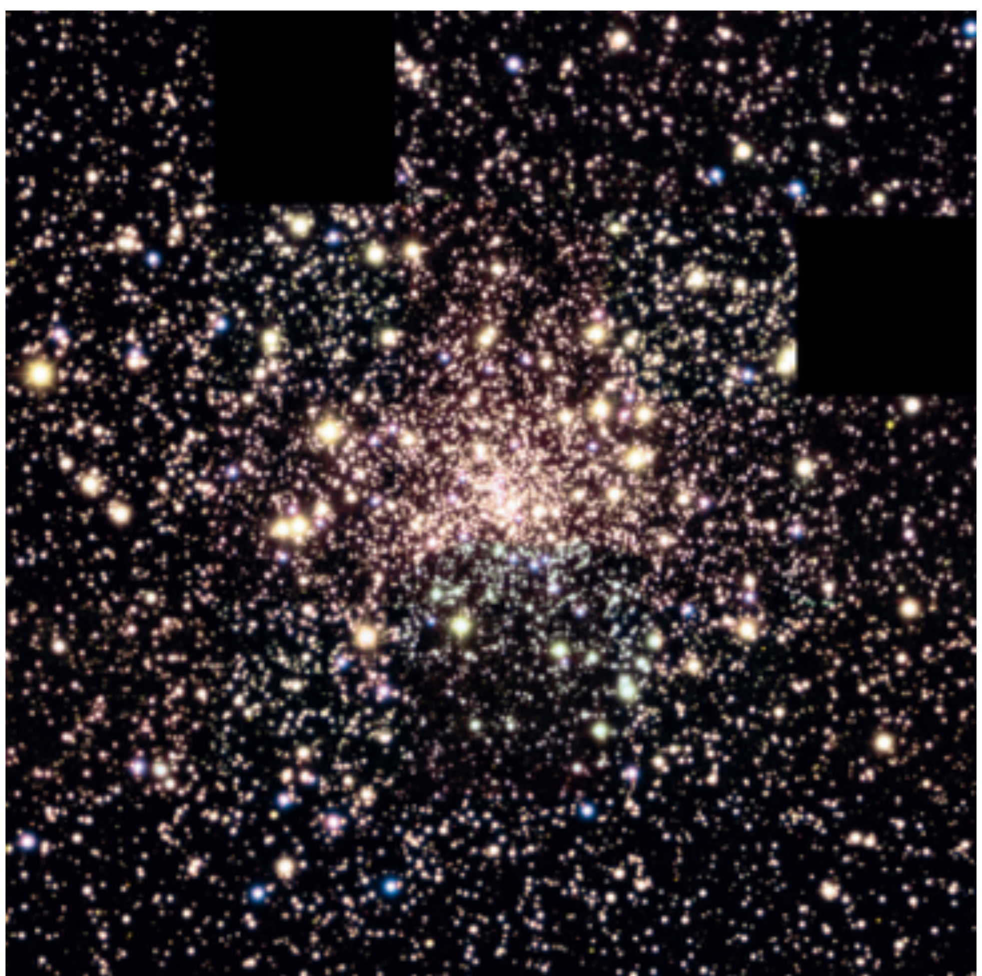
Moving to a better basis set for the continuum fit (in this case MILES spectra), made much more of a difference than changes in the algorithm used for fitting.

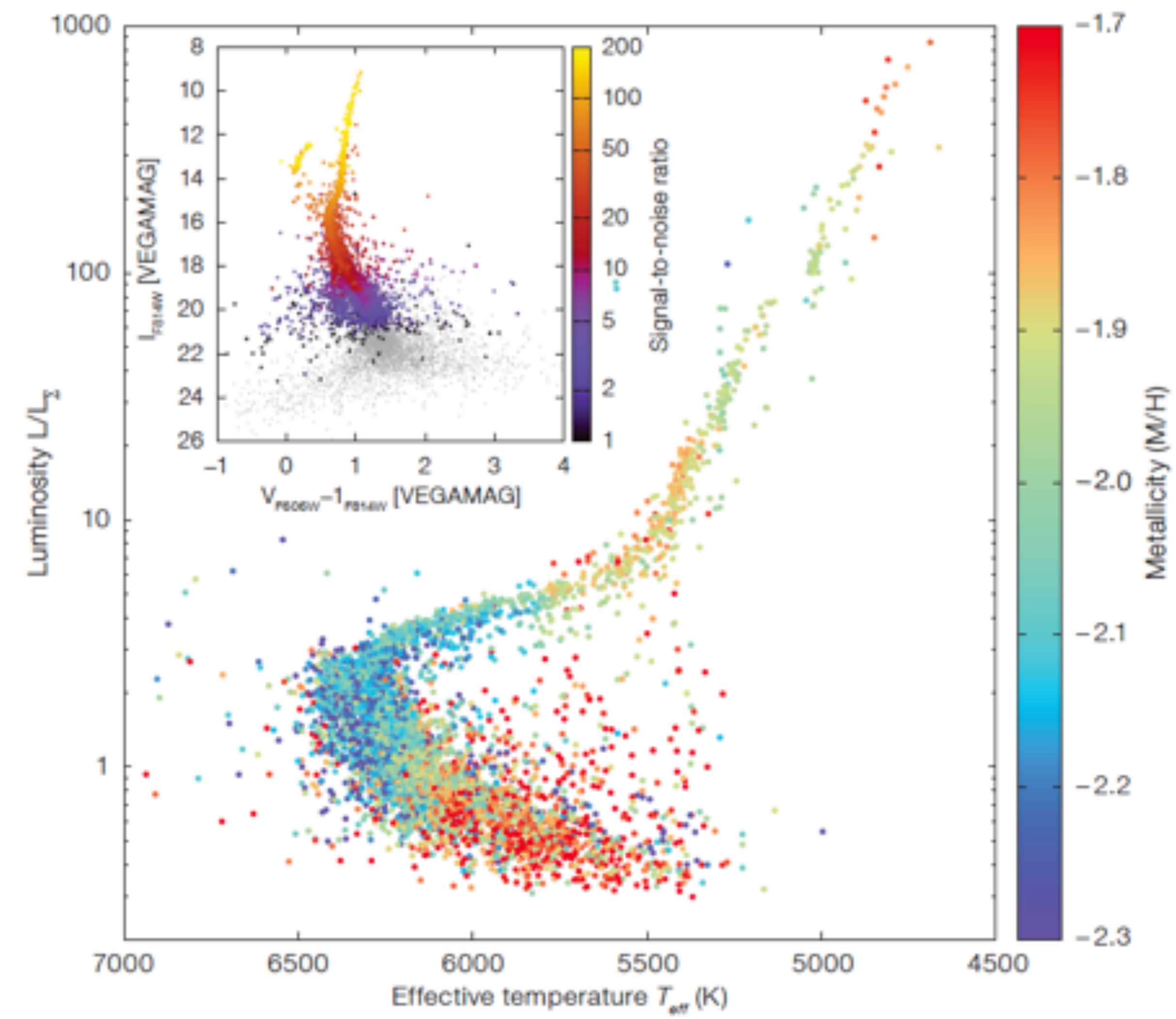
CHECKING THE INGREDIENTS

Classical: Milky Way globulars - e.g. Maraston et al (2003); Barber et al (2014)

Next step?: IFU spectroscopy of stellar systems & HST CMDs

Kamann et al (2016) - A Stellar Census in NGC 6397 with MUSE





Preliminary results (MSc thesis Pietrow): SED modelling consistent with CMD

FULL-SPECTRUM FITTING

The move from index-index plots has led to a proliferation of methods to fit the full spectrum of galaxies:

FULL-SPECTRUM FITTING

The move from index-index plots has led to a proliferation of methods to fit the full spectrum of galaxies:

MOPED k-correct STECKMAP Pipe3D
platefit STARLIGHT ppxf
VESPA NBURSTS ULySS BEAGLE
GASPEX SEDFIT FADO +++++

LePhare, GalMC, CIGALE, MagPhys, BayeSED, FAST, HyperZ, GP-CV, Annz, Stable-GP, EAZY, BPZ, ++++++

FULL-SPECTRUM FITTING

The move from index-index plots has led to a proliferation of methods to fit the full spectrum of galaxies:

MOPED k-correct STECKMAP Pipe3D
platefit STARLIGHT ppxf
VESPA NBURSTS ULySS BEAGLE
GASPEX SEDFIT FADO +++++

LePhare, GalMC, CIGALE, MagPhys, BayeSED, FAST, HyperZ, GP-CV, Annz, Stable-GP, EAZY, BPZ, ++++++

These have used a range of methods: Principal Component Analysis, Non-negative least squares, non-negative matrix factorisation, bounded value least-squares, MCMC + χ^2, constrained minimisation with generalised CV etc.

But beyond these optimisation methods the main progression has been from minimum χ^2 to more rigorous Bayesian approaches.

THE APPEAL OF BAYESIAN APPROACHES

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Model evidence}}$$

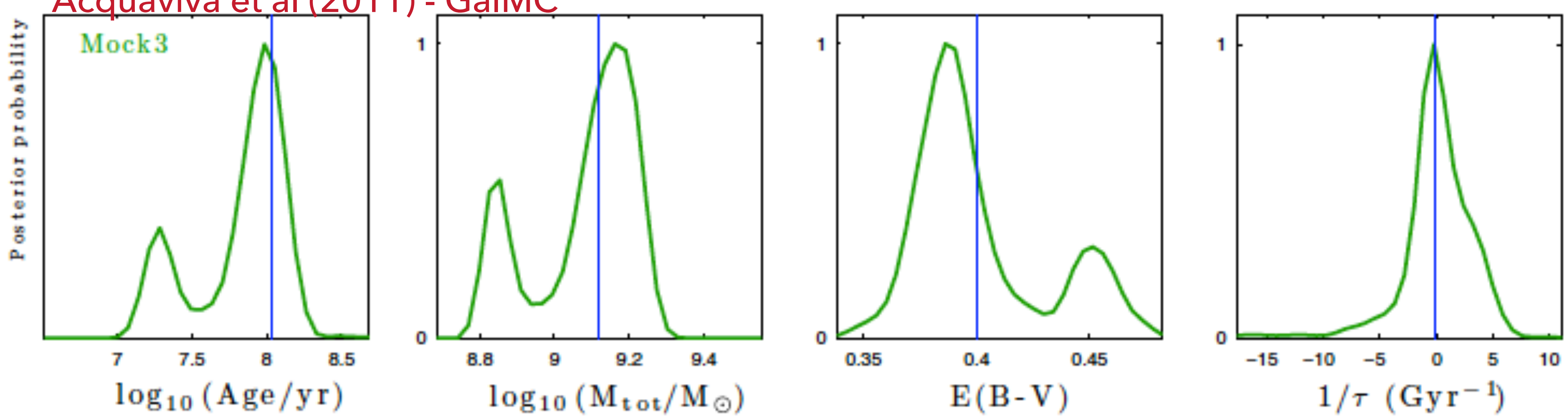
✓ Need to explicitly state (many) assumptions

✓ Complex posterior distributions can be handled

✓ Efficient MCMC methods exist to sample high-dimensional problems

✓ Potentially flexible framework to handle systematic uncertainties (SFHs?)

Acquaviva et al (2011) - GalMC



- Can be time-consuming
- Can be appear misleadingly “rigorous”

THE LIKELIHOOD

Standard approach:

$$f^{\text{obs}}(\lambda) = f^{\text{true}}(\lambda) + \epsilon \quad \& \quad \epsilon \sim N\left(0, \sigma^2\right)$$

so

$$\log L_j = \sum_i \frac{\left[f(\lambda_i) - M_j(\lambda_i)\right]^2}{\sigma_i^2}$$

but this supposes that we know the variance - if it is estimated from the data or uncertain, this is not correct and you need a Student t-distribution:

$$L\left(x_i|\nu, \sigma_i\right) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}\sigma_i}\left(1 + \frac{1}{\nu}\left(\frac{x_i}{\sigma_i}\right)^2\right)^{-\frac{\nu+1}{2}}$$

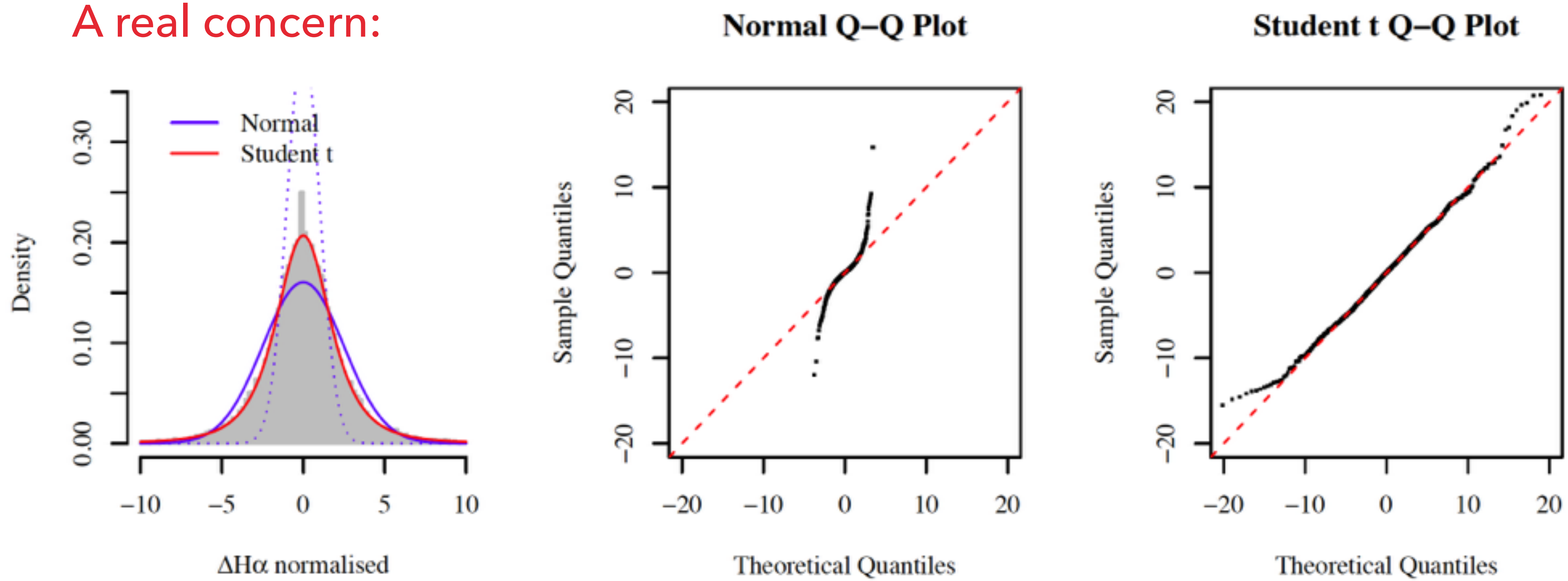
$$x_i = f(\lambda_i) - M_j(\lambda_i)$$

THE LIKELIHOOD

Standard approach:

$$f^{\text{obs}}(\lambda) = f^{\text{true}}(\lambda) + \epsilon \quad \& \quad \epsilon \sim N\left(0, \sigma^2\right)$$

