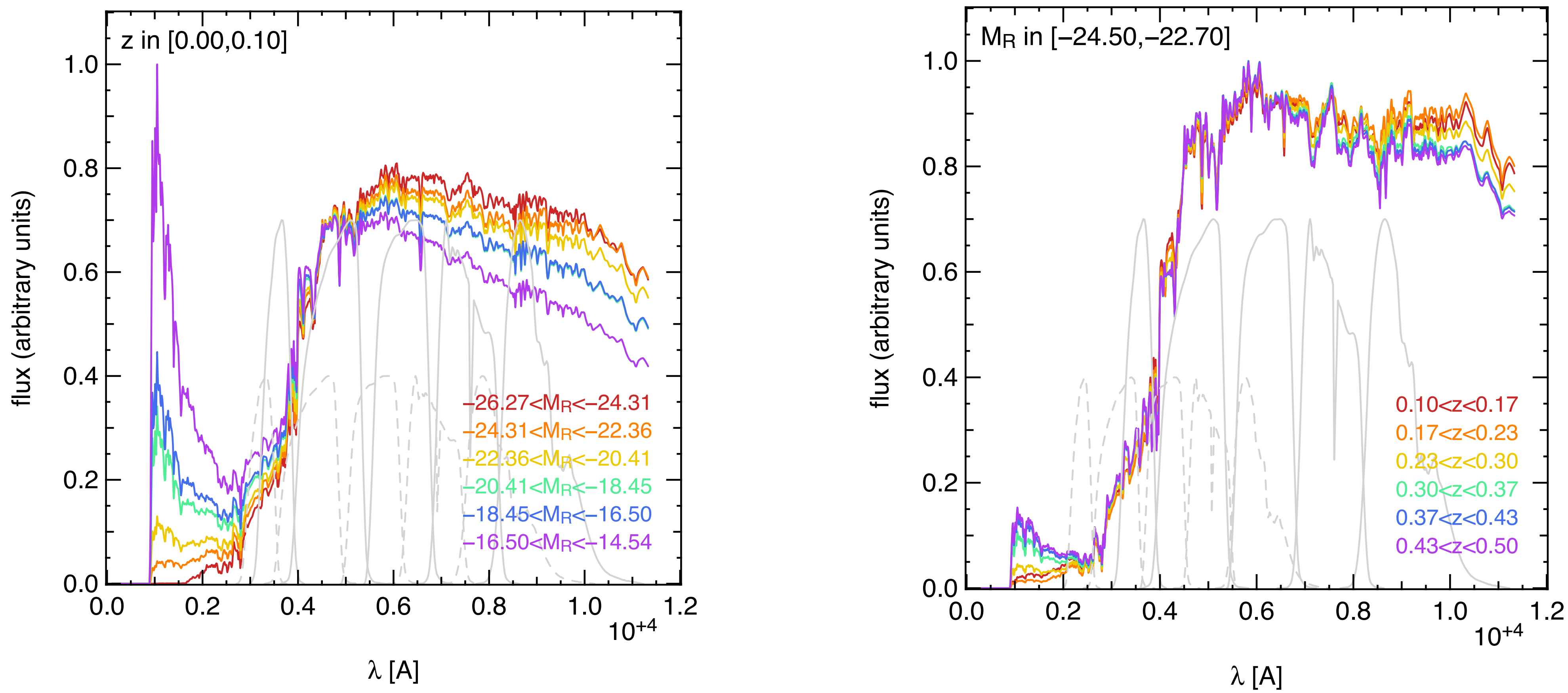
Photometric Redshifts

R.P. Saglia, MPE

(not really an expert on photometric redshifts, nor on Bayesian statistic!)

