Some applications of Bayesian methods in galaxy evolutionary studies

Stefano Andreon
INAF-Oss. Di Brera
stefano.andreon@brera.inaf.it

Most measurements that I perform deal with:

1) Measurements near boundaries (fractions, completeness, hardness ratios, sources/features with few counts …)

-> accounting for boundaries

2) nuisance parameters

I.e. I would like to measure an interesting parameter, without precise knowledge of another parameter that I known to influence the measurement:

-> parameter estimation in presence of a nuisance parameters

3) ask whether my models or hypothesis have to be rejected

e.g.

i) do my data provide evidence for a luminosity evolution ?

ii) What fits better: a de Vaucouleurs or an exponential law?

$\ln(I(R))=R^{1/n}$ with N=4 (de Vaucouleurs) or N=1 (exponential)

->model selection

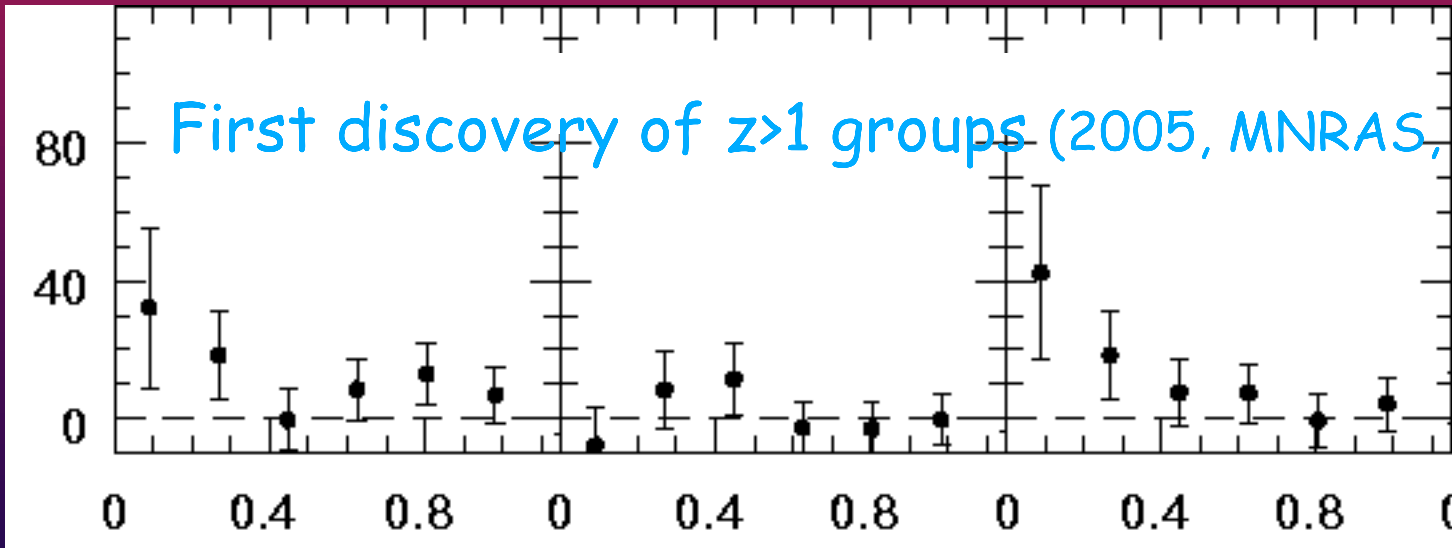
models are often not hierarchically nested.

Bayesian methods address all three above topics.

Galaxies come in positive units. Maybe.

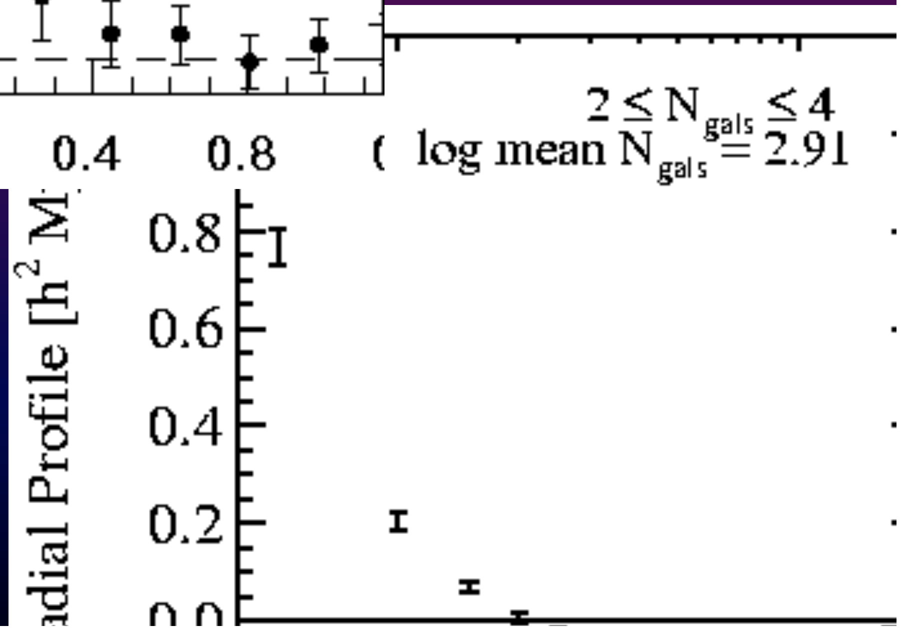
First discovery of z>1 groups (2005, MNRAS, 357, 1357)