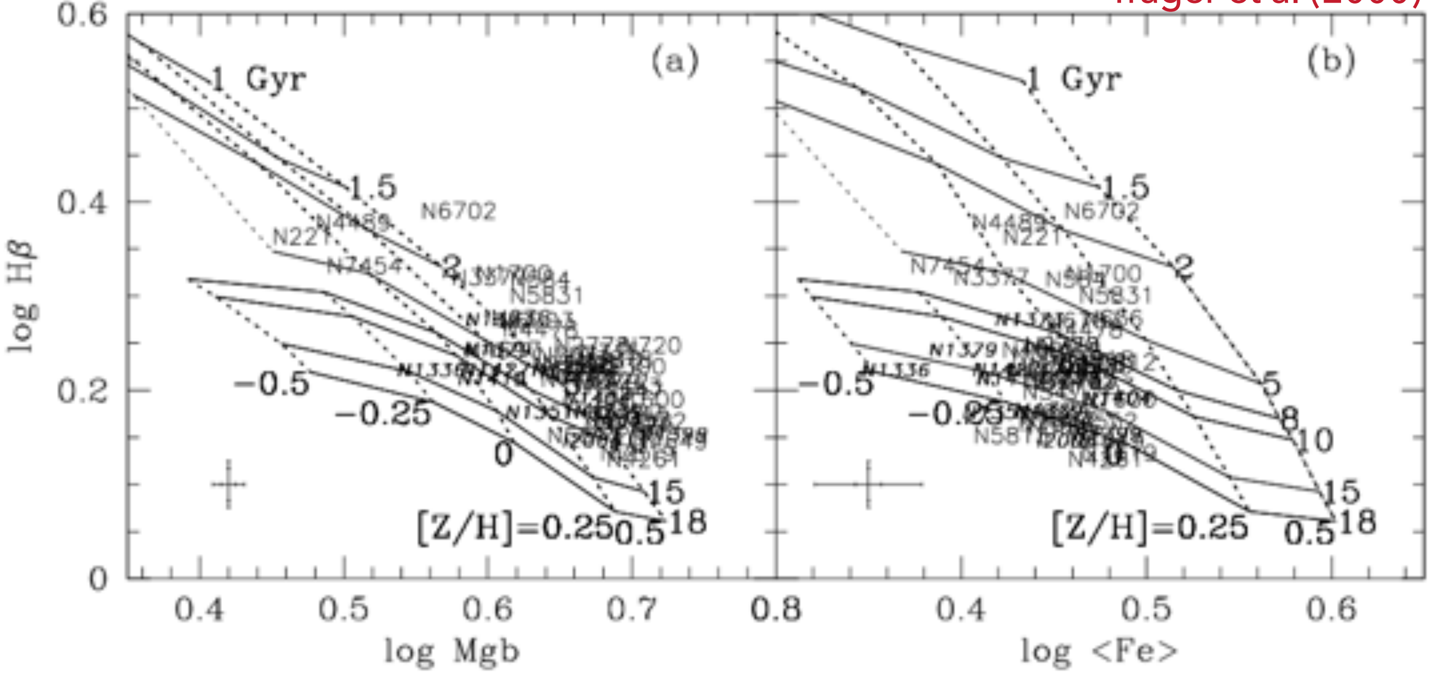
A real concern:



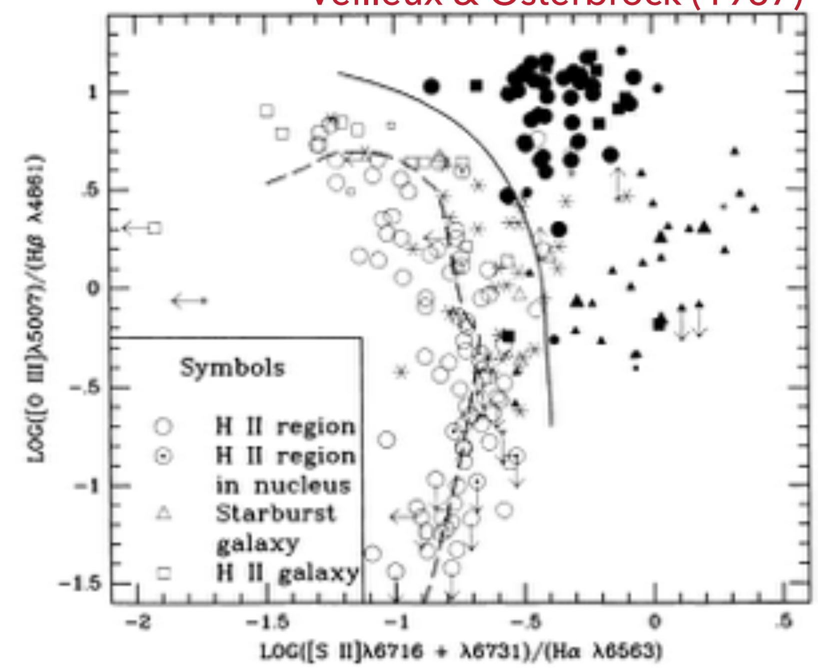
$$x_i = f(\lambda_i) - M_j(\lambda_i)$$

INDEX–INDEX & DIAGNOSTIC DIAGRAMS

Trager et al (2000)



Veilleux & Osterbrock (1987)



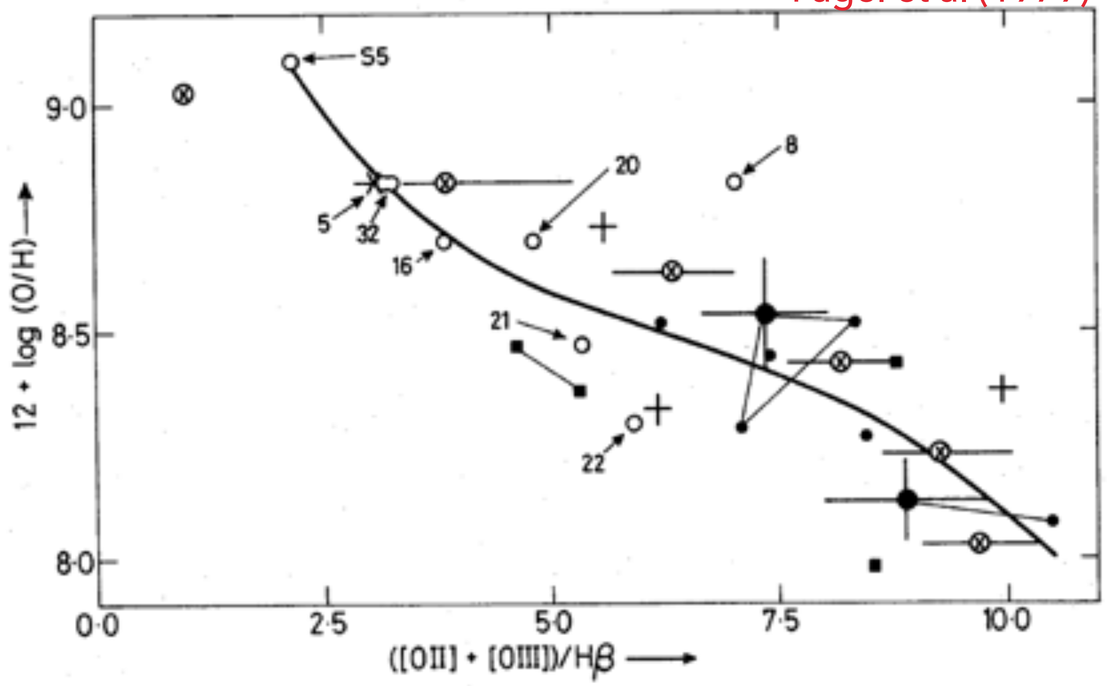
Are these now a waste of time?

Certainly not - they provide physical insight, usually much more than a triangle plot of posterior samples from an MCMC chain.

EMISSION LINES

INFERRING IONISED GAS PROPERTIES

Pagel et al (1979)



While the theoretical models have advanced, many studies today use analysis techniques similar to this.

$$12 + \log \mathrm{O/H} = 7.056 + 0.767\mathrm{R}_{23} + 0.602\mathrm{R}_{23}^2 - \mathrm{O}_{32}\left(0.29 + 0.332\mathrm{R}_{23} - 0.331\mathrm{R}_{23}^2\right)$$

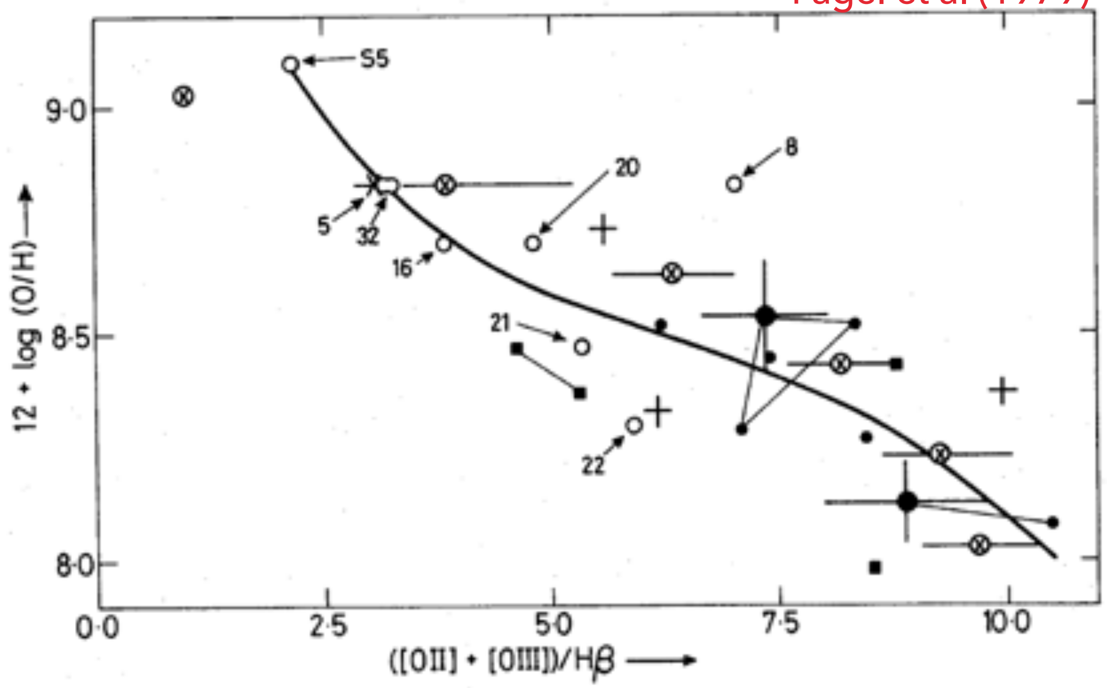
$$\begin{aligned}\log[\mathrm{N\,II}]/[\mathrm{O\,II}] =& 1106.8660 - 532.1545112 + \log \mathrm{O/H} + \\ & 96.3732612 + \log \mathrm{O/H}^2 - 7.810612312 + \log \mathrm{O/H}^3 + 0.2392824712 + \log \mathrm{O/H}^4\end{aligned}$$

$$12 + \log \mathrm{O/H} = 8.73 - 0.32\mathrm{O3N2}$$

$$12 + \log \mathrm{O/H} = \frac{\mathrm{R}_{23} + 726.1 + 842.2P + 337.5P^2}{85.96 + 82.76P + 43.98P^2 + 1.793\mathrm{R}_{23}}$$

INFERRING IONISED GAS PROPERTIES

Pagel et al (1979)



While the theoretical models have advanced, many studies today use analysis techniques similar to this.

$$12 + \log \mathrm{O/H} = 7.056 + 0.767\mathrm{R}_{23} + 0.602\mathrm{R}_{23}^2 - \mathrm{O}_{32}\left(0.29 + 0.332\mathrm{R}_{23} - 0.331\mathrm{R}_{23}^2\right)$$

$$\log[\mathrm{N\,II}]/[\mathrm{O\,II}] = 1106.8660 - 532.1545112 + \log \mathrm{O/H} + 96.3732612 + \log \mathrm{O/H}^2 - 7.810612312 + \log \mathrm{O/H}^3 + 0.2392824712 + \log \mathrm{O/H}^4$$

$$12 + \log \mathrm{O/H} = 8.73 - 0.32\mathrm{O3N2}$$

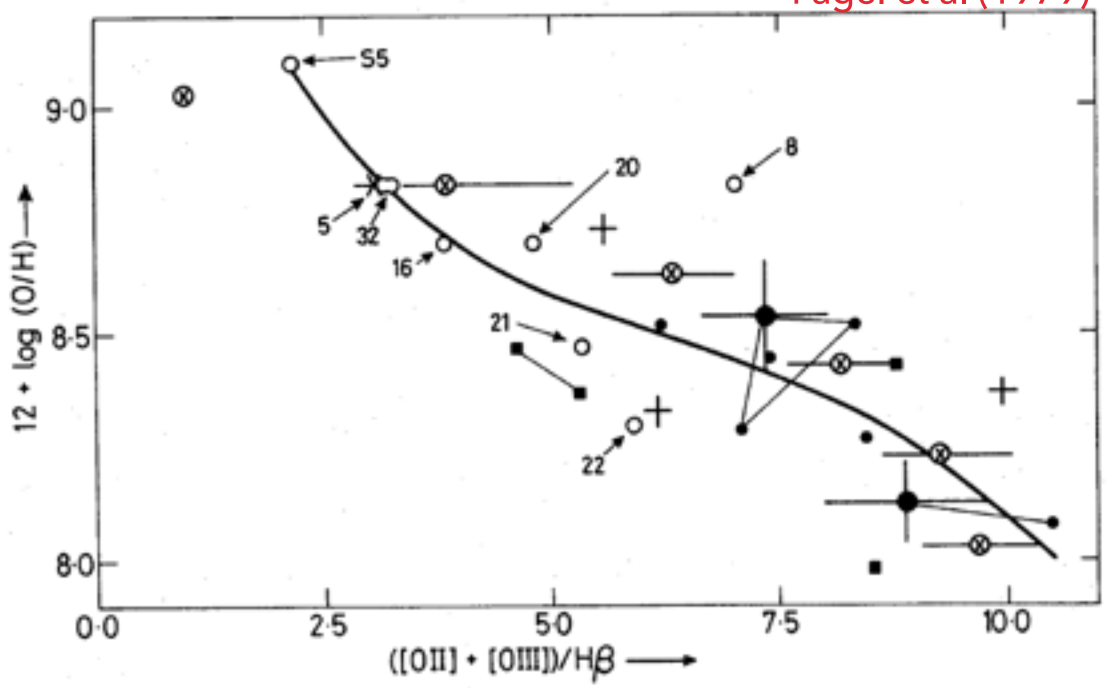
$$12 + \log \mathrm{O/H} = \frac{\mathrm{R}_{23} + 726.1 + 842.2P + 337.5P^2}{85.96 + 82.76P + 43.98P^2 + 1.793\mathrm{R}_{23}}$$

"STRONG-LINE METHODS"

Models → O/H = f(line ratios)

INFERRING IONISED GAS PROPERTIES

Pagel et al (1979)



While the theoretical models have advanced, many studies today use analysis techniques similar to this.

$$12 + \log \mathrm{O/H} = 7.056 + 0.767\mathrm{R}_{23} + 0.602\mathrm{R}_{23}^2 - \mathrm{O}_{32}\left(0.29 + 0.332\mathrm{R}_{23} - 0.331\mathrm{R}_{23}^2\right)$$

$$\begin{aligned}\log[\mathrm{N\,II}]/[\mathrm{O\,II}] =& 1106.8660 - 532.1545112 + \log \mathrm{O/H} + \\ & 96.3732612 + \log \mathrm{O/H}^2 - 7.810612312 + \log \mathrm{O/H}^3 + 0.2392824712 + \log \mathrm{O/H}^4\end{aligned}$$

$$12 + \log \mathrm{O/H} = 8.73 - 0.32\mathrm{O3N2}$$

$$12 + \log \mathrm{O/H} = \frac{\mathrm{R}_{23} + 726.1 + 842.2P + 337.5P^2}{85.96 + 82.76P + 43.98P^2 + 1.793\mathrm{R}_{23}}$$

"STRONG-LINE METHODS"

Models ⟶ O/H = f(line ratios)

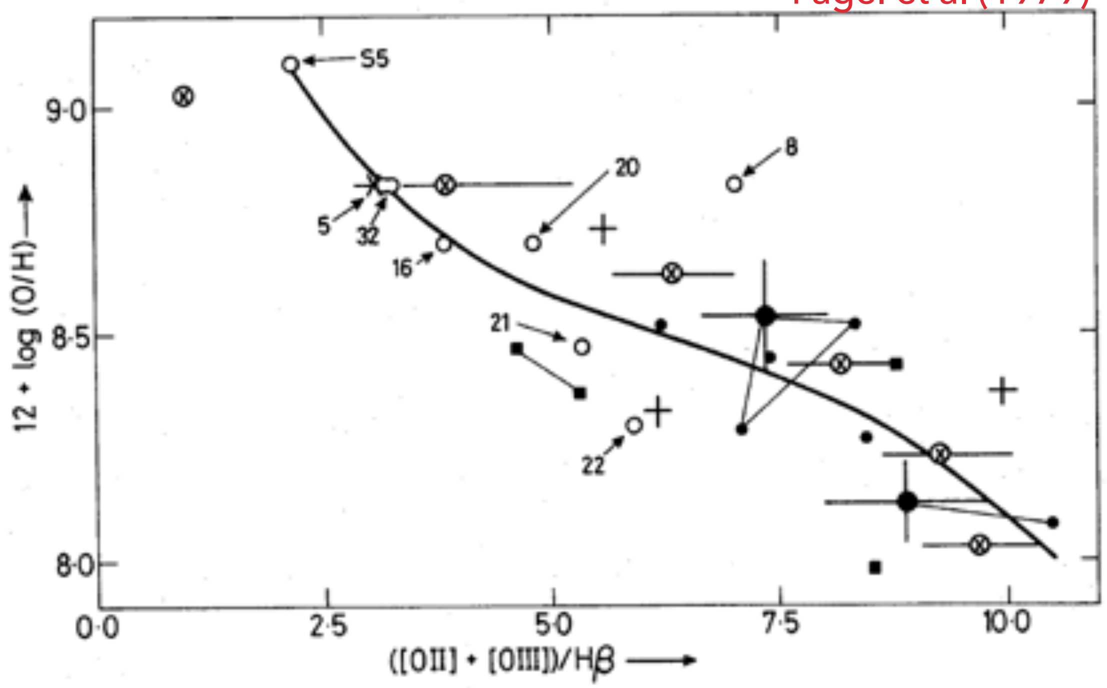
"EMPIRICAL METHODS":