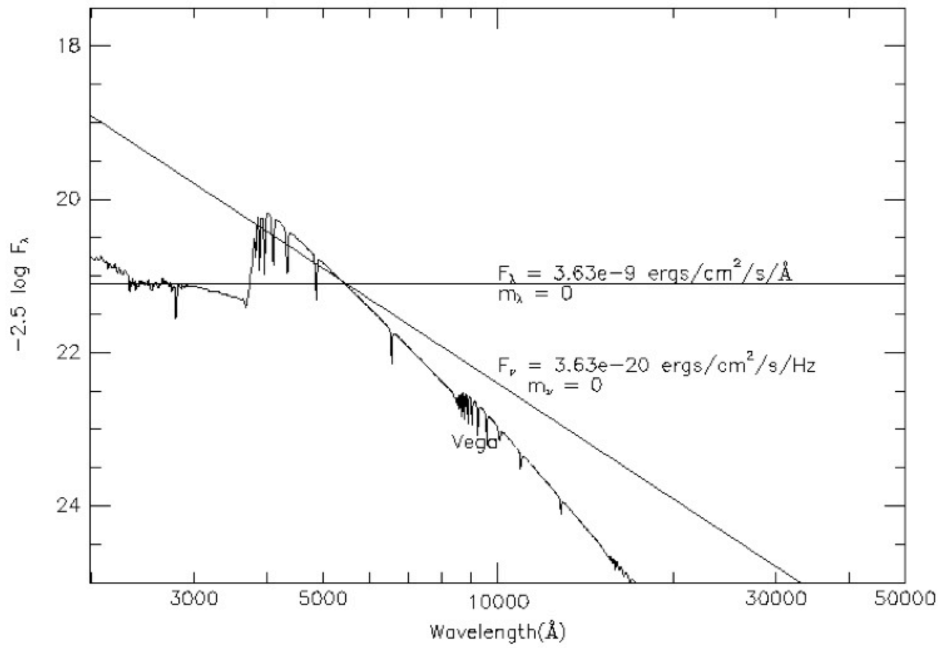
with R. Bender, N. Greisel, S. Seitz, R. Senger, J. Snigula