N



Radius

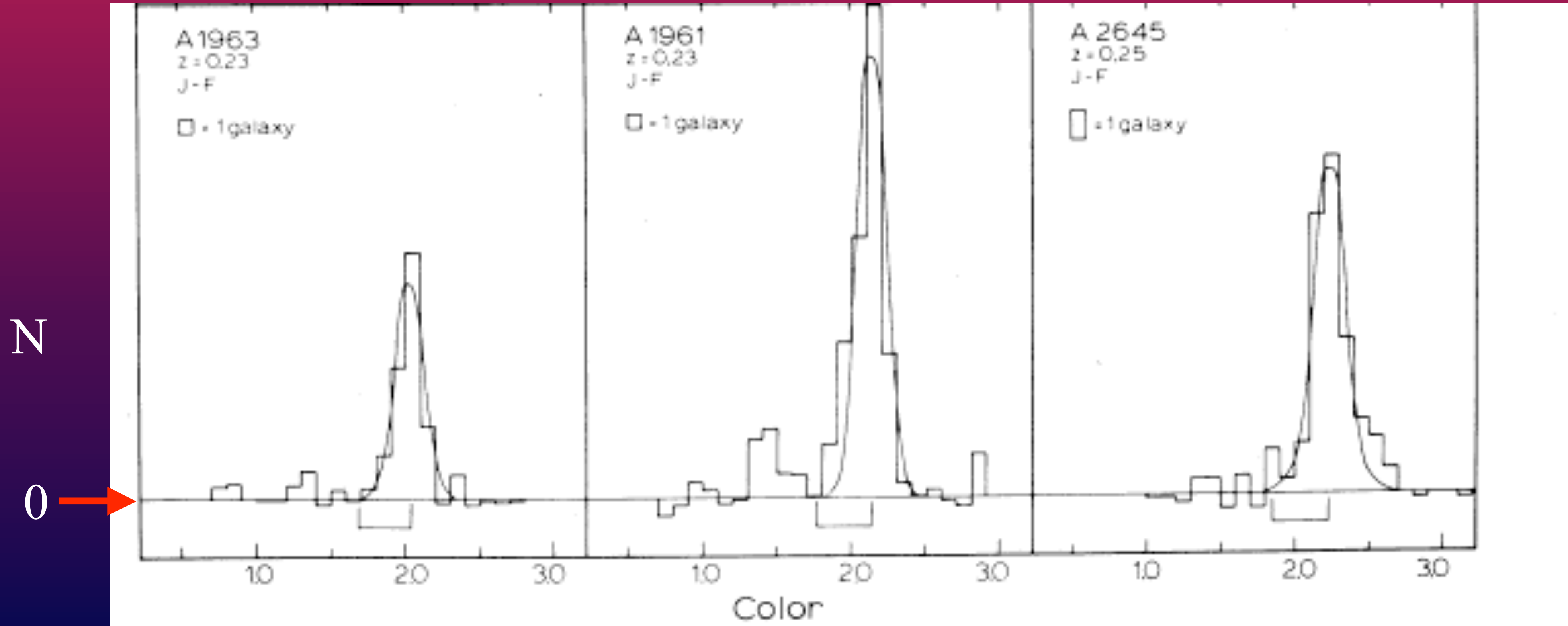
Radius



Average of many local groups

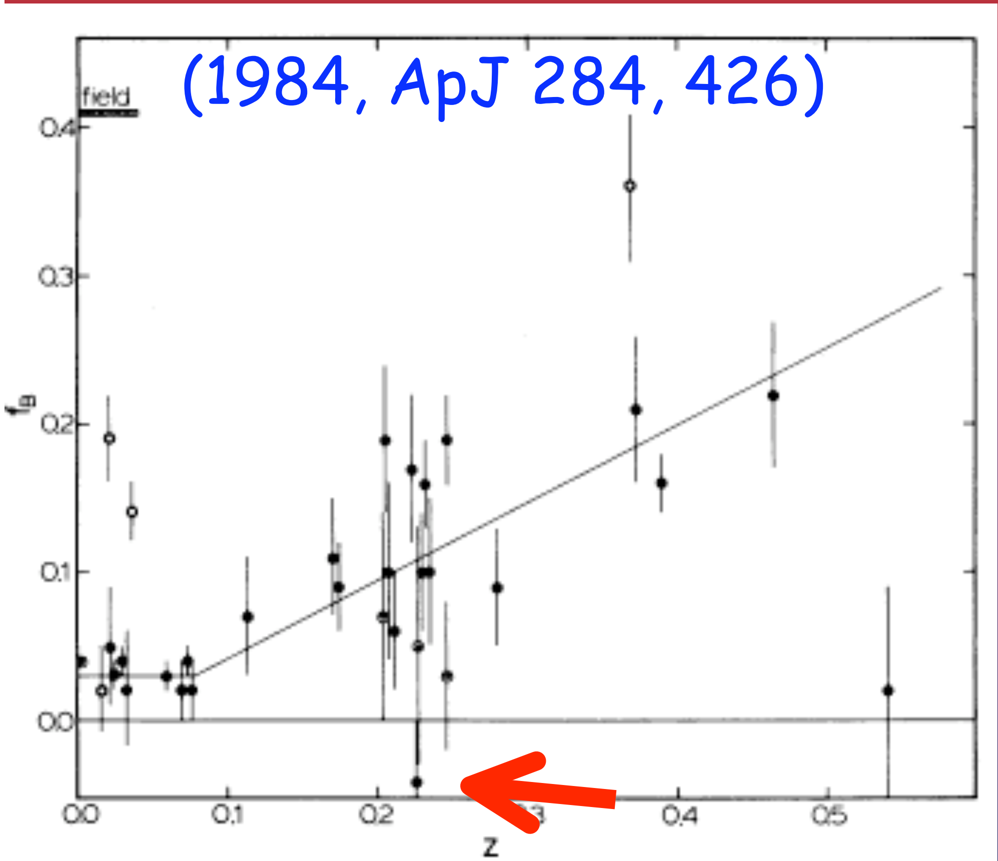
(2005, ApJ 633, 122)

colour distribution

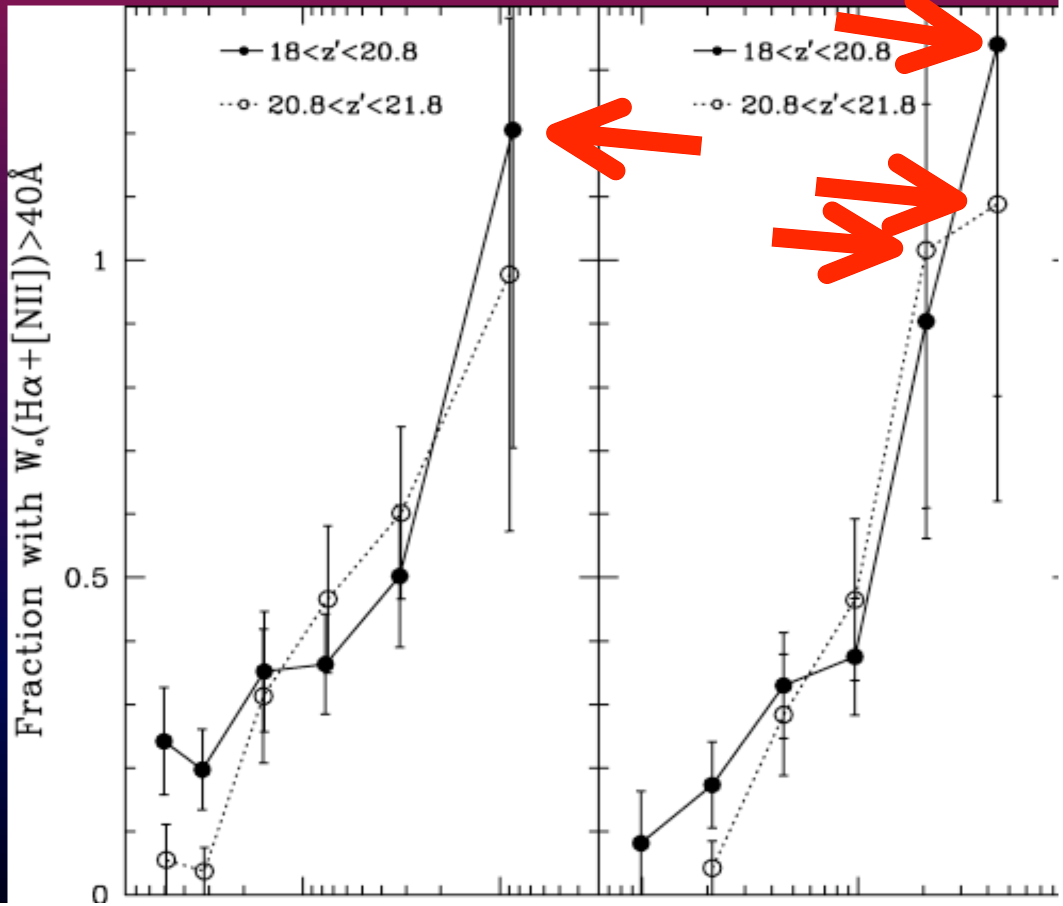


(1984, ApJ 284, 426)