Atomic data ⟶ suitable data (e.g. [O III]4363) ⟶ O/H

O/H = f(line ratios)

INFERRING IONISED GAS PROPERTIES

Pagel et al (1979)



While the theoretical models have advanced, many studies today use analysis techniques similar to this.

$$12 + \log \mathrm{O/H} = 7.056 + 0.767\mathrm{R}_{23} + 0.602\mathrm{R}_{23}^2 - \mathrm{O}_{32}\left(0.29 + 0.332\mathrm{R}_{23} - 0.331\mathrm{R}_{23}^2\right)$$

$$\log[\mathrm{N\,II}]/[\mathrm{O\,II}] = 1106.8660 - 532.1545112 + \log \mathrm{O/H} + 96.3732612 + \log \mathrm{O/H}^2 - 7.810612312 + \log \mathrm{O/H}^3 + 0.2392824712 + \log \mathrm{O/H}^4$$

$$12 + \log \mathrm{O/H} = 8.73 - 0.32\mathrm{O3N2}$$

$$12 + \log \mathrm{O/H} = \frac{\mathrm{R}_{23} + 726.1 + 842.2P + 337.5P^2}{85.96 + 82.76P + 43.98P^2 + 1.793\mathrm{R}_{23}}$$

"STRONG-LINE METHODS"

Models → O/H = f(line ratios)

"EMPIRICAL METHODS"

Atomic data → suitable data (e.g. [O III]4363) → O/H

O/H = f(line ratios)

THESE METHODS CONTAIN IMPLICIT PRIORS

INFERRING IONISED GAS PROPERTIES

THESE METHODS CONTAIN IMPLICIT PRIORS

But they still work well - sometimes

Analogous to Bell & de Jong (2001) M/L as a function of colour method.

Problems:

INFERRING IONISED GAS PROPERTIES

THESE METHODS CONTAIN IMPLICIT PRIORS

But they still work well - sometimes

Analogous to Bell & de Jong (2001) M/L as a function of colour method.

Problems:

- ✓ Uncertainty analysis based on fitting formulae is questionable.
  - ➡ Full model grids should be used.

INFERRING IONISED GAS PROPERTIES

THESE METHODS CONTAIN IMPLICIT PRIORS

But they still work well - sometimes

Analogous to Bell & de Jong (2001) M/L as a function of colour method.

Problems:

- ✓ Uncertainty analysis based on fitting formulae is questionable.
  - ➡ Full model grids should be used.
- ✓ Unspecified prior assumptions might be invalid in some conditions, but so hidden as to be overlooked.
  - ➡ Make your assumptions explicit - for instance using a Bayesian analysis framework.

INFERRING IONISED GAS PROPERTIES

THESE METHODS CONTAIN IMPLICIT PRIORS

But they still work well - sometimes

Analogous to Bell & de Jong (2001) M/L as a function of colour method.

Problems:

✓ Uncertainty analysis based on fitting formulae is questionable.

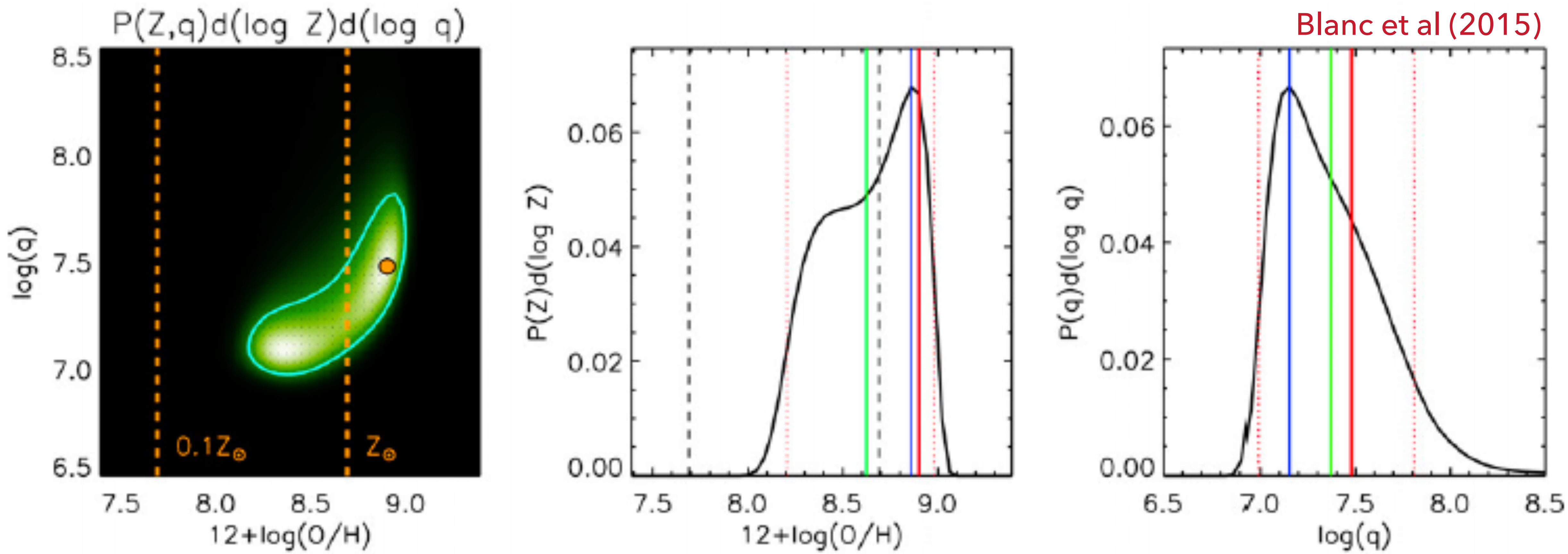
➡ Full model grids should be used.

✓ Unspecified prior assumptions might be invalid in some conditions, but so hidden as to be overlooked.

➡ Make your assumptions explicit - for instance using a Bayesian analysis framework.

E.g.: pyqz (Dopita et al 2013), IZI (Blanc et al (2015), HII-CHI-mistry (Perez-Montero 2014), MPA-JHU code (Brinchmann et al 2004; Tremonti et al 2004).

EXAMPLE: IZI – SINGLE HII REGION



Note in particular the **double-peaked** nature of the PDFs & the **correlation** between ionisation parameter (q) and metallicity (12 + log O/H).

HANDLING OF PDFS

PDFs for e.g. O/H are often bimodal/complex in shape - this requires some care:

Best(?): Use the full PDF

HANDLING OF PDFS

PDFs for e.g. O/H are often bimodal/complex in shape - this requires some care:

Best(?): Use the full PDF

Alternative: summary statistics

HANDLING OF PDFS

PDFs for e.g. O/H are often bimodal/complex in shape - this requires some care:

Best(?): Use the full PDF

Alternative: summary statistics

Bimodality:

$$\beta = \frac{\mathrm{skewness}^2 + 1}{\mathrm{kurtosis}}$$

Entropy:

$$S(P) = \int P(x) \ln P(x) \; dx$$

Quantiles:

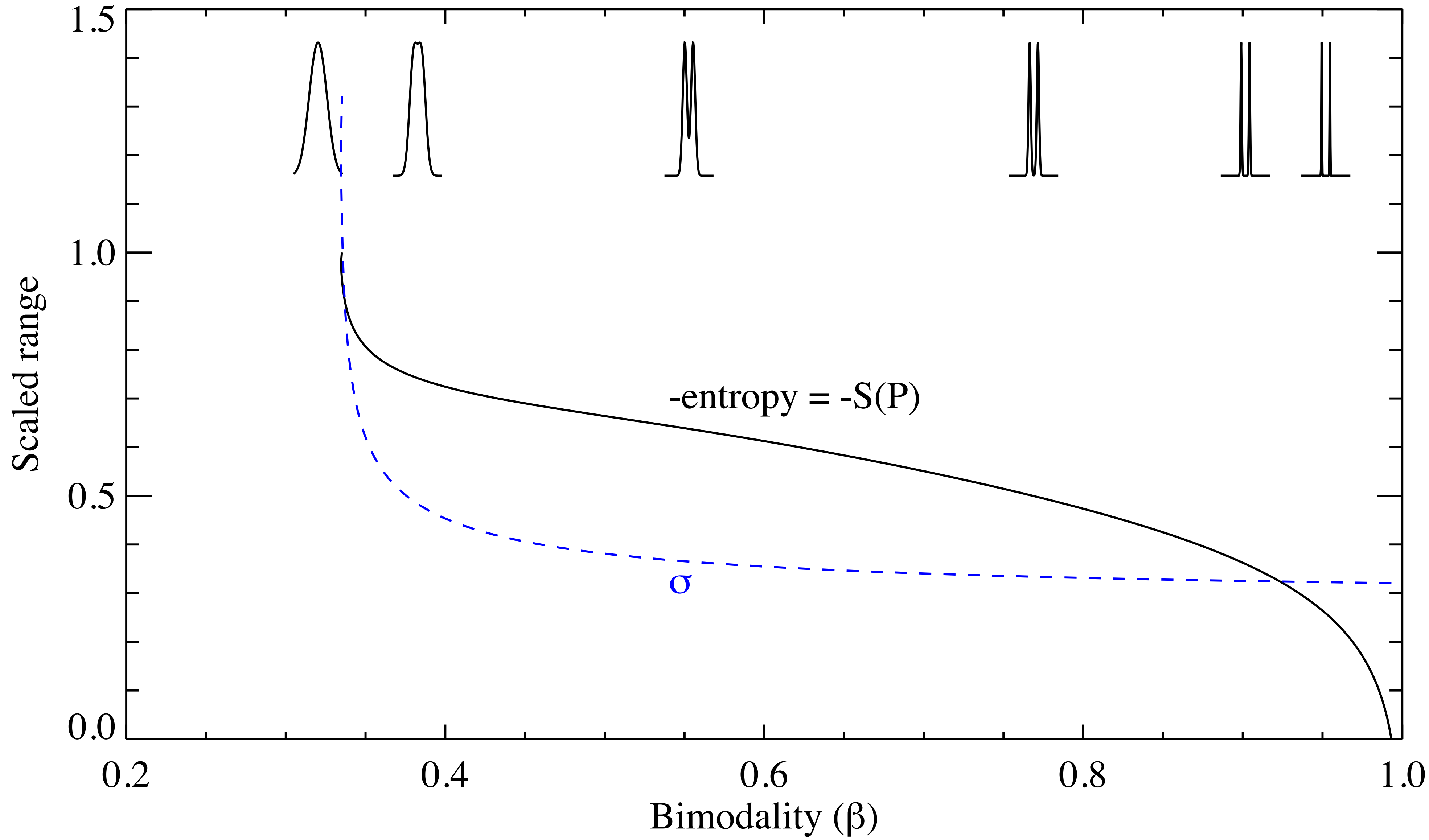
$$\left\{ x_q : q = \int^{x_q} P(t) dt \right\}$$

HANDLING OF PDFS

PDFs for e.g. O/H are often bimodal/complex in shape - this requires some care:

$$\beta = \frac{\mathrm{skewness}^2 + 1}{\mathrm{kurtosis}}$$

$$S(P) = \int P(x) \ln P(x)\, dx$$



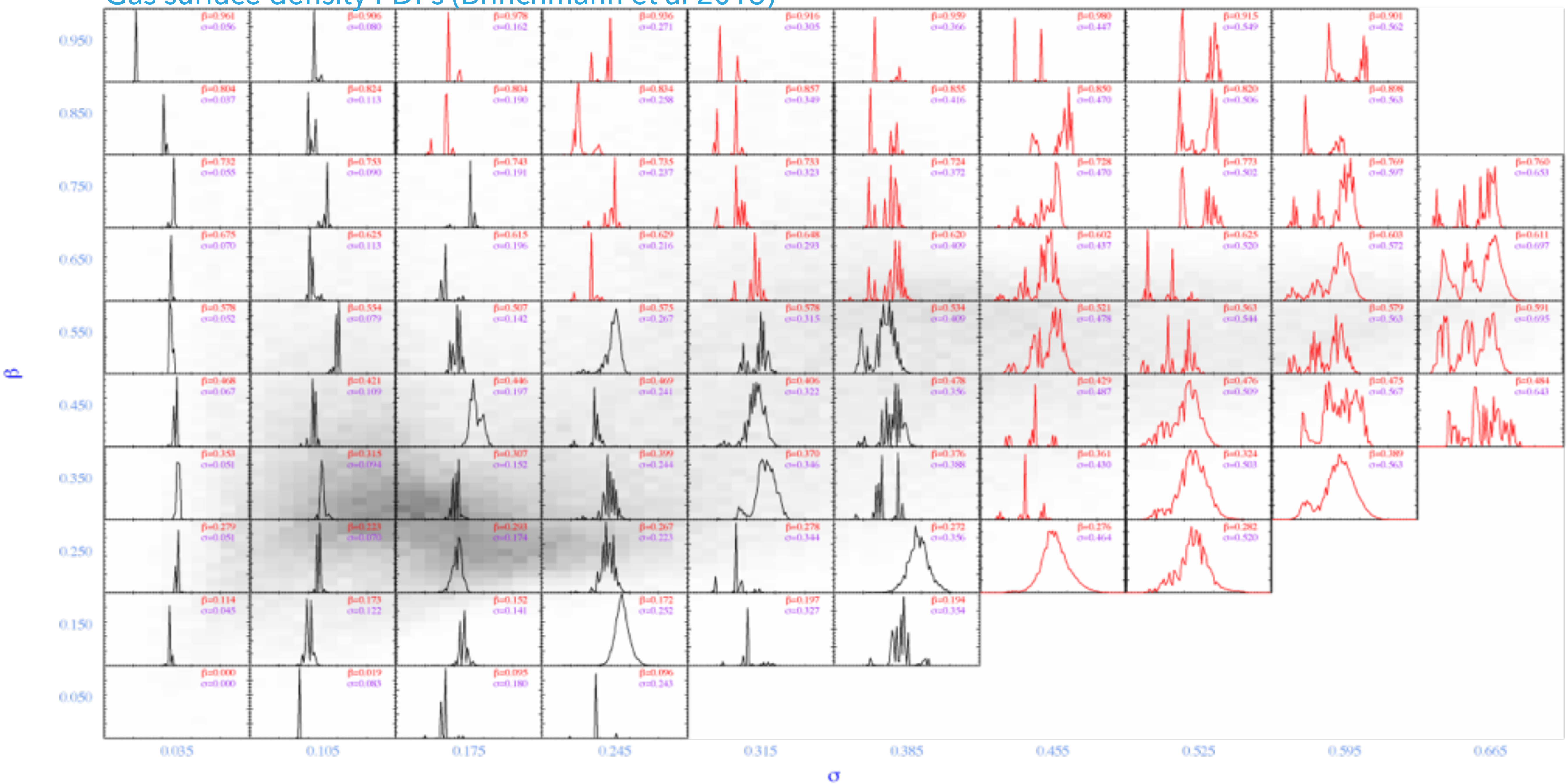
HANDLING OF PDFS

PDFs for e.g. O/H are often bimodal/complex in shape - this requires some care:

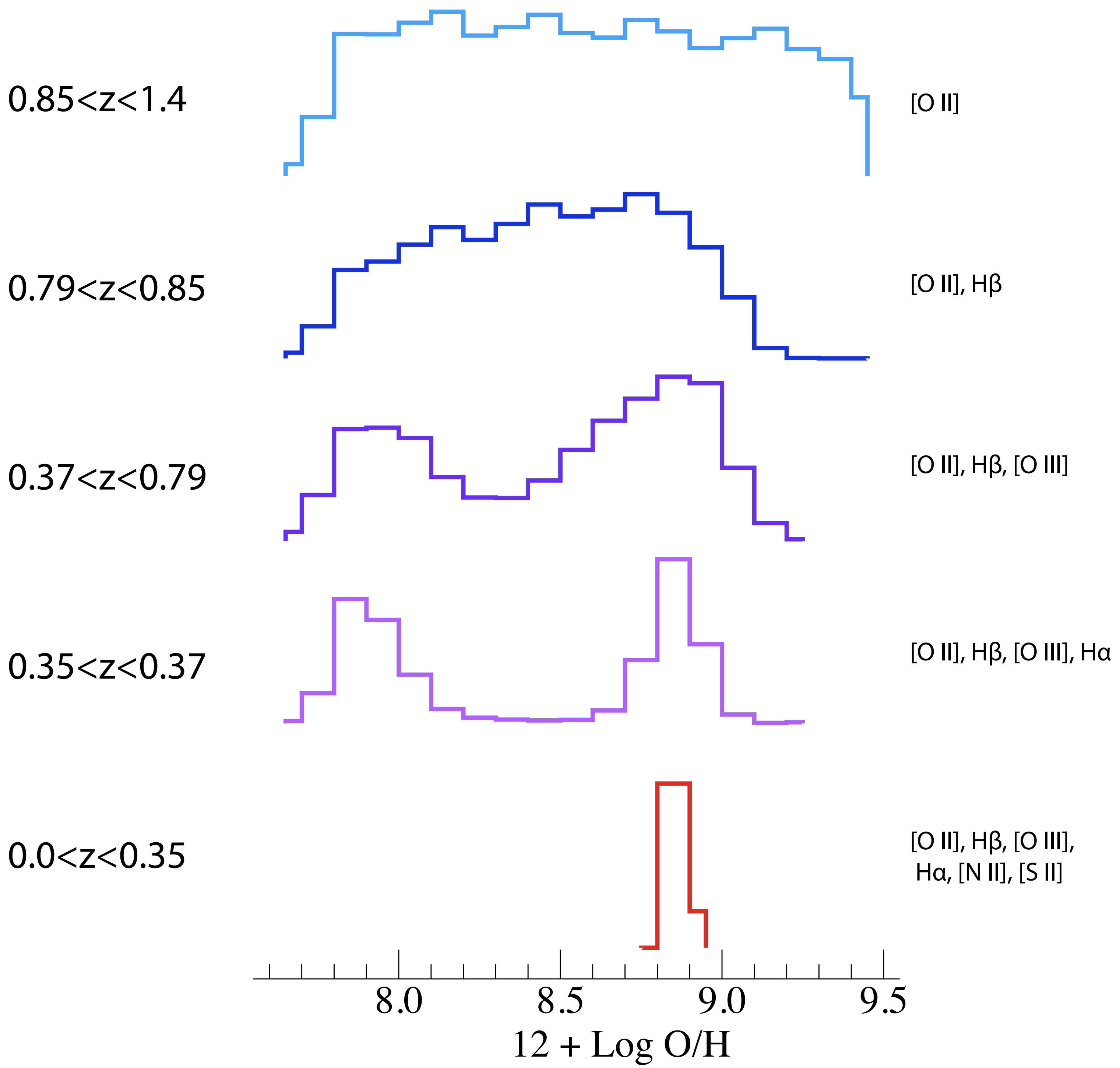
$$\beta = \frac{\text{skewness}^2 + 1}{\text{kurtosis}}$$

$$S(P) = \int P(x) \ln P(x)\, dx$$

Gas surface density PDFs (Brinchmann et al 2013)



SO IS A BAYESIAN ANALYSIS THE WAY FORWARD?

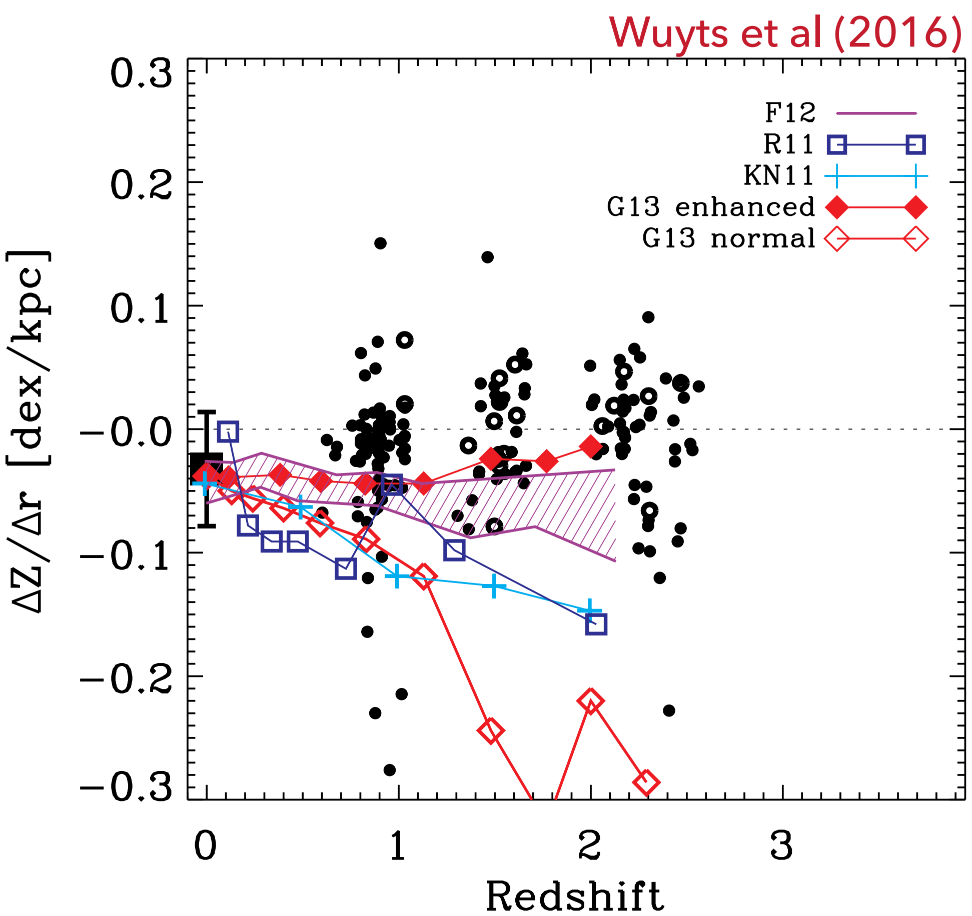


It might be - but priors might have to be carefully chosen to get useful results.

True when comparing different redshifts, but also when comparing *within* a galaxy

Optical spectrograph: [3700Å, 9000Å]

METALLICITY GRADIENTS DETERMINATION



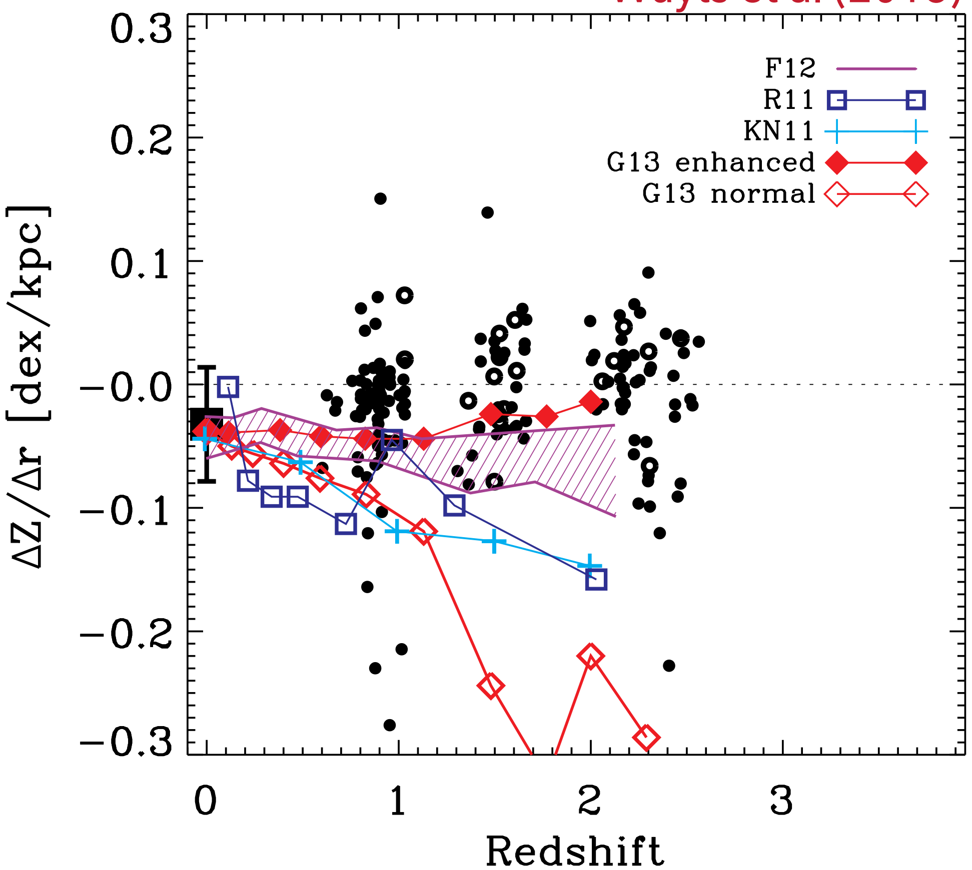
One approach (e.g. KMOS-3D, Wuyts et al 2016):

- strong lines close in wavelength
- beam smearing correction from models

Works well - but hard to quantify uncertainties in a rigorous way.

METALLICITY GRADIENTS DETERMINATION

Wuyts et al (2016)

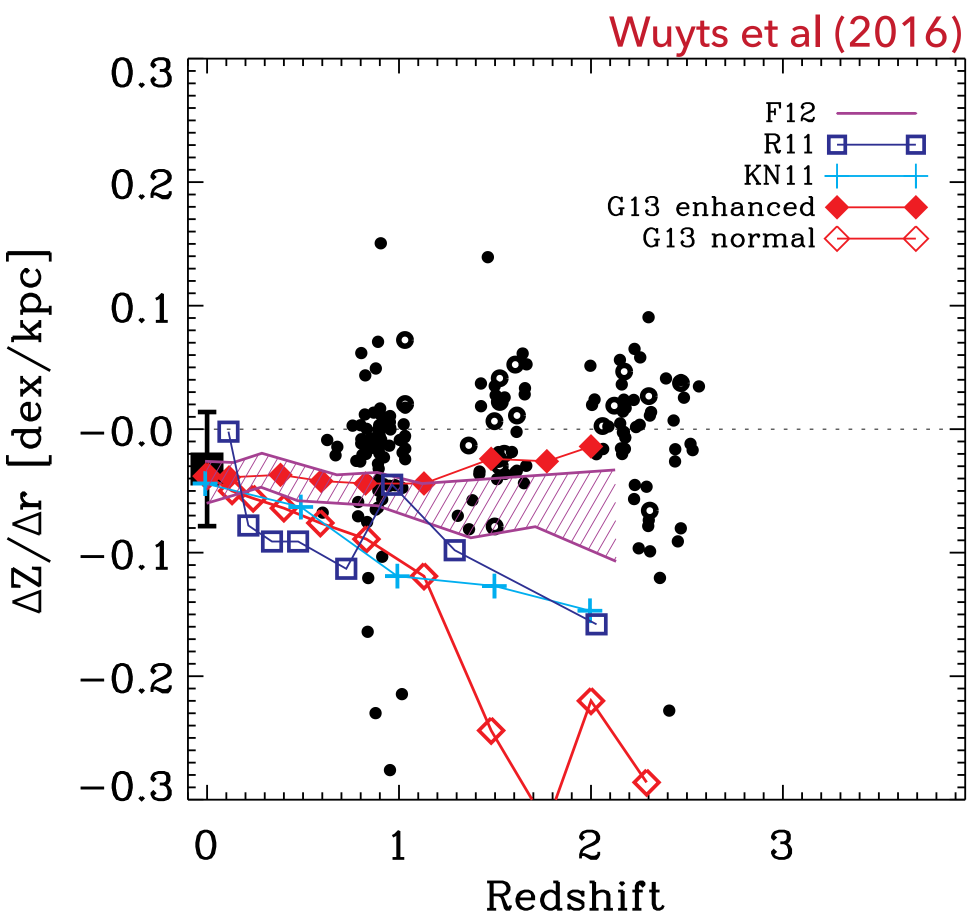


One approach (e.g. KMOS-3D, Wuyts et al 2016):

- strong lines close in wavelength
- beam smearing correction from models

Works well - but hard to quantify uncertainties in a rigorous way.

METALLICITY GRADIENTS DETERMINATION



One approach (e.g. KMOS-3D, Wuyts et al 2016):

- strong lines close in wavelength
- beam smearing correction from models

Works well - but hard to quantify uncertainties in a rigorous way.

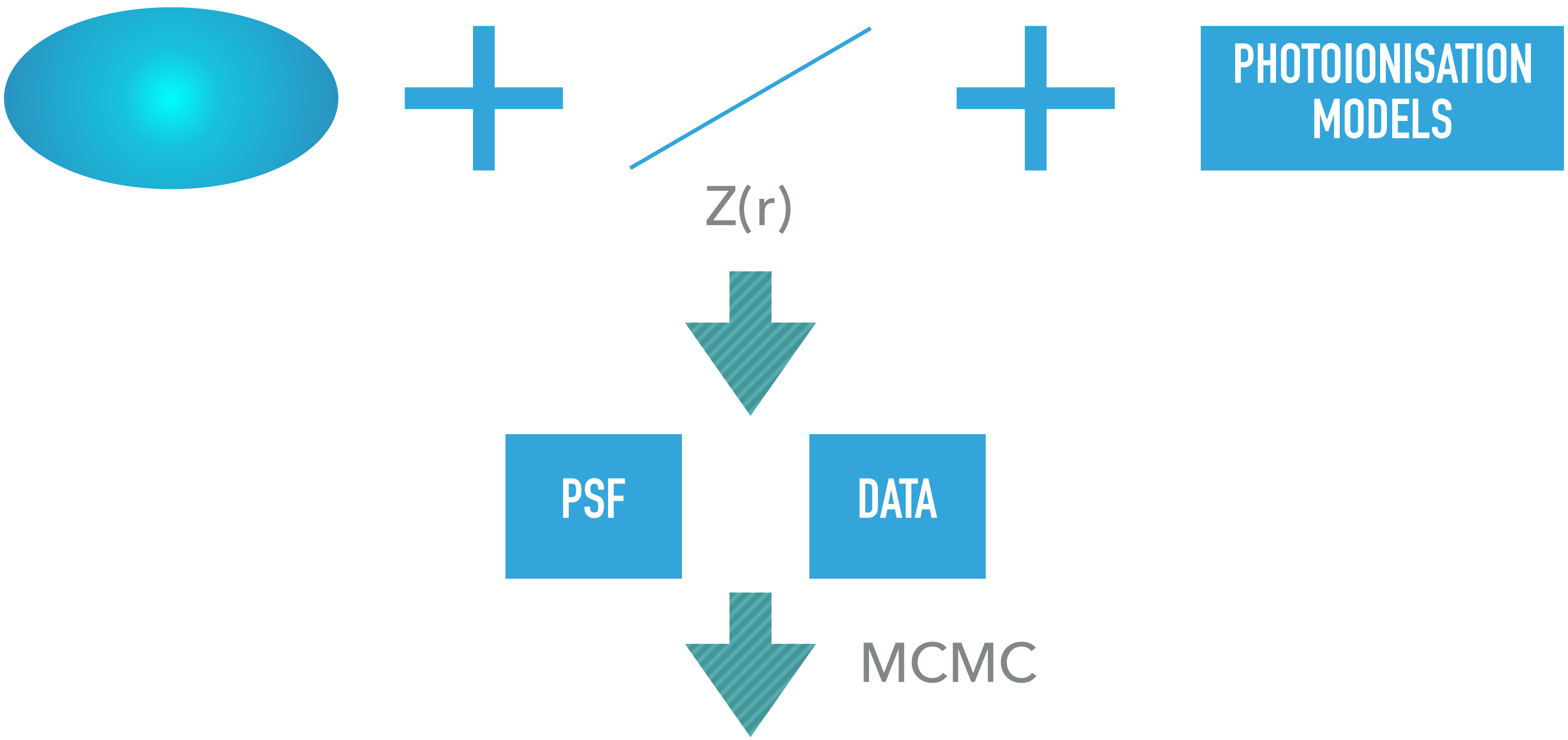
With MUSE:

- a range of lines, weaker in the outskirts
- different sets of lines at different redshifts
- wavelength dependent PSF

METALLICITY GRADIENTS DETERMINATION

David Carton

IFU data: need to exploit spatial correlations



Metallicity gradient, central metallicity

METALLICITY GRADIENTS DETERMINATION

IFU data: need to exploit spatial correlations

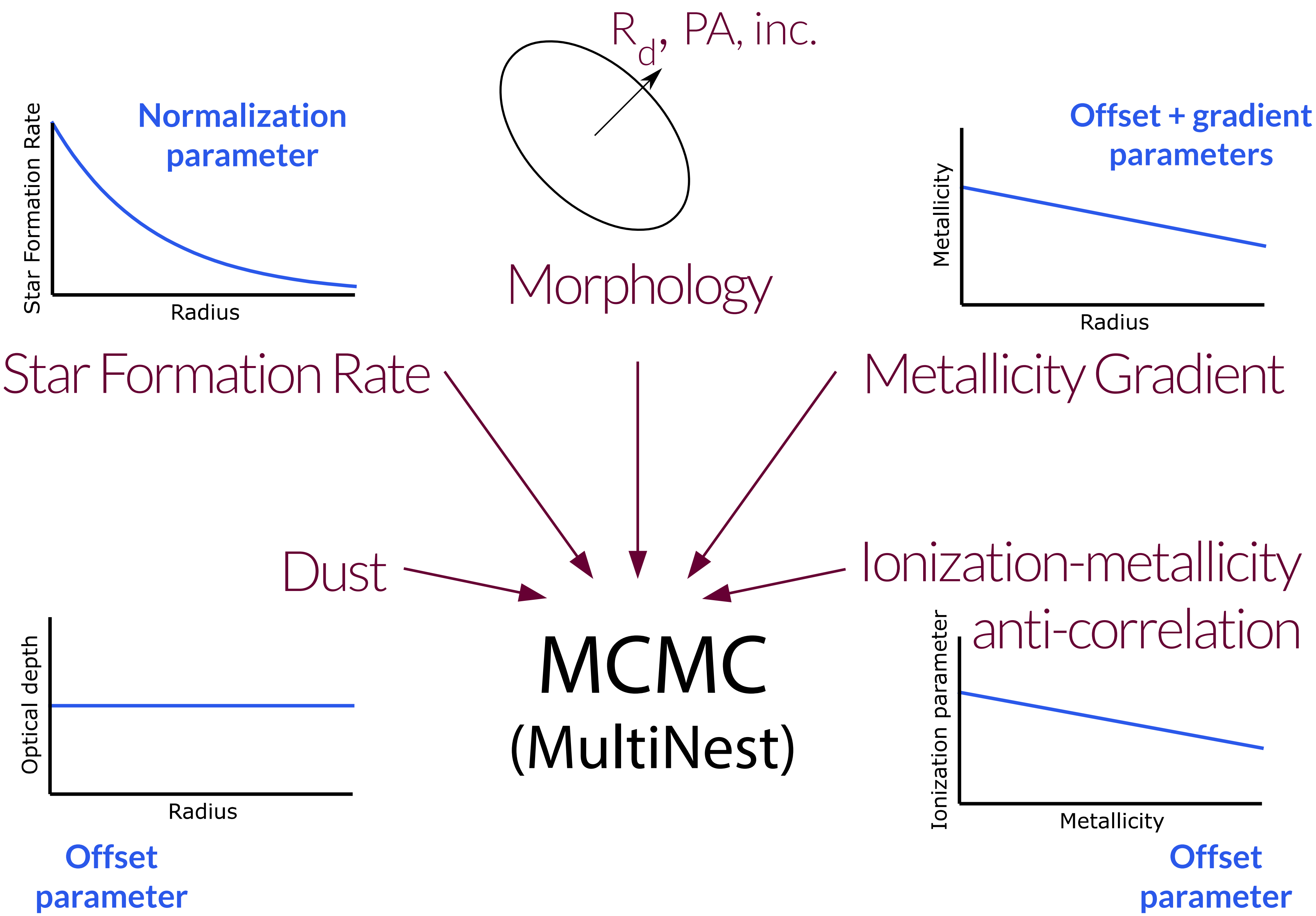
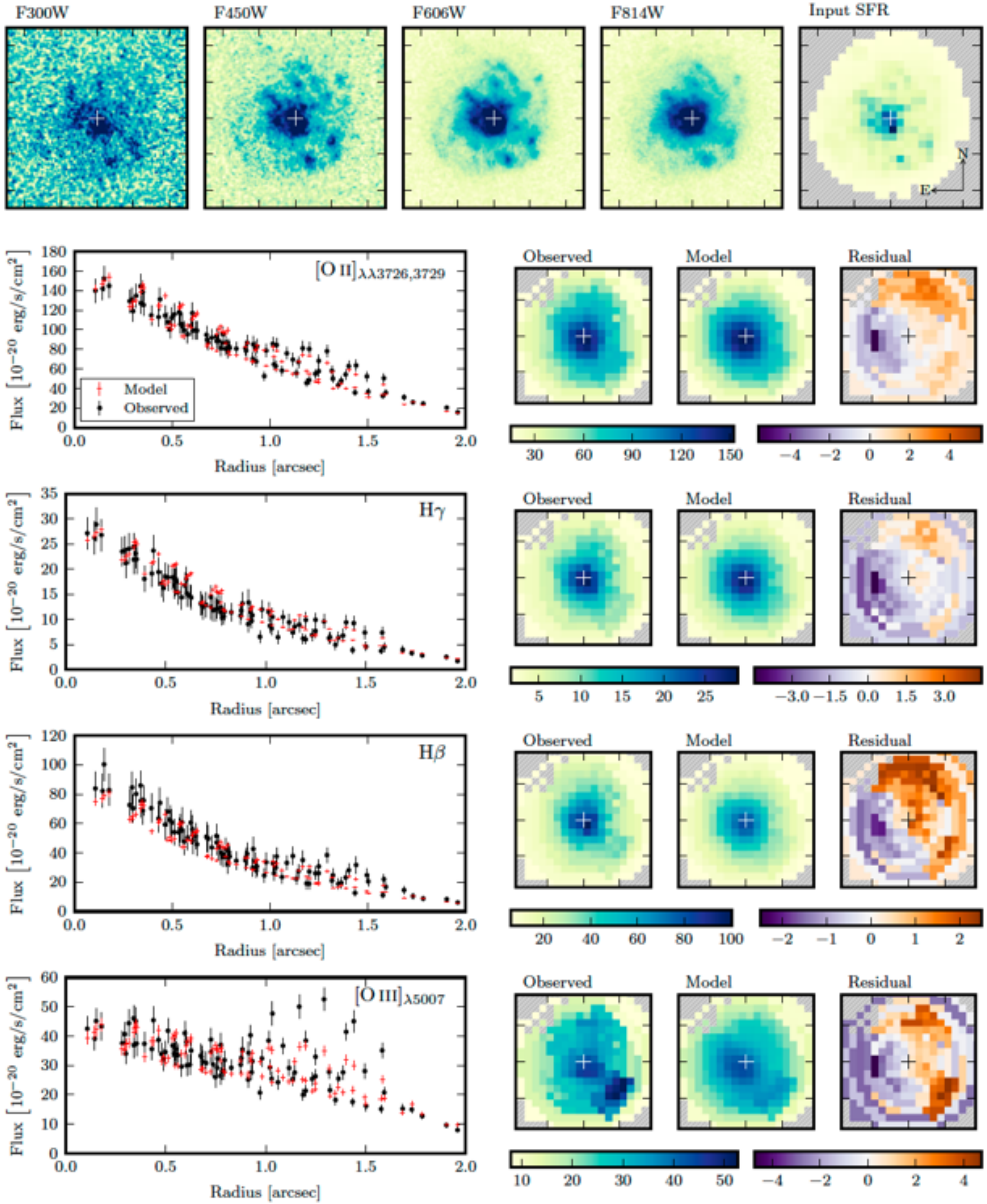
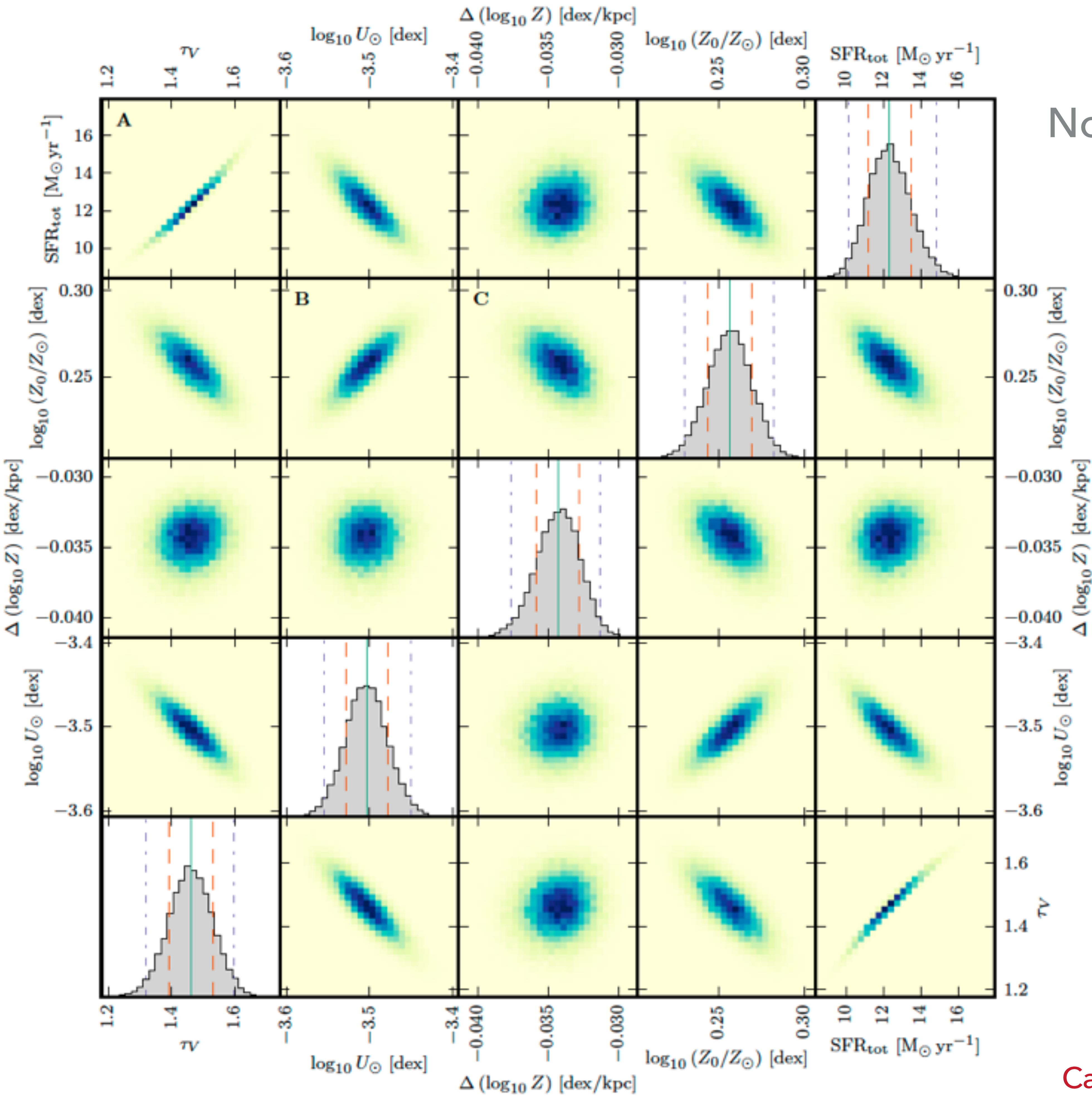


Figure from David Carton



Carton et al (2016, submitted)

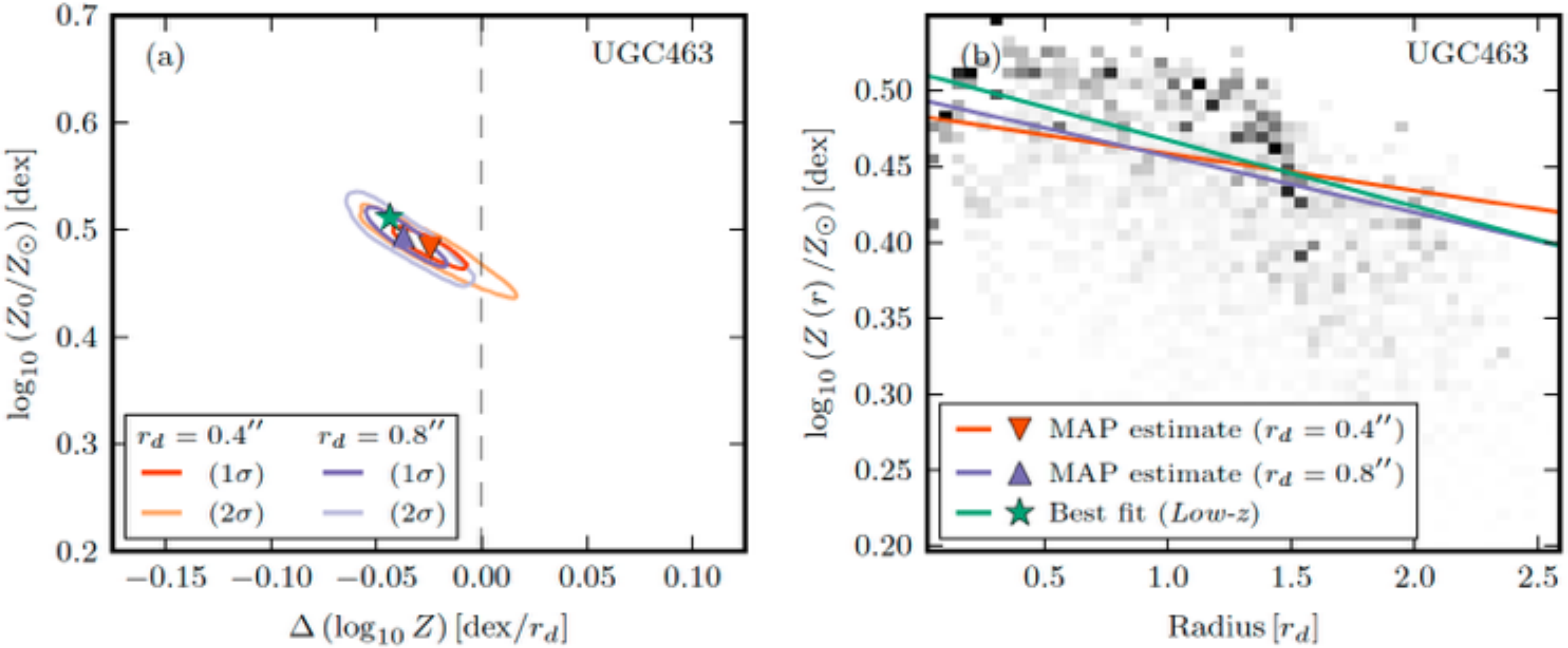
Not always this nice!