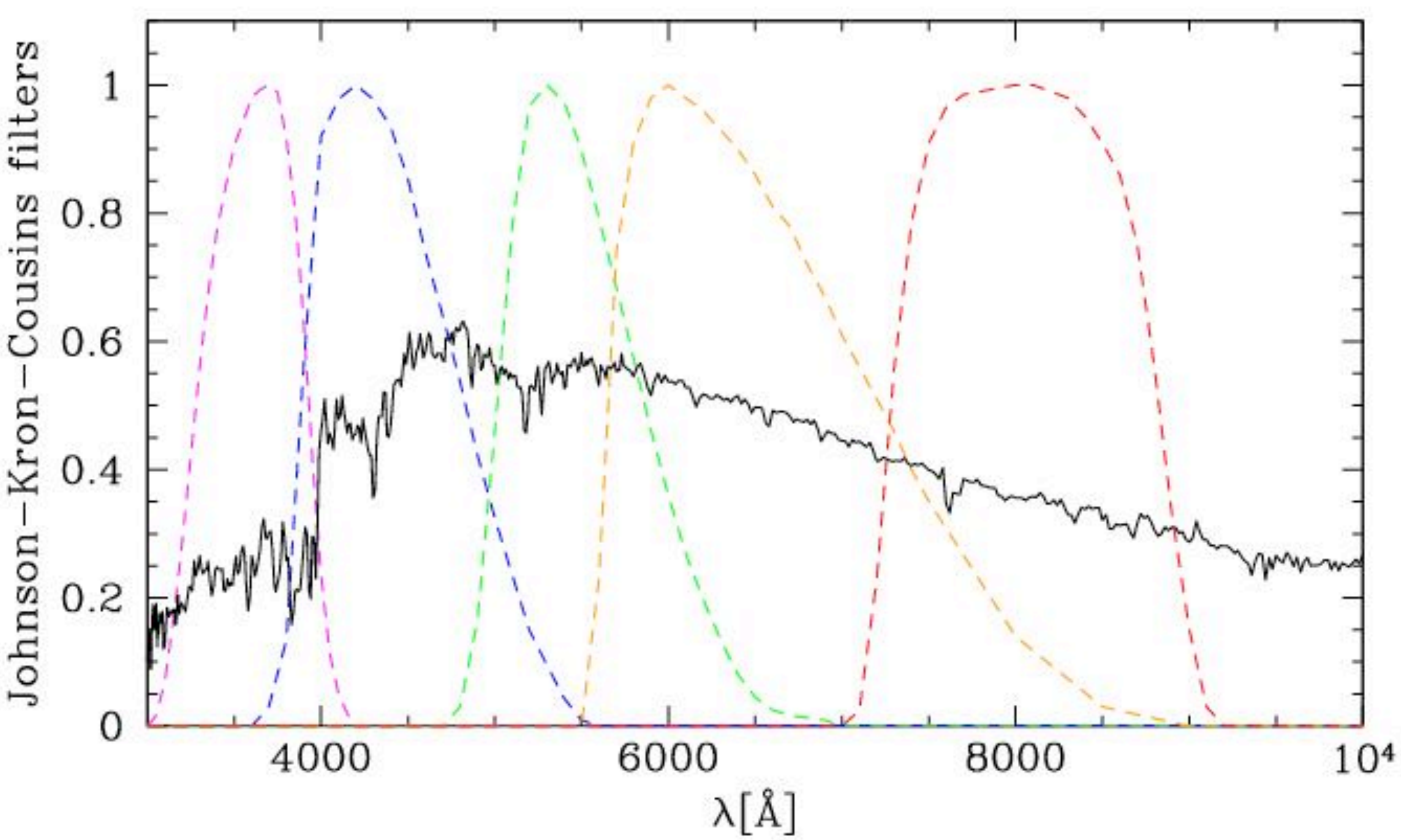


Outline

- Introduction
- Empirical methods
- Template (Bayesian) fitting
- Luminous red galaxies
- Conclusions

Definitions





Redshift z:

$$z = \frac{\lambda_{obs} - \lambda_{lab}}{\lambda_{lab}}$$

Color:

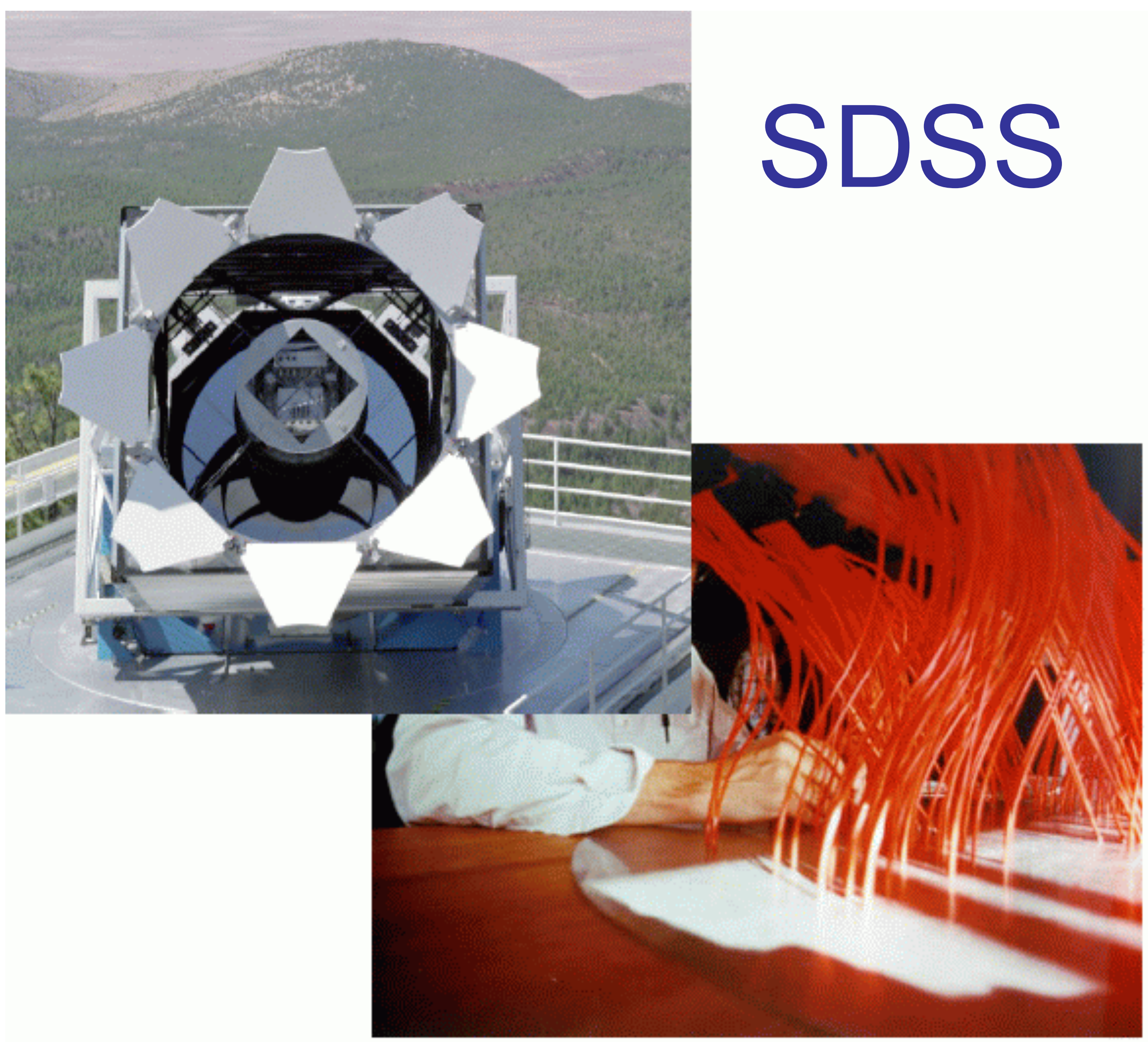
U-B = $m_U - m_B$

B-V = $m_B - m_V$

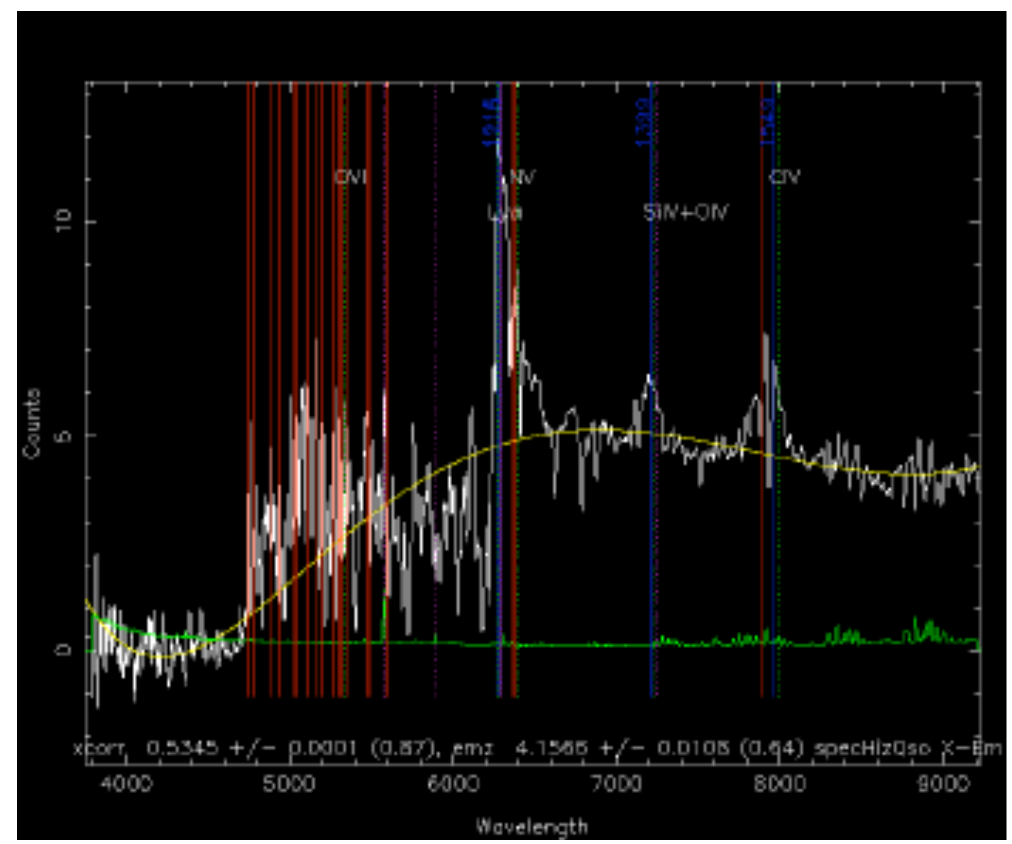
$$m_x = -2.5 \log \left[\frac{\int f_\nu T_x(\nu) d\nu}{\int f_{\nu,Vega} T_x(\nu) d\nu} \right]$$