… fractions

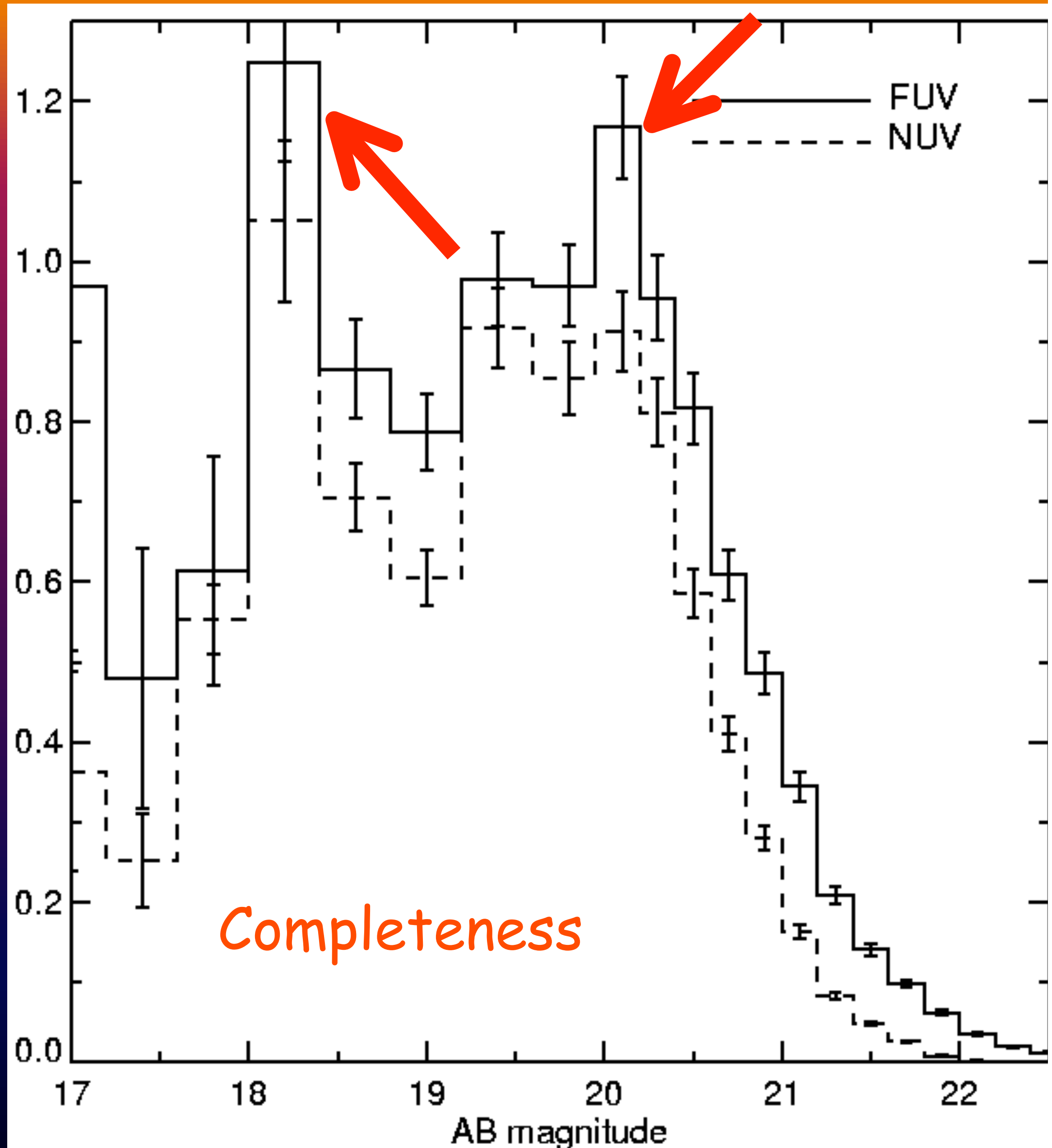


(2004, MNRAS 354, 1103



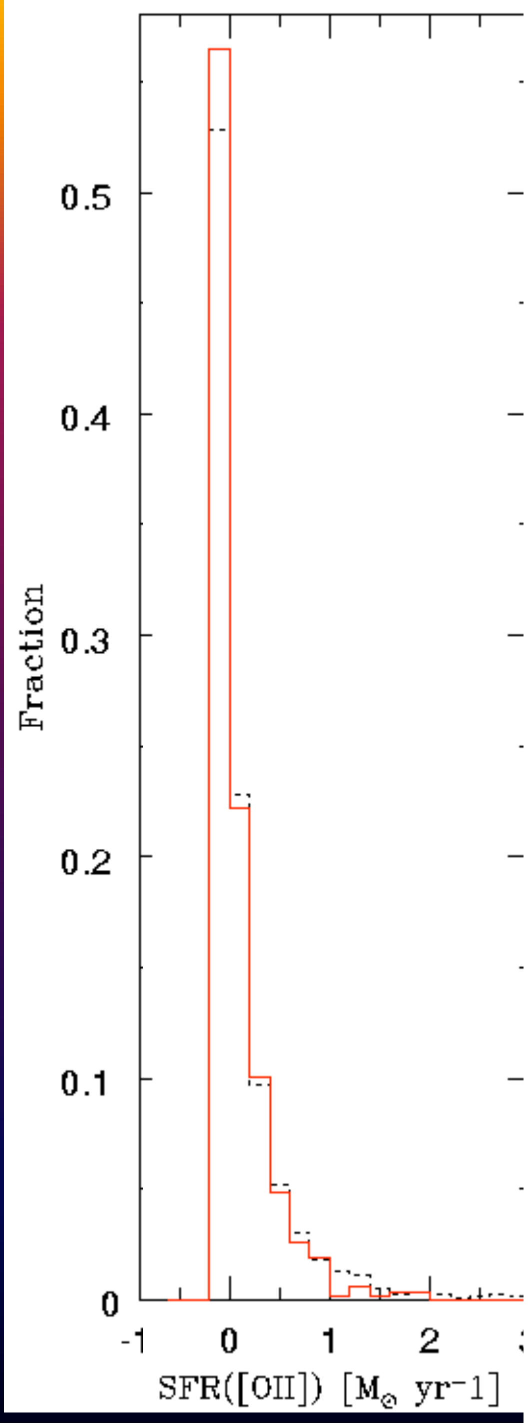
Every physically acceptable value is better than the claimed best value!