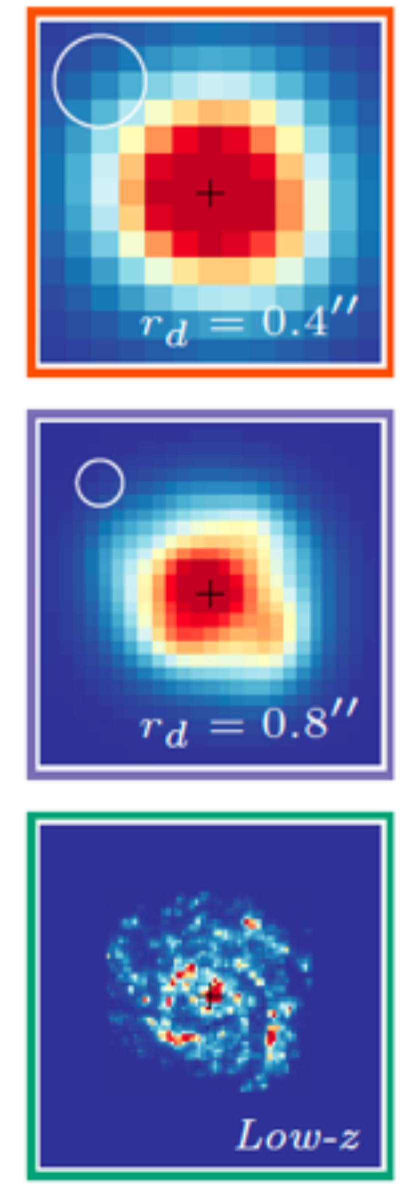


Carton et al (2016, submitted)

When galaxies are well-behaved it works

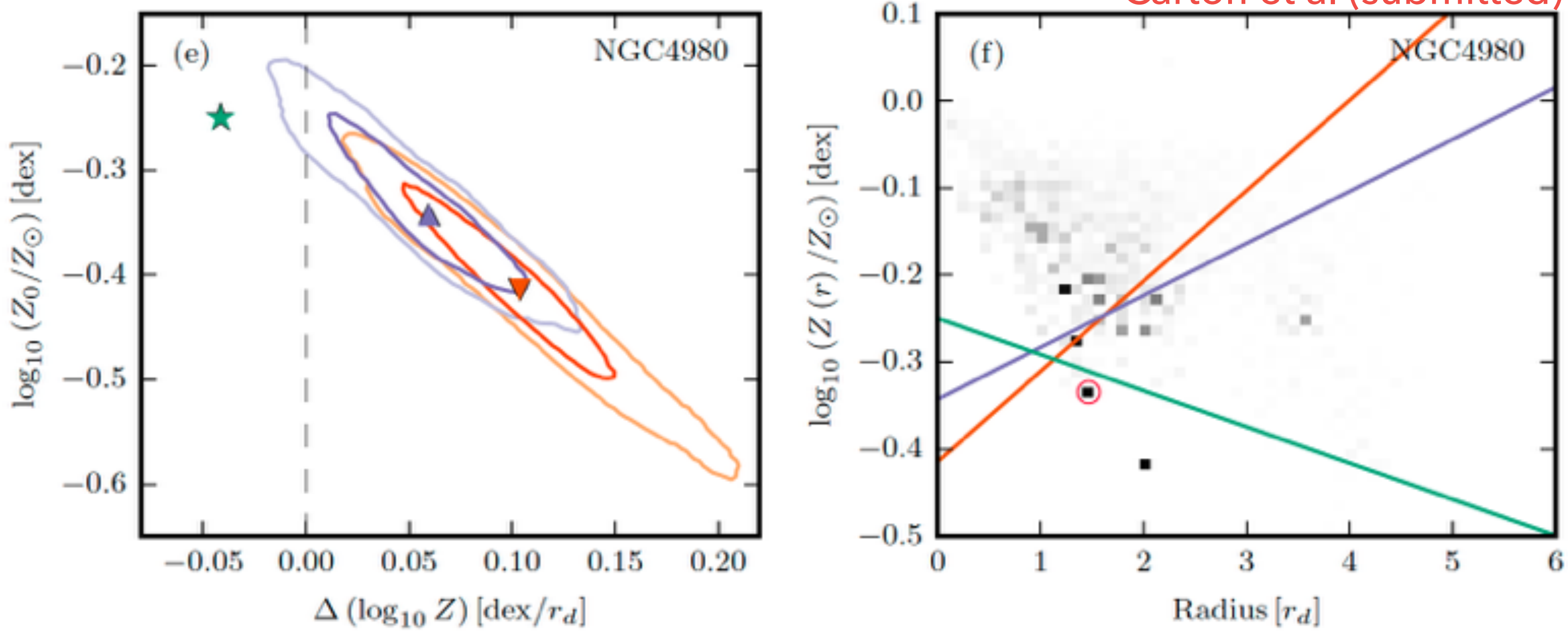
Carton et al (submitted)

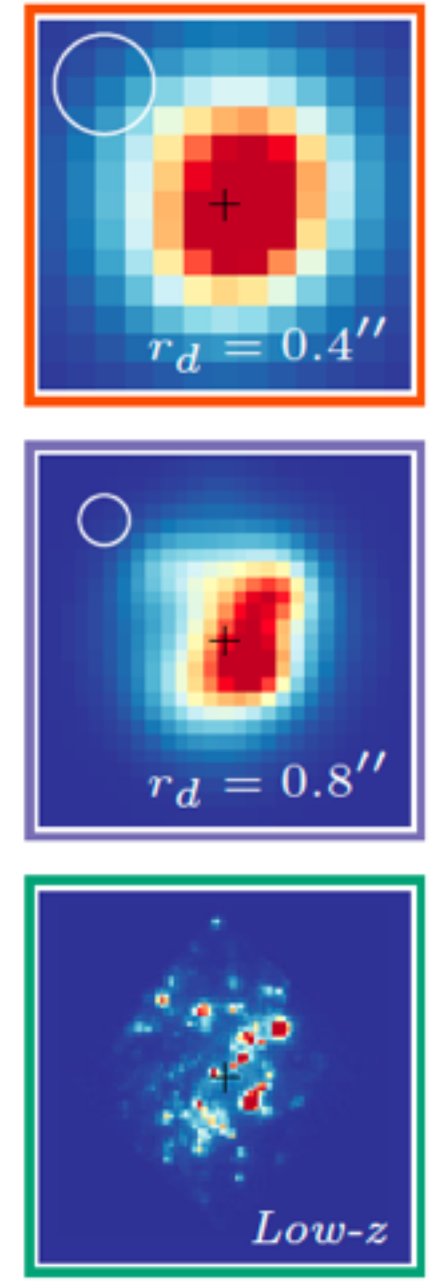




When they are not - it is a problem

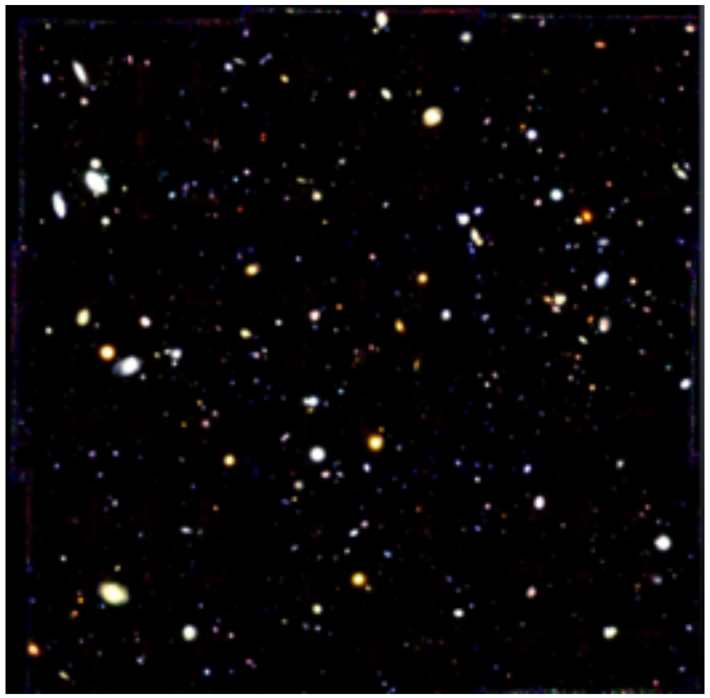
Carton et al (submitted)



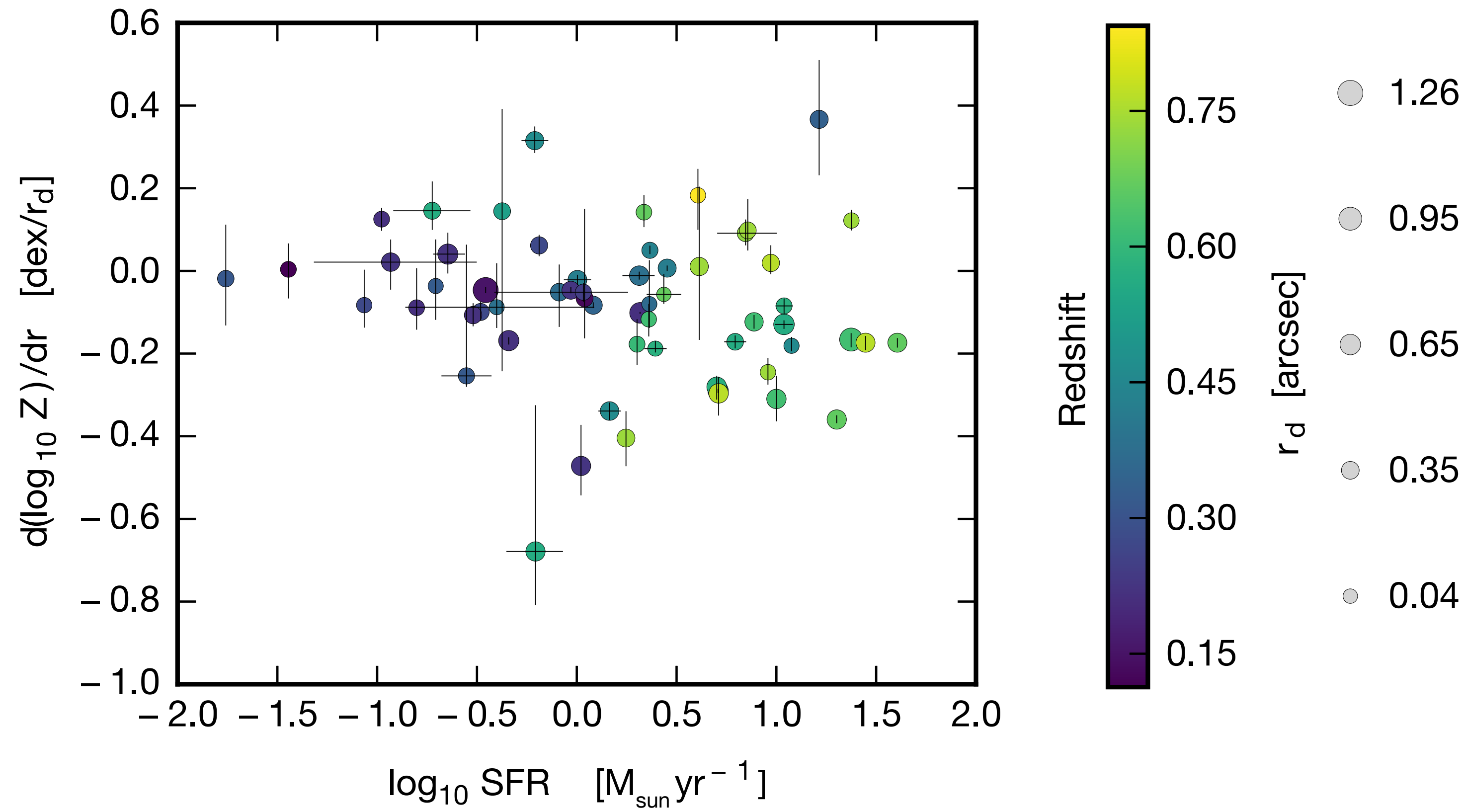


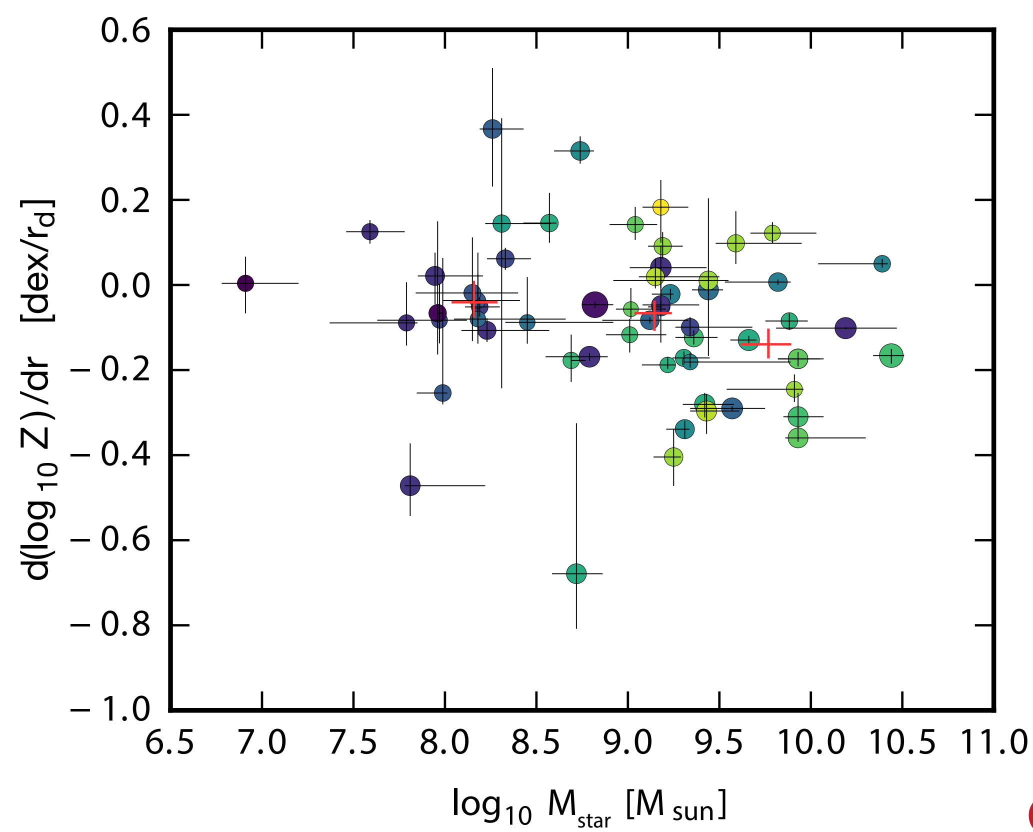
Clumpy structure means comparison of low & high-z metallicity profiles can be problematic.

APPLICATION: MUSE DEEP FIELDS



3′x3′ mosaic in the UDF





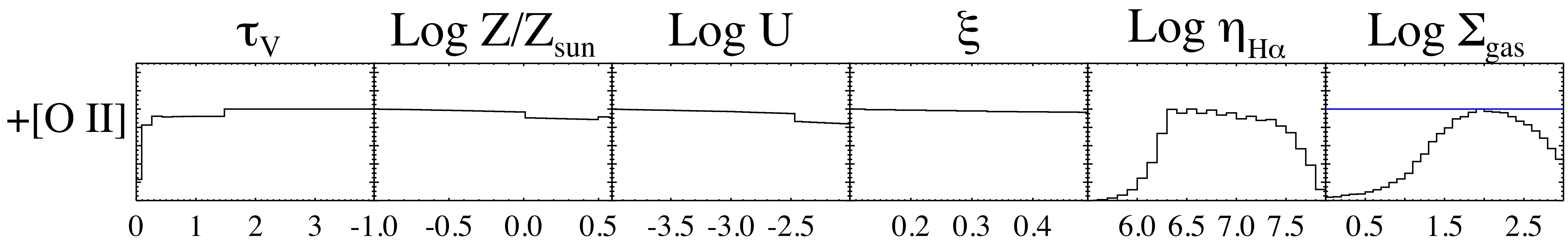
Carton et al (2016, in prep)

SUMMARY

- The basic ingredients are increasingly well known - time for the statistical methodology to take advantage of this.
- Bayesian modelling, particularly of emission lines, should become the norm for the field over the coming years.
- With the steady increase in resolved spectroscopy, flexible ways of integrating spatial coherence would be useful extensions to SED/emission line models.
- A full forward modelling of metallicity profiles allows uncertainties to be well characterised but also shows that low-z and high-z metallicity gradients should be compared with care

ADDITIONAL MATERIAL

FITTING EMISSION LINES - A LOT TO LEARN



Fit models (Charlot & Longhetti 2001, also Dopita et al (2013) for metallicities):