Introduction

- Spectra at moderate/medium/high resolution provide typing (star/galaxy/quasar) and distances (redshifts) to astronomical objects with better than permille precision
- Spectroscopy is however costly and increasingly difficult at faint magnitudes



600× multi-object spectrograph

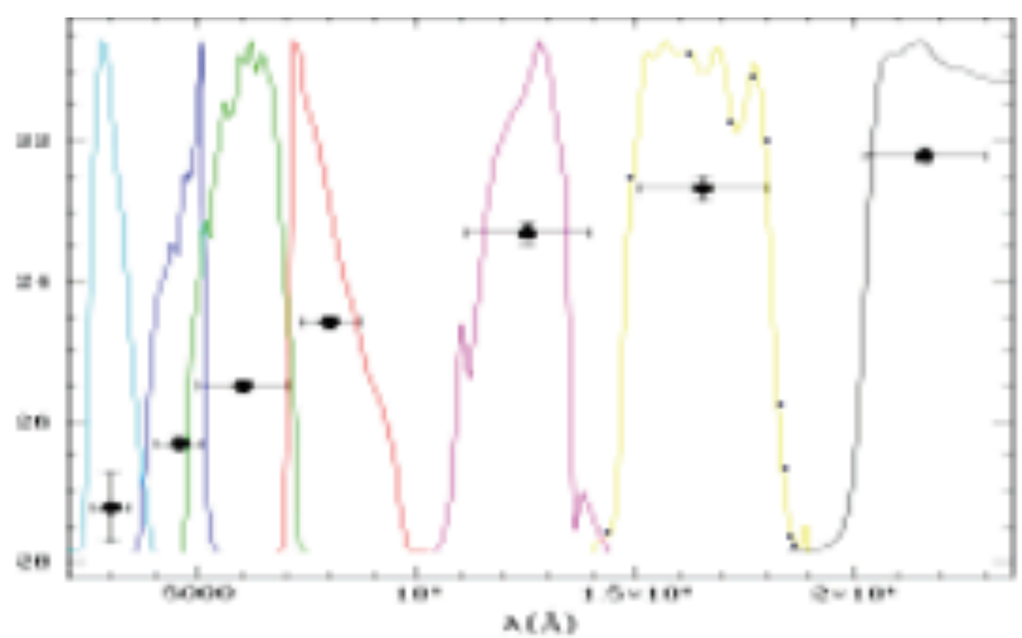


Introduction

- Broad multi-band photometry can be seen as very low-resolution spectroscopy and provides an alternative to spectroscopy, when classifications and distances of large numbers of (faintish) objects are needed, and precisions at the percent level suffice.

- The photometric SDSS provides ugriz photometry for millions of galaxies down to g~22, PanSTARSS1 will provide grizy photometry down to slightly fainter magnitudes on twice the sky coverage

- Running (KIDS) or soon starting (DES) ground-based, or approved (EUCLID) space-based lensing surveys aim at measuring weak lensing signatures down to 24th magnitude, relying on photometric redshifts for distance estimations



How to measure a redshift

- With spectra: identify known features (emission/absorption lines, breaks)
- With photometry, two approaches:

1. Empirical: search for the mapping of fluxes and colors (and possibly additional information like morphology, concentration, etc...) into redshift

   Need training sets with spectroscopic redshifts that map 'all' existing galaxy types, extrapolations to fainter magnitude limits doubtfull. Galaxy properties (absolute magnitudes, spectral types, stellar masses...) need to be computed using template fitting techniques.

2. Template fitting methods (Maximum Likelihood, Bayesian)

   Need representative template sets incorporating galaxy evolution
   Allow extrapolation to fainter magnitudes
   Not only 'best fitting' redshift, rather full probability distribution and galaxy products as a byproduct.

A list of empirical methods