Completeness



Negative masses accelerating the expansion of the universe? (joke)

SFR



Bayesian methods force assumptions to be listed and be obvious.

Astro: the determination of the luminosity evolution of galaxies by mean of luminosity function (LF, i.e. n(L)dL) determinations (a recurrent theme).

Suppose we would like to measure the LF and its evolution with redshift. LF computation: two galaxies with the same L but different redshifts end up at the same L or different L's? In order to know it, we need to known how L evolves, which is derived from n(L)! Risk of a circular raisoning: assumed evolution to compute n(L) to derive evolution!

Most of LF determinations (listed in Andreon, 2004, A&A 416, 865) assume no luminosity evolution, find a luminosity evolution and do not revisit the analysis, although conclusions are in contradiction with hypothesis under which conclusions are derived.

Bayesian solution to nuisance parameters

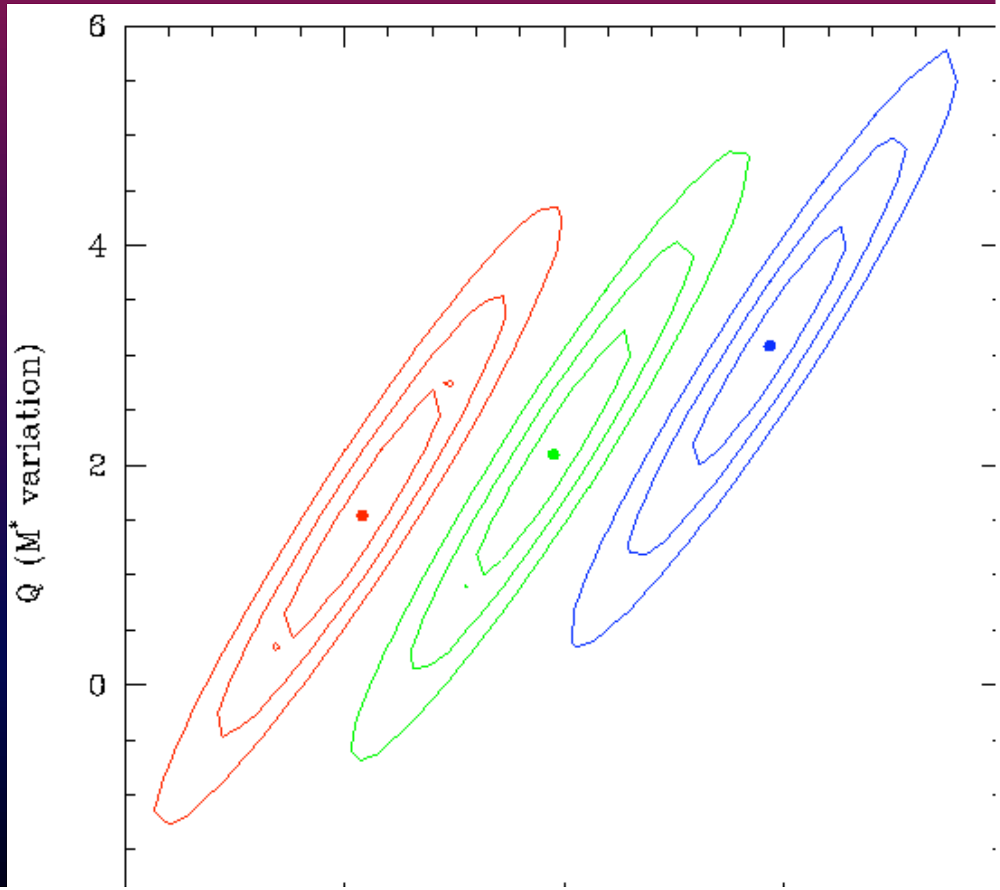
Empirically, claimed LF errors are underestimated by a factor 2 at least (Andreon, 2004, A&A 416, 865)

consequence of sum rule of probabilites: $p(x|I) = \int p(x,y|I)dy$

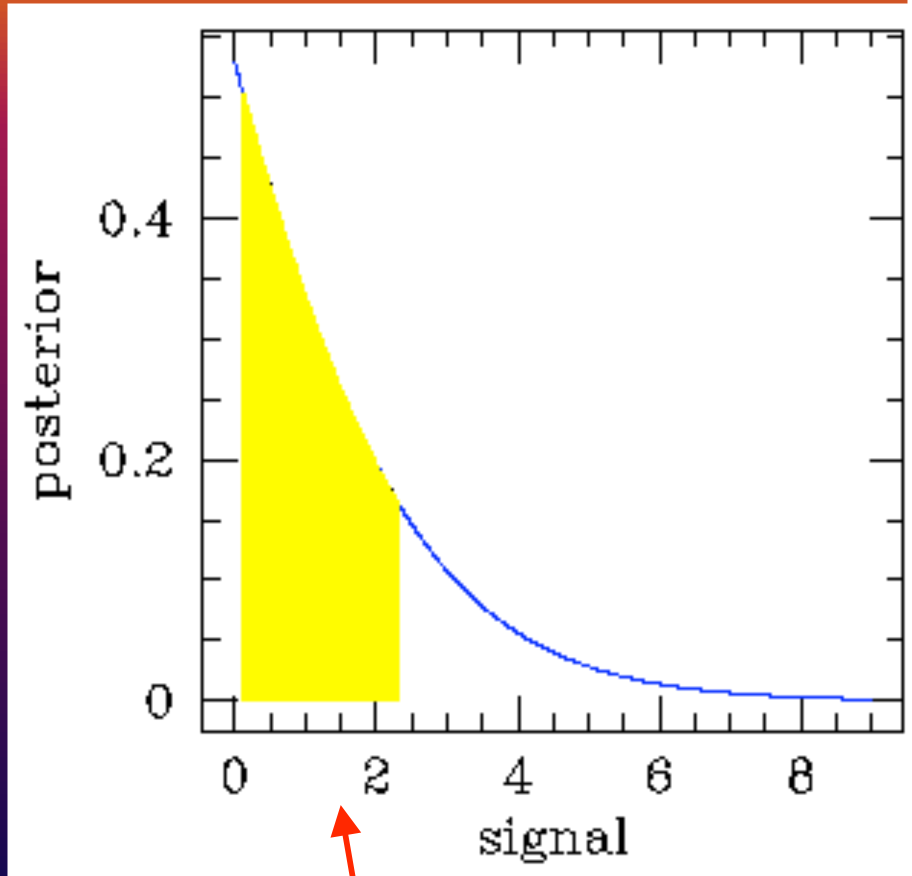
Std derivation assumes no evolution on M* in order to derive it, i.e.

$$Q = \frac{\partial M^*}{\partial z} = 0$$

the unknown nuisance parameter Q has been taken fixed instead of marginalizing over it.