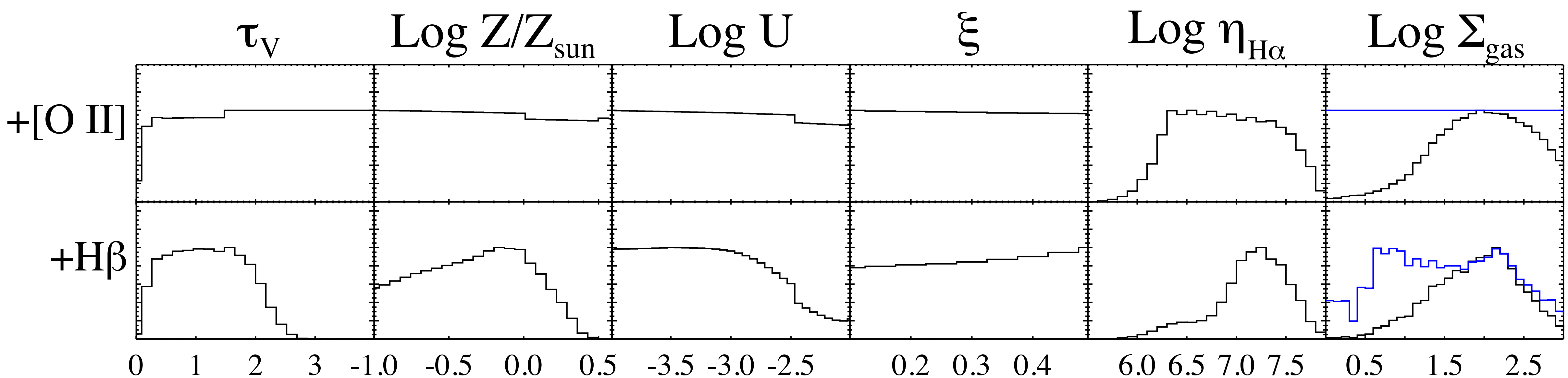
$$\ln P(\mathcal{M}|\{L_i\}) = -\frac{1}{2}\sum_{i\in\{L_i\}}\frac{(f_i - Af_{\mathcal{M}})^2}{\sigma_i^2} + \ln \mathrm{Pr},$$