Bayesian solution to boundary problem: the Bayes theorem



the classical estimate of a Poisson signal in presence of a background

n_tot = 3, n_bkg = 5

Astro recipe:

n_net=n_tot-n_bkg=3-5=-2

Bayes:

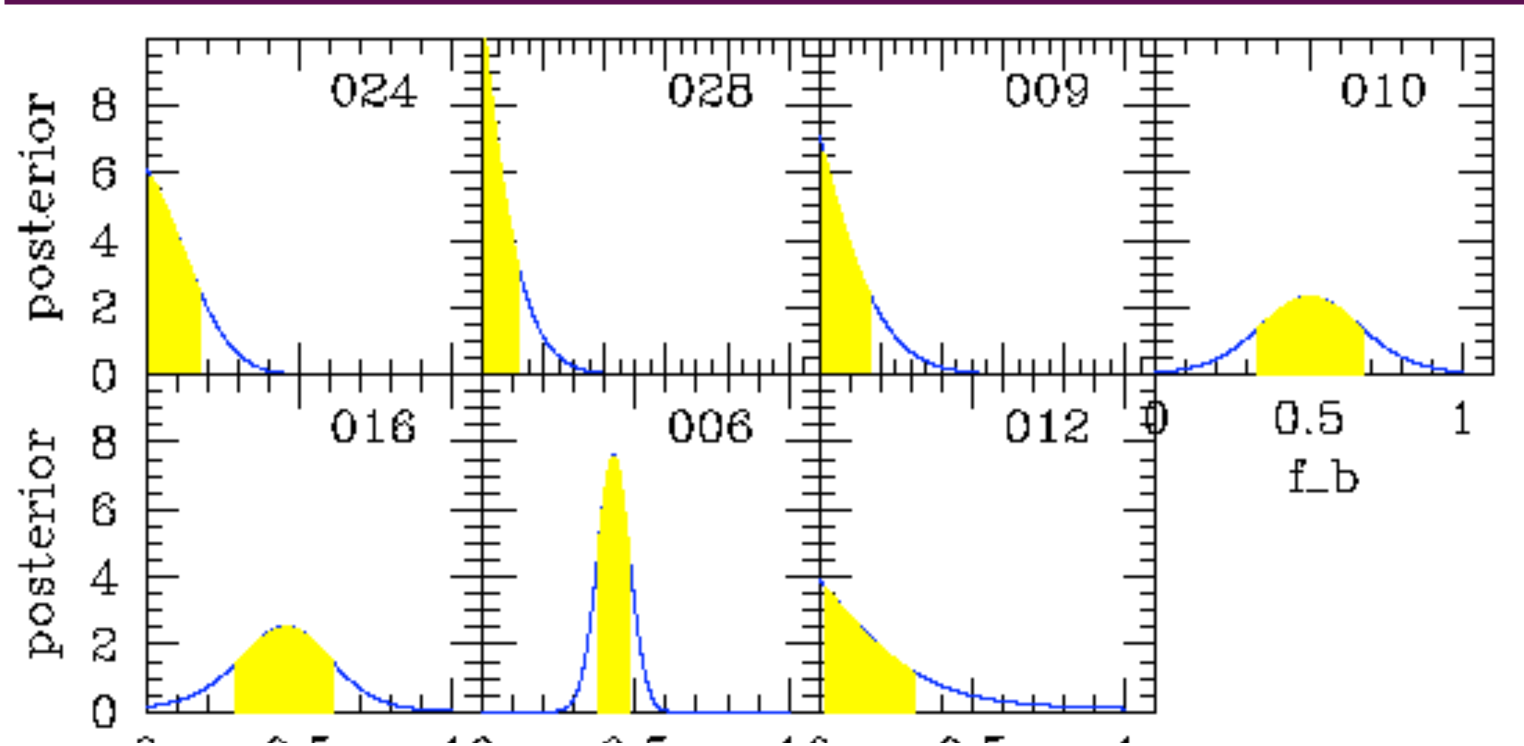
Bayes theorem: p(y|x,I)=c*p(x|y,I)*p(y|I)
Posterior = Poisson(n_tot; n_bkg+n_net) * p(n_net)

Since p(n_net)=0 if n_net>0, posterior($\forall$ n_net<0)=0

Life is hard: boundaries and nuisance parameters often got toghether.

Astro: The evolution of the blue fraction (Butcher-Oemler effect) measures the change with cosmic times of the star formation activity in clusters.

Stats: Fraction in presence of a background: D'Agostini (2004, physics/0412069)

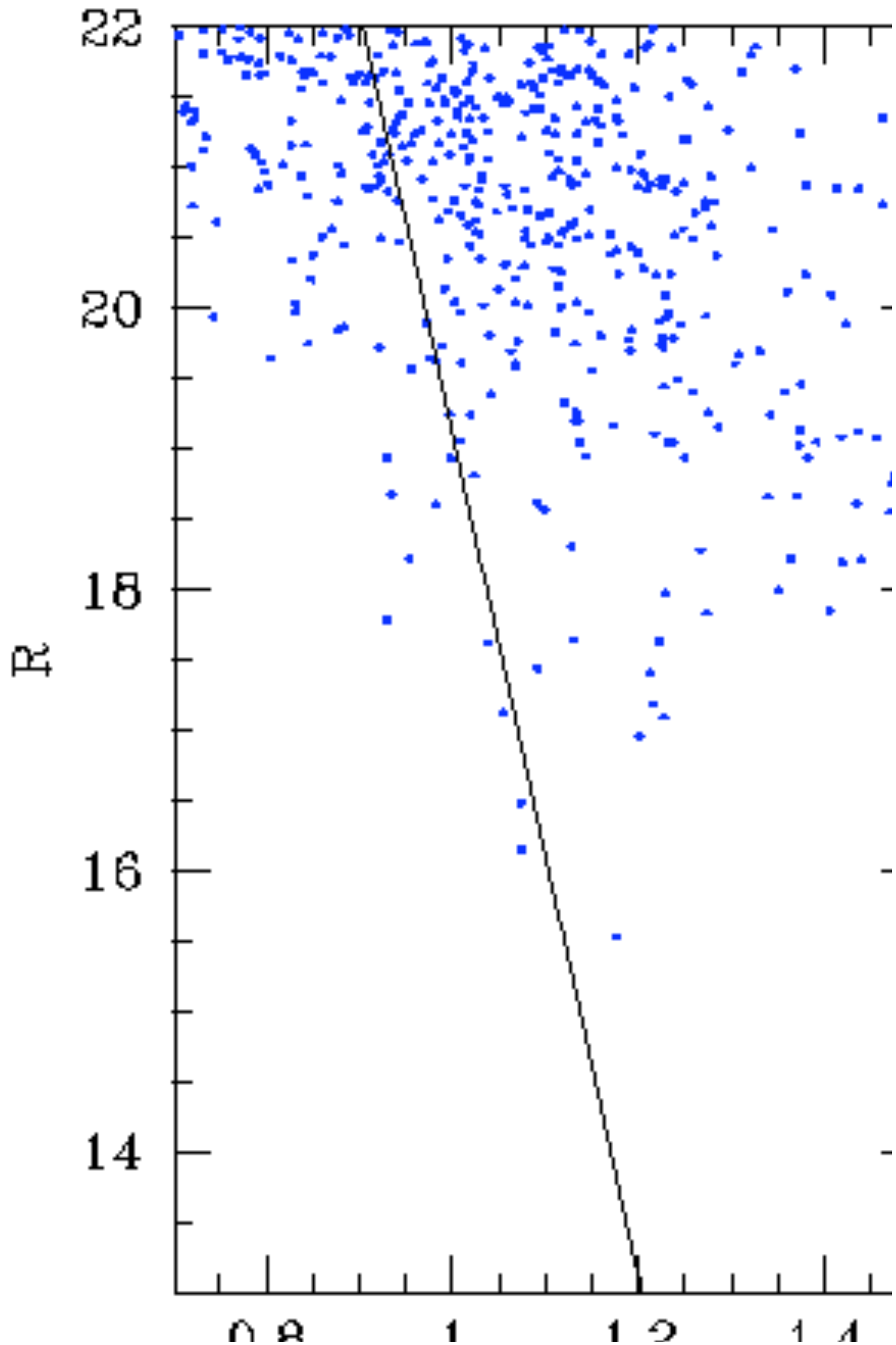


Andreon et al. (2006, MNRAS 365, 915).

Harder and harder.

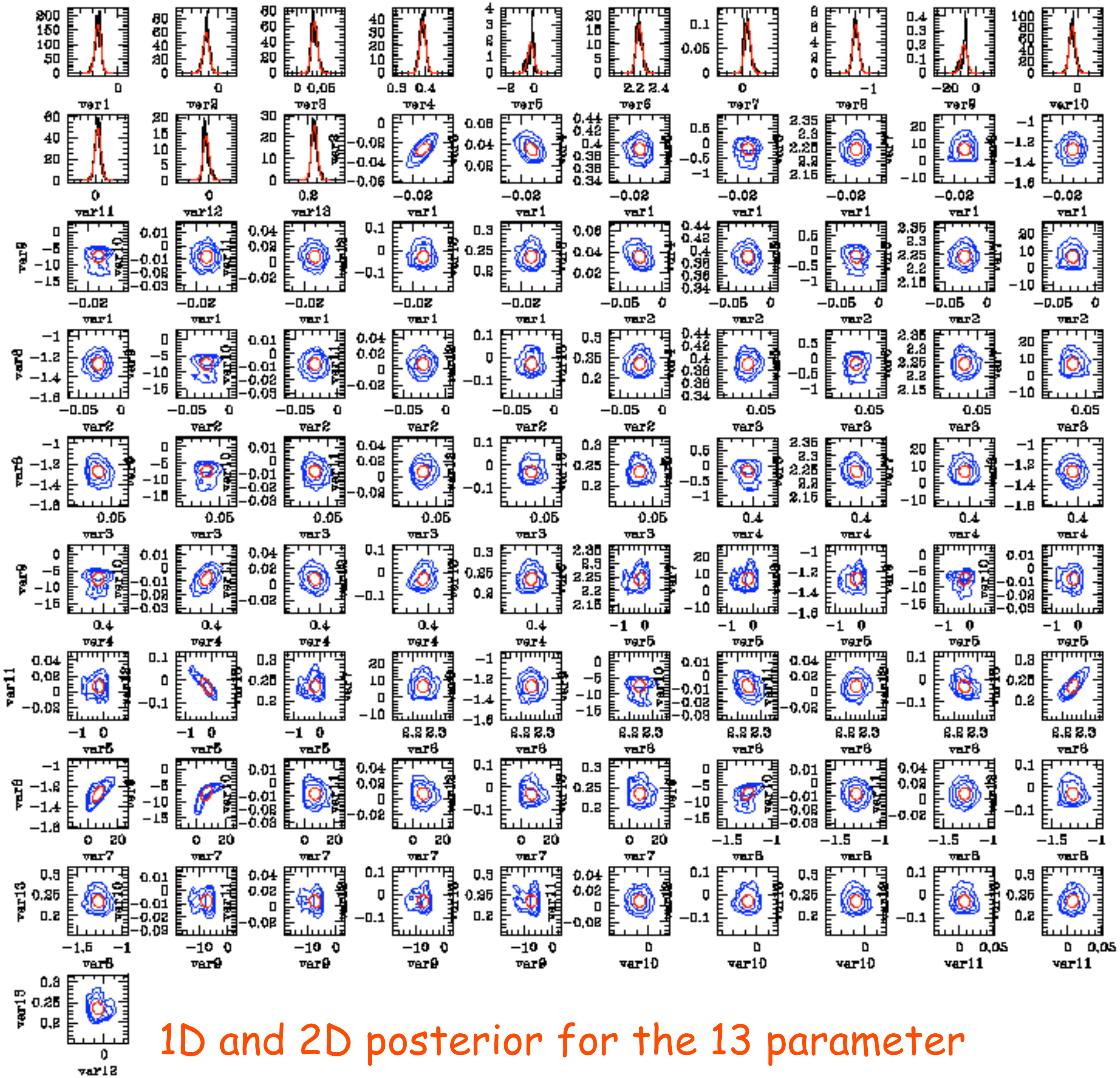
Astro: the scatter around the colour-magnitude relation put a strong constraint on the age of stars in these objects and, indirectly, on the ages of the galaxies themselves.

Problem: background galaxies (about 4 for every cluster galaxy).



Stats: determination of the intrinsic scatter of a correlation in presence of heteroscedastics errors on both x and y (solution due to D'Agostini 2005, physics/0511182) without the precise knowledge of which galaxies are cluster members, i.e. in presence of a background possibly displaying another correlation (solution due to Andreon 2006, MNRAS, in press, astro-ph/0603605).

Marginalization in a large dimensional space by using MCMC stochastical computations.



1D and 2D posterior for the 13 parameter

A real case: boundaries, marginalization and model selection

Astro: How the assembly of galaxy masses proceed? Evolution of the 3.6 micron LF measures the growth history of galaxy masses.

Stats: determination of the 3.6 micron LF and model selection among various possible mass growth histories.