f_M=model

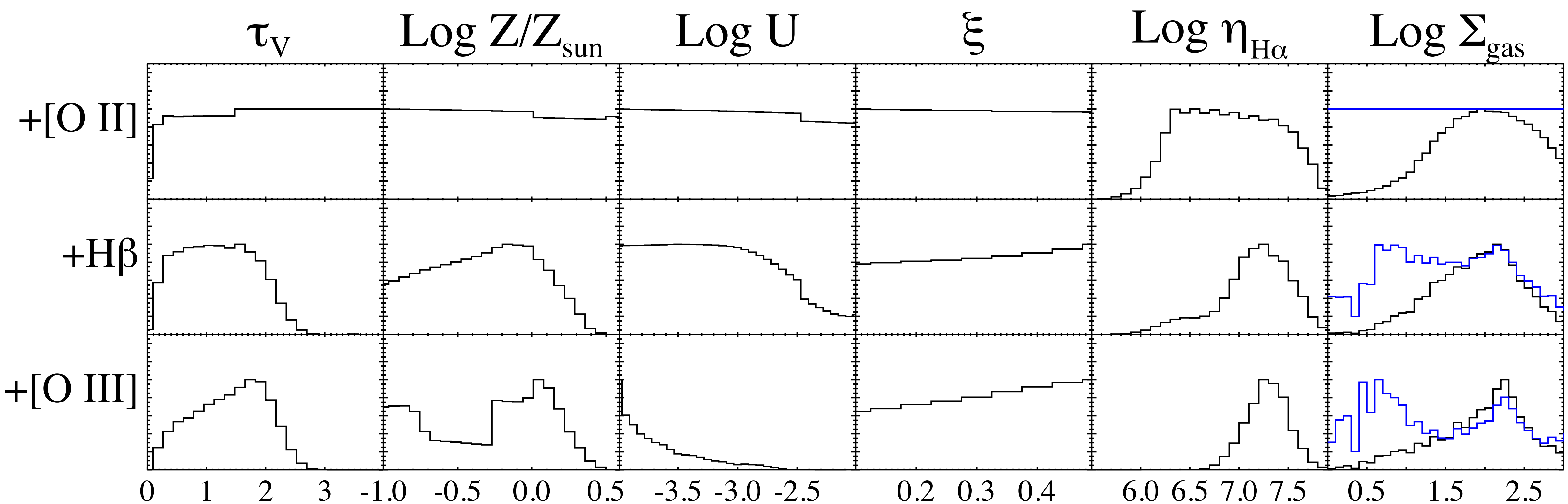
Prior

likelihood

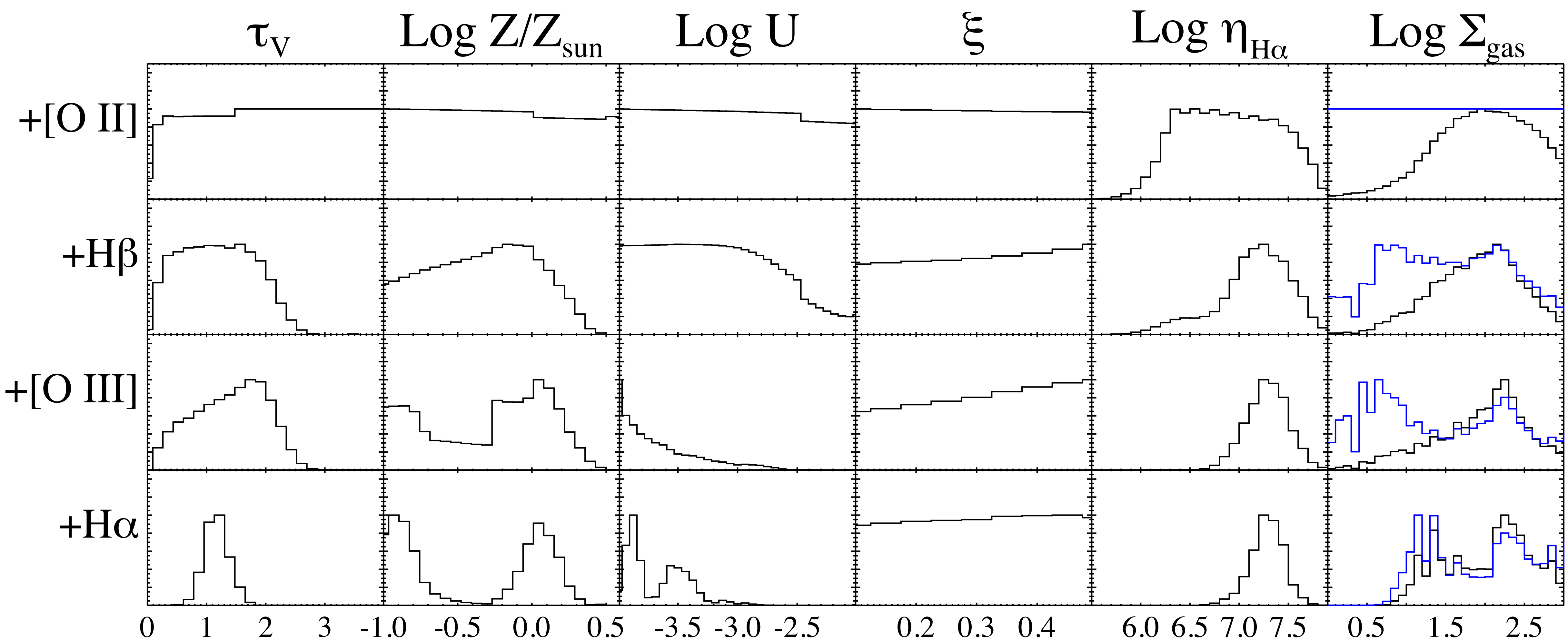
FITTING EMISSION LINES – A LOT TO LEARN



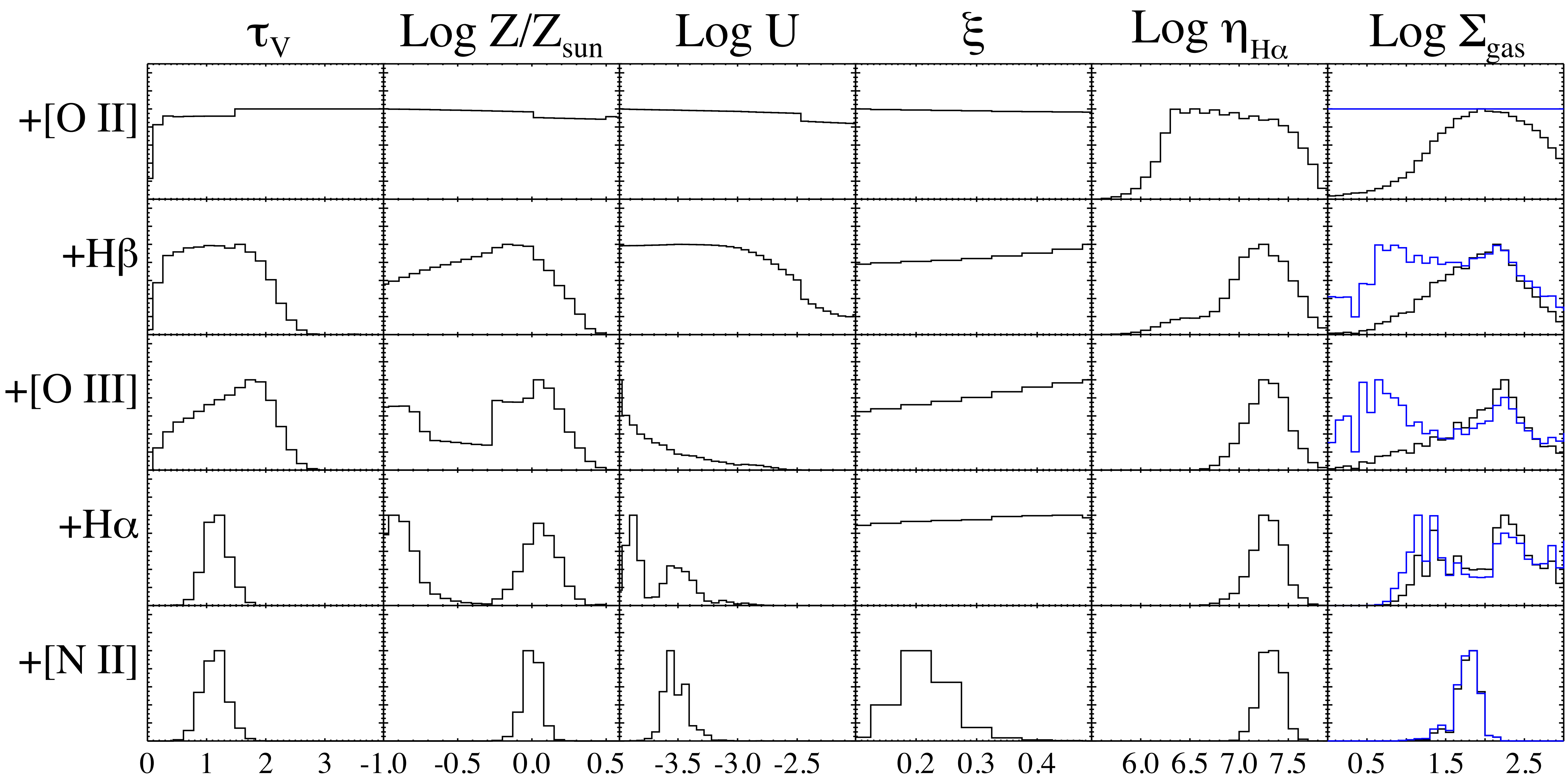
FITTING EMISSION LINES – A LOT TO LEARN



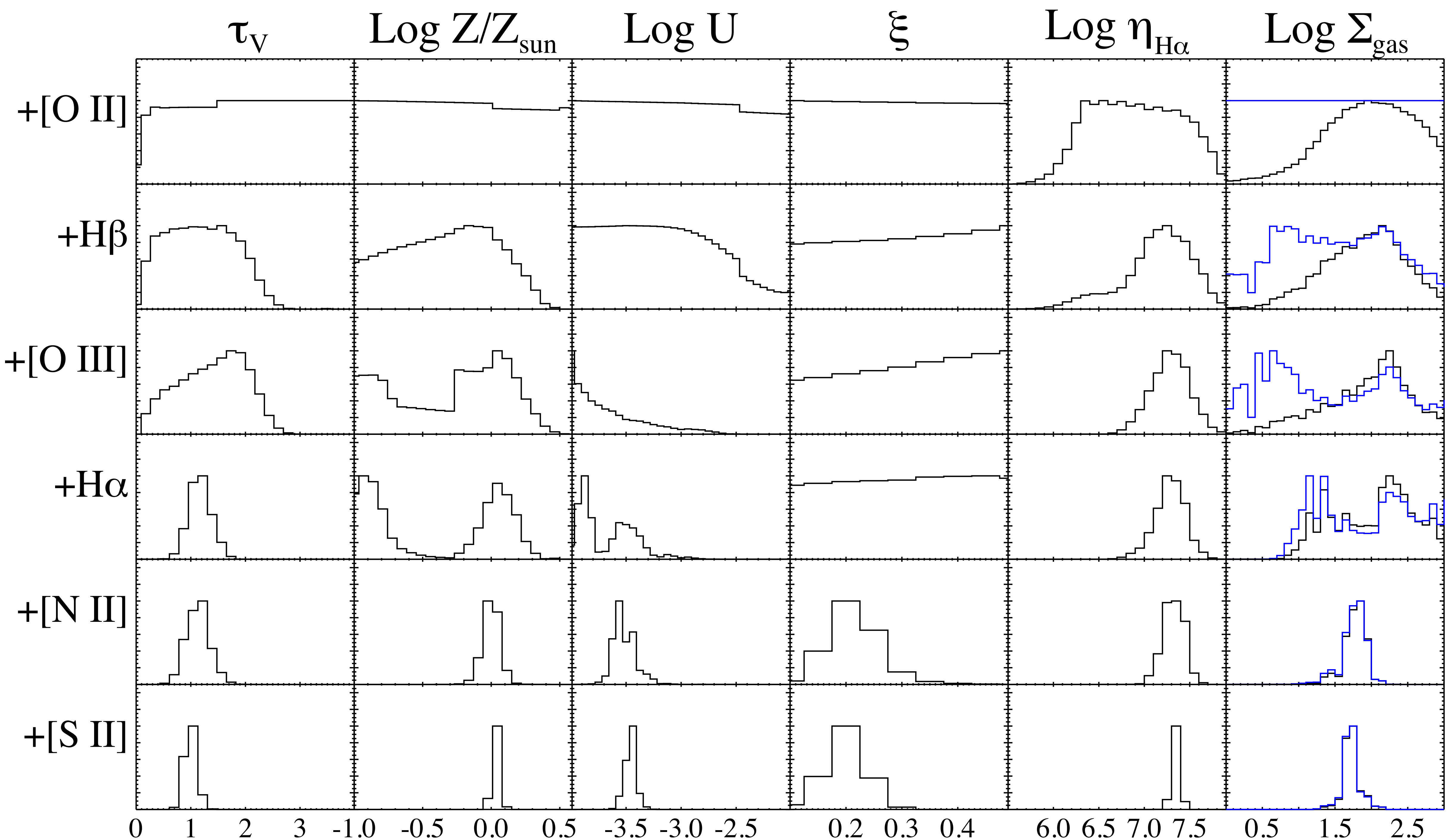
FITTING EMISSION LINES – A LOT TO LEARN



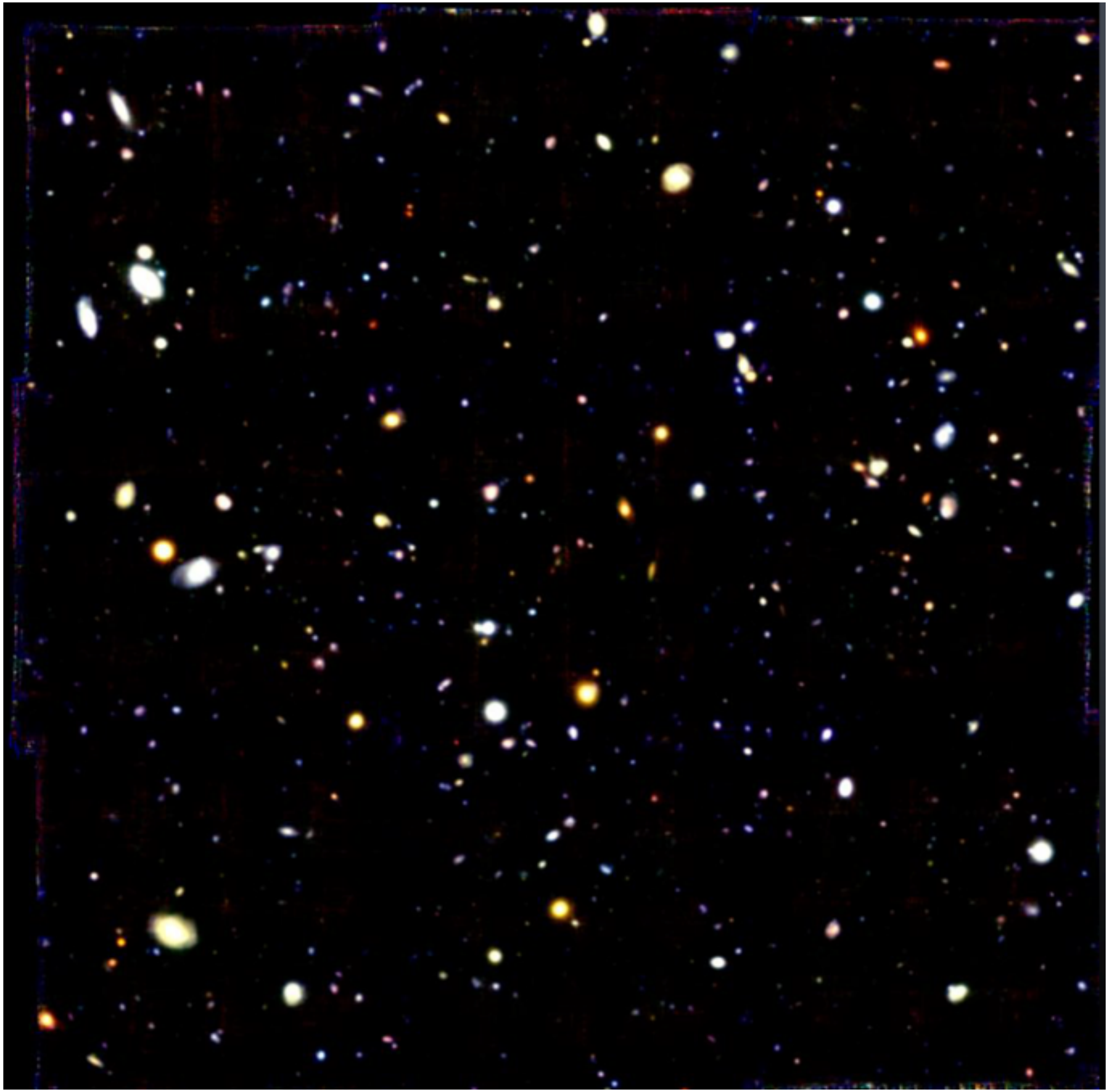
FITTING EMISSION LINES – A LOT TO LEARN



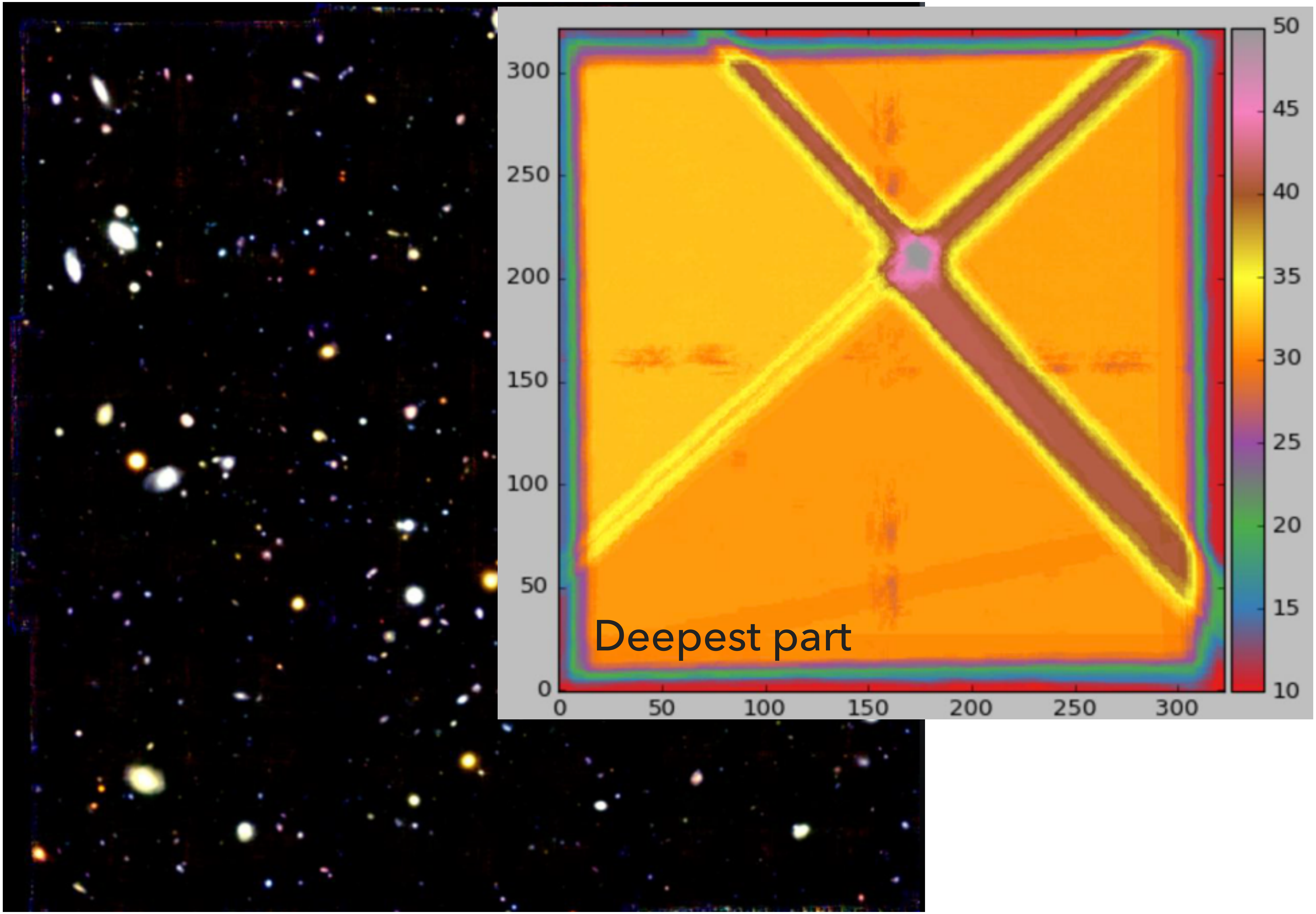
FITTING EMISSION LINES – A LOT TO LEARN



UDF Mosaic : 3′x3′ - >10hr per pointing



UDF Mosaic : 3′x3′ - >10hr per pointing

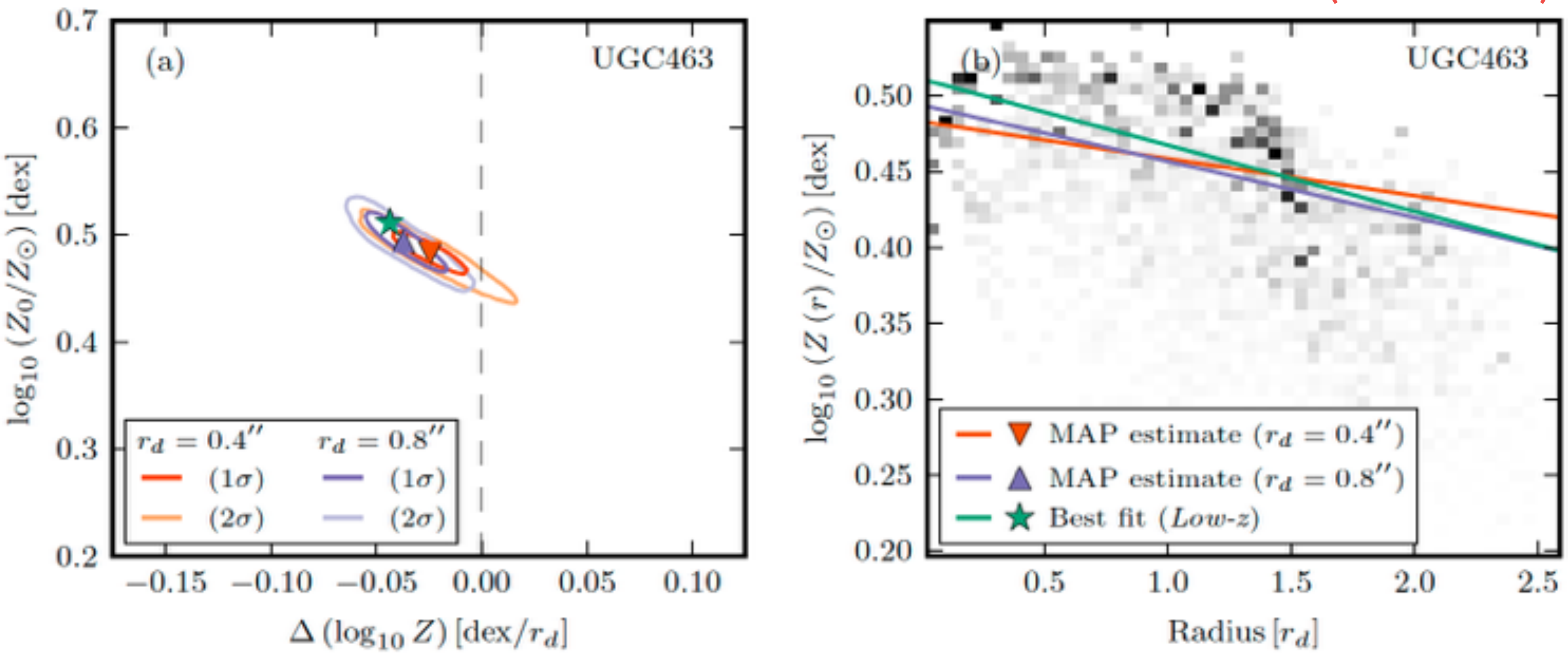


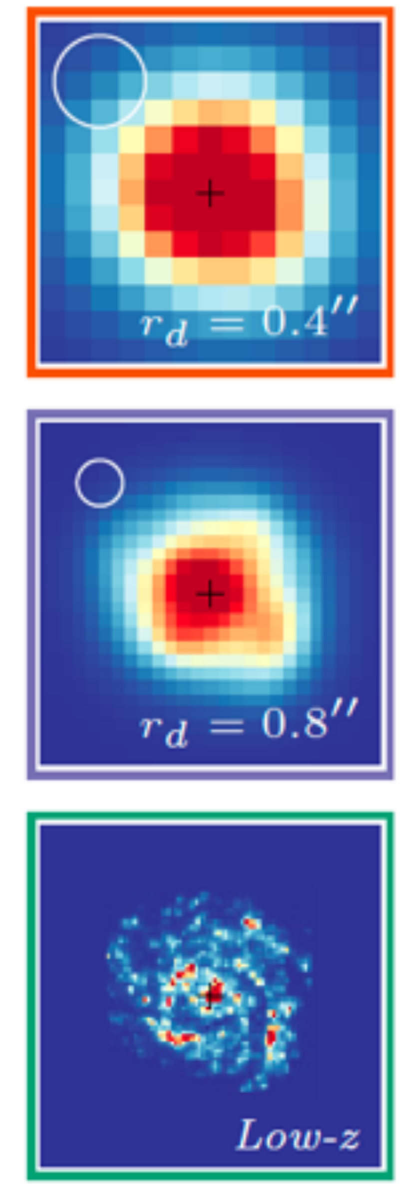
THE ENEMY: SEEING

Way forwards: Forward modelling

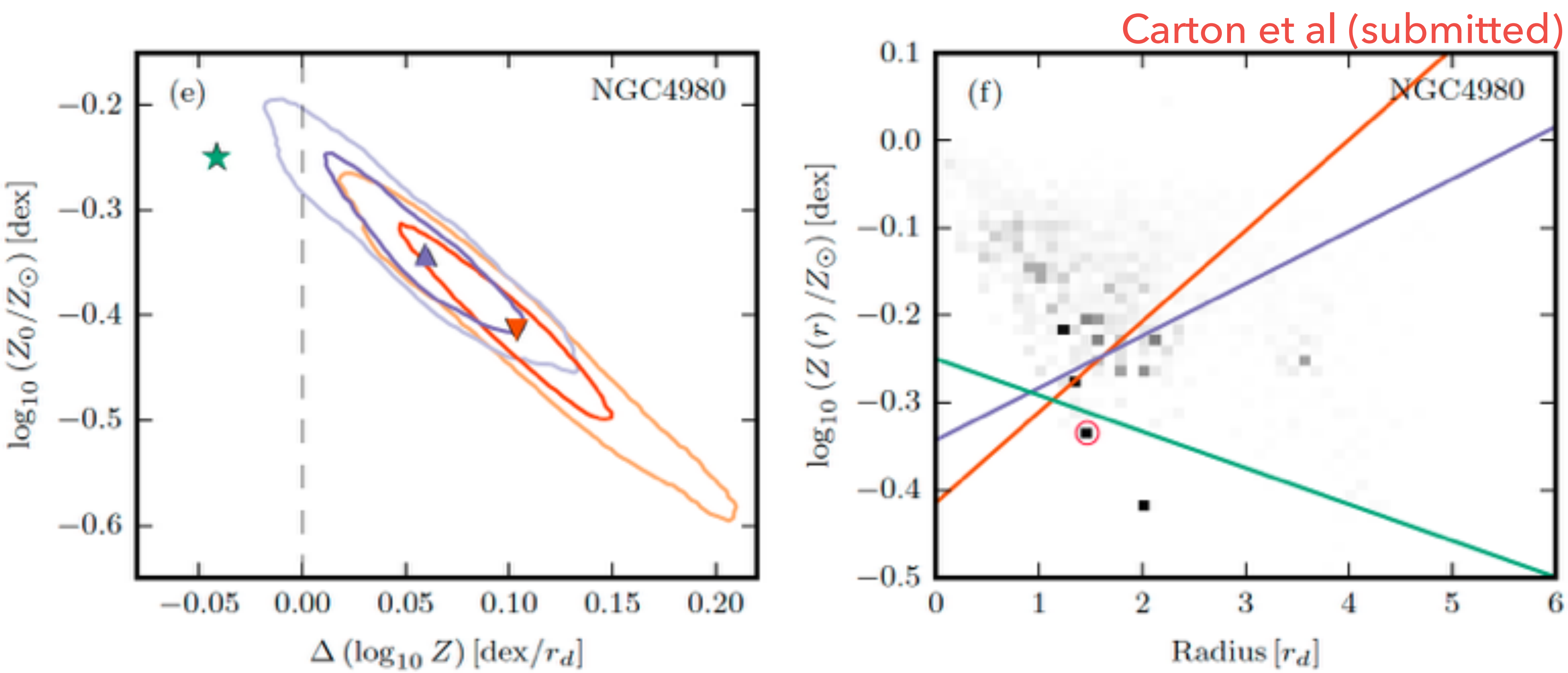
When galaxies are well-behaved it works

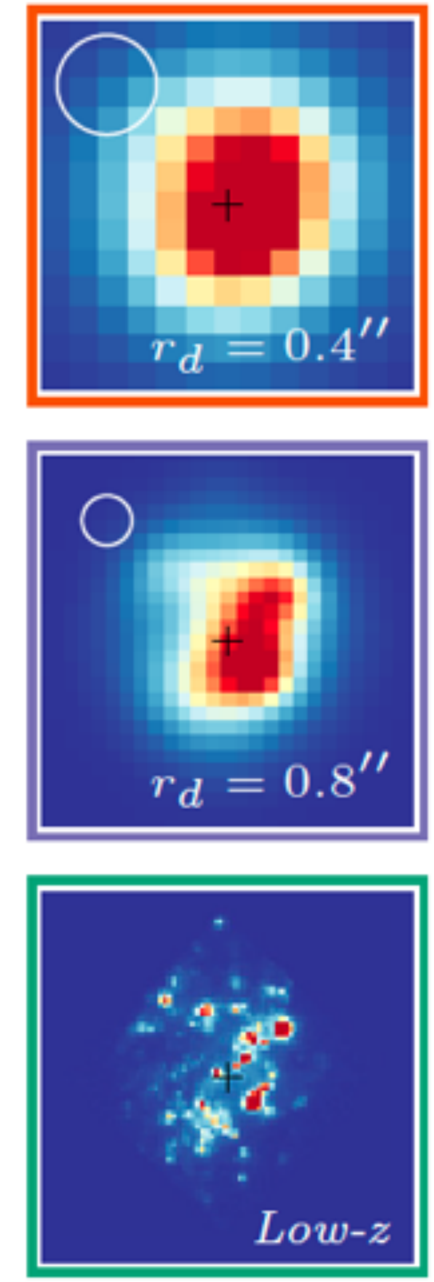
Carton et al (submitted)





When they are not - it is a problem





Clumpy structure means comparison of low & high-z metallicity profiles can be problematic.

METALLICITY PROFILES