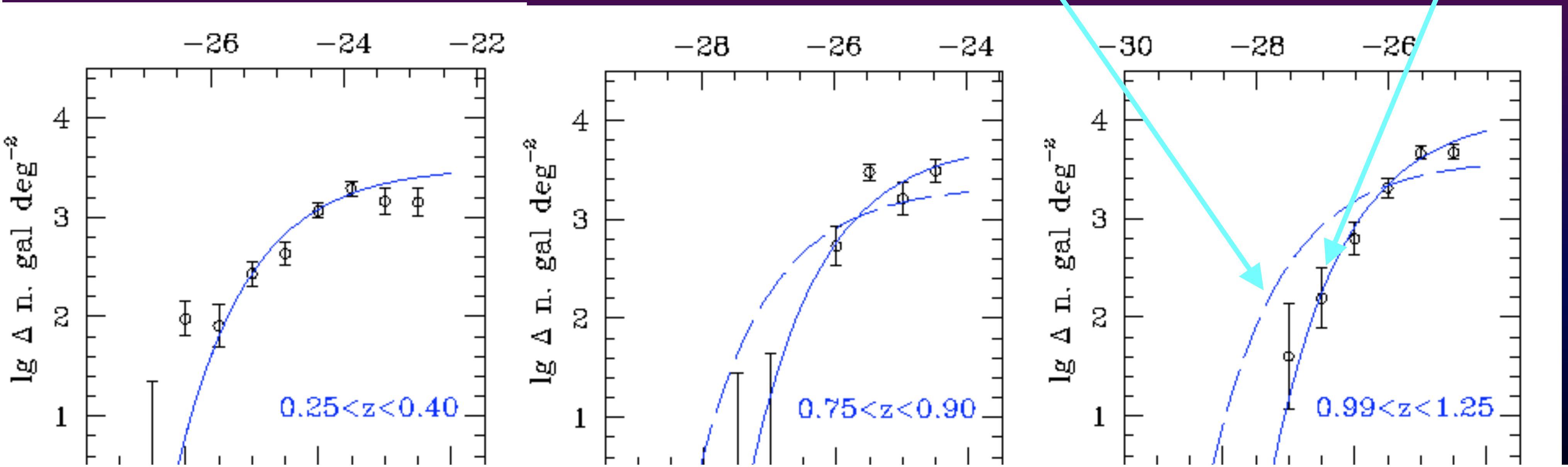
Data: 1000 member galaxies (plus a 4500 background galaxies, whose distribution is estimated from a larger background formed by 107000 galaxies) from one of Legacy Spitzer surveys (SWIRE)

Two derivations: standard and bayes, both published in Andreon (2006, A&A 448, 447), so I can blame without bless anyone ...

Simplistic (std) analysis.

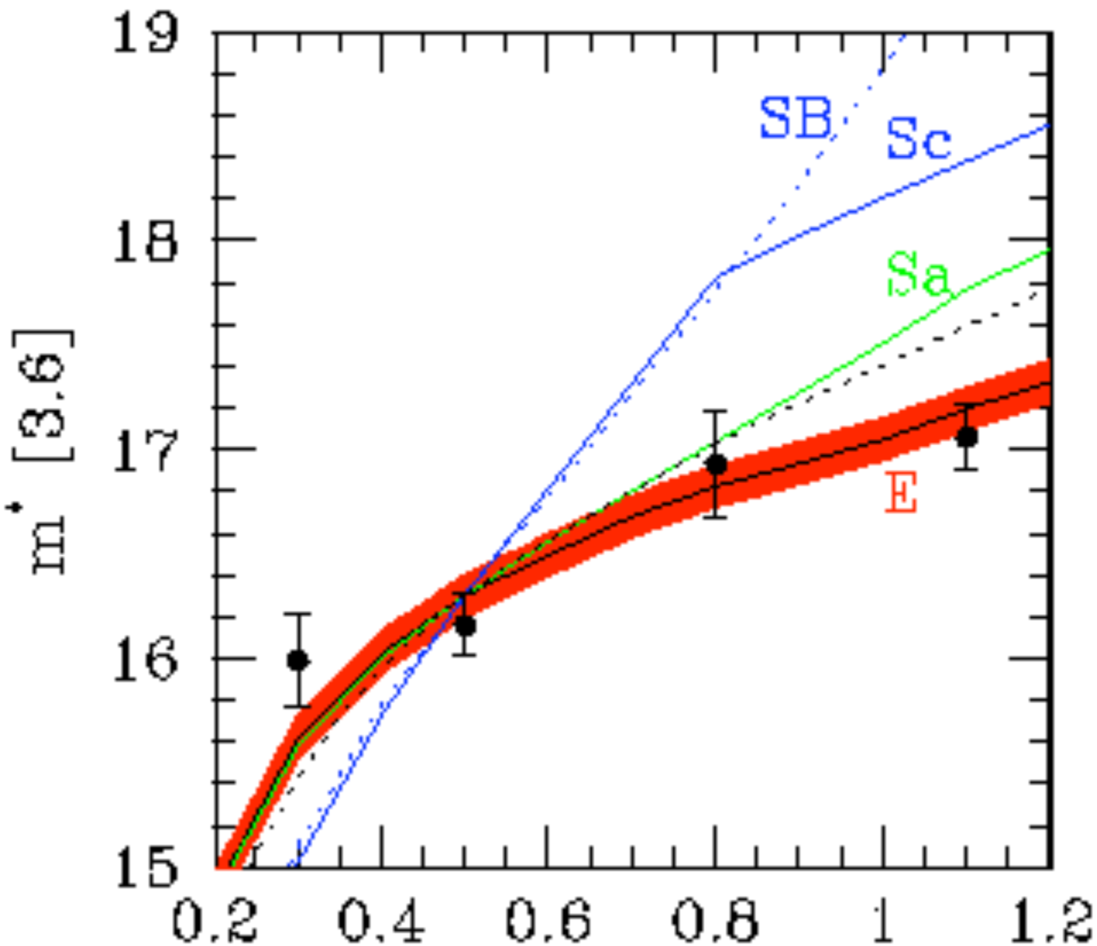
Step 1: parameter estimation

- Bin in z and mag, don't care if bins are optimally chosen
- Don't care if n(L) is defined to be positive and found negative (positive background fluctuation)
- Assume no evolution, don't worry the risk of a circular raisoning
- take unconstrained parameters (alpha) fixed, i.e. neglect the role of nuisance parameters



Simplistic (std) analysis

Step 2: simplistic model selection. Compare data and models in order to select the best model.



E models are obviously best, but does other model are rejected?

Do the best model need to be refined? (model complexity issue)

At this point of the talk, we known that problems are there. Therefore I stop with the simplistic analysis, and I use bayesian methods.

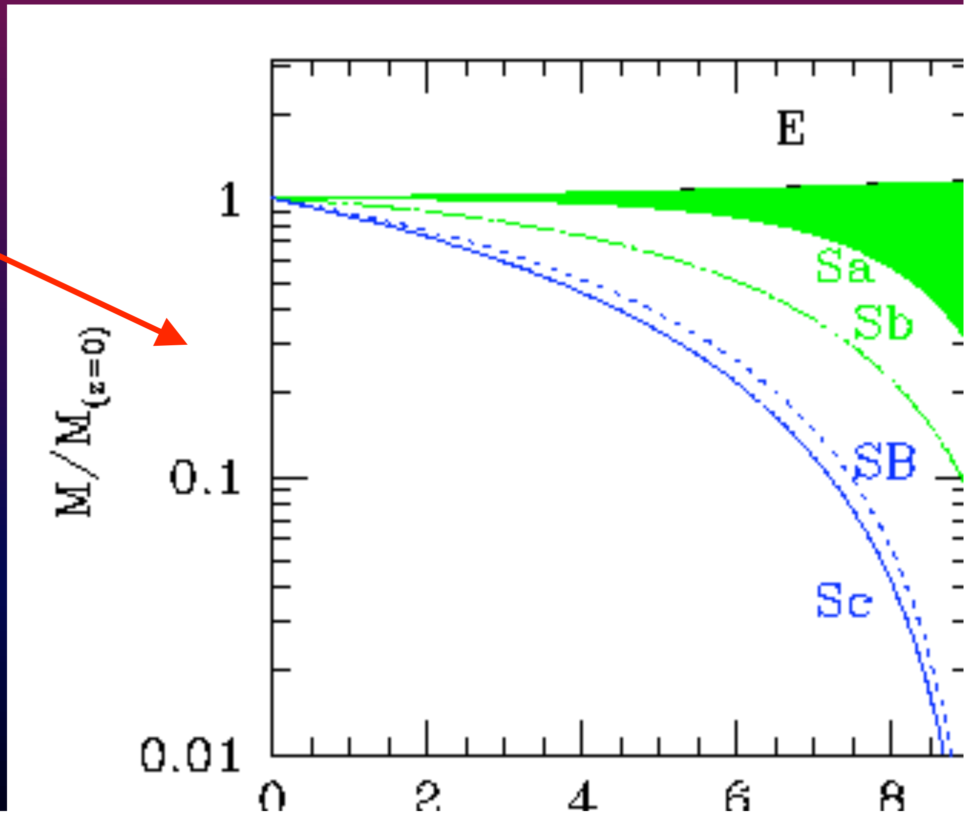
Bayesian analysis

- Don't bin in z and mag, bayes don't require bins,
- Account for boundaries: n(Mass)dMass is positively defined.
- Model evolution and, eventually

  refine the model
- Marginalize over nuisance parameters

  (alpha).

Mass growth histories, converted in 3.6 micron luminosity evolution by using



Bayesian analysis

Step 1: model selection.

Models are not hierarchically nested, likelihood ratio test cannot be used.

Questions to answer:

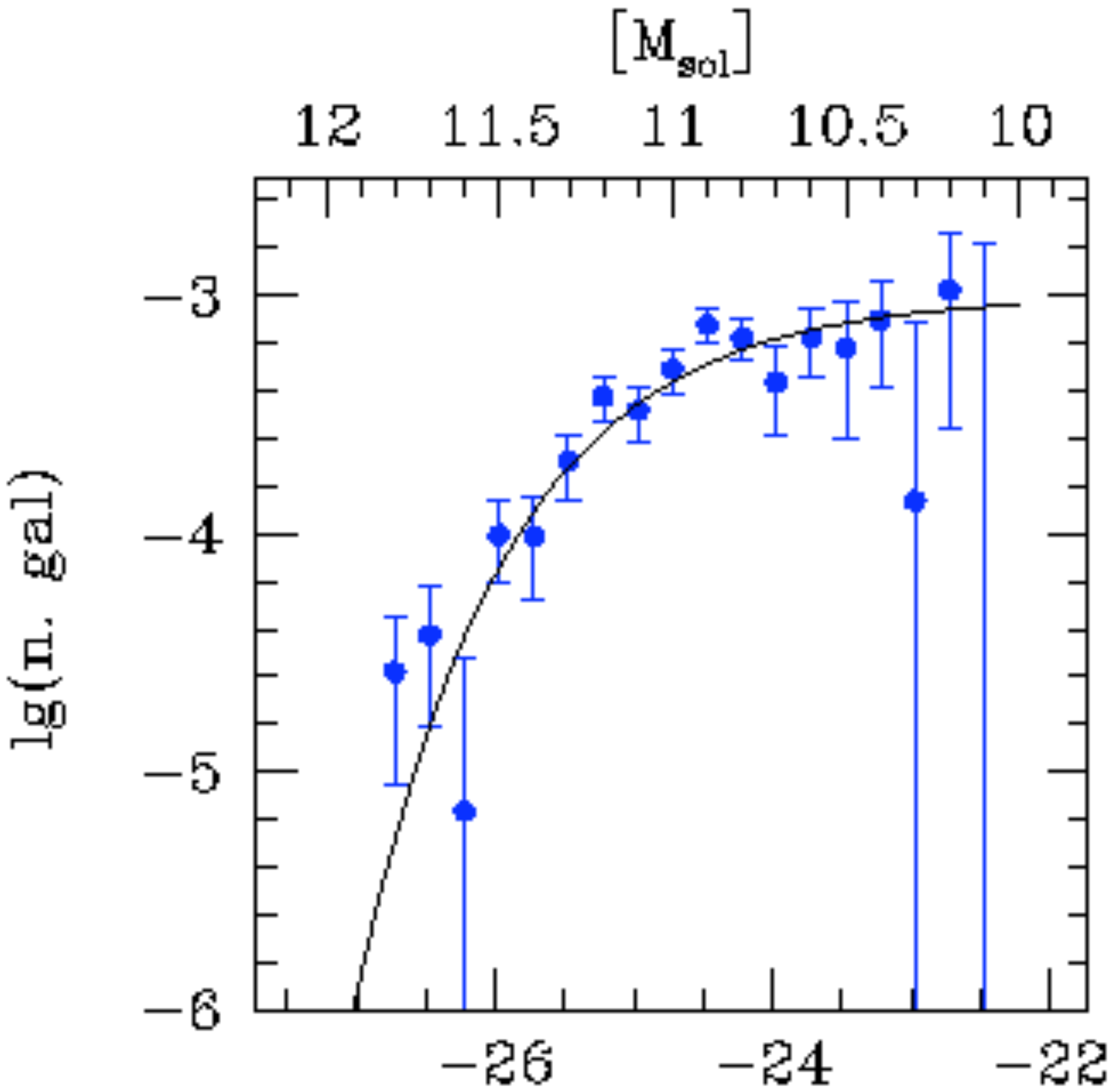
1) Which model best describes the data? Which models are rejected?
2) Do models need to be refined by a further evolutionary term (taken prop to z)? (model complexity)

I used the $BIC = -2 \ln L_{best} + k \ln N$

R1: E (no mass growth) models are preferable to all the other ($\Delta BIC > 5$)
R2: no

Mass Function

… I known the mass evolution and I known at which mass I should put galaxies observed to have $mass_i$ at z_i .



Summary

In my field, Bayesian methods are almost unknow.

Imaginary values of velocity dispersions

Possible negative M/L ratios of clusters

Negative star formation rates

Fractions of blue galaxies outside [0-1]

Spectroscopic of photometric completeness larger than 1

Number density profiles of cluster systematically negative

Conclusions in contradictions with hypothesis

Thank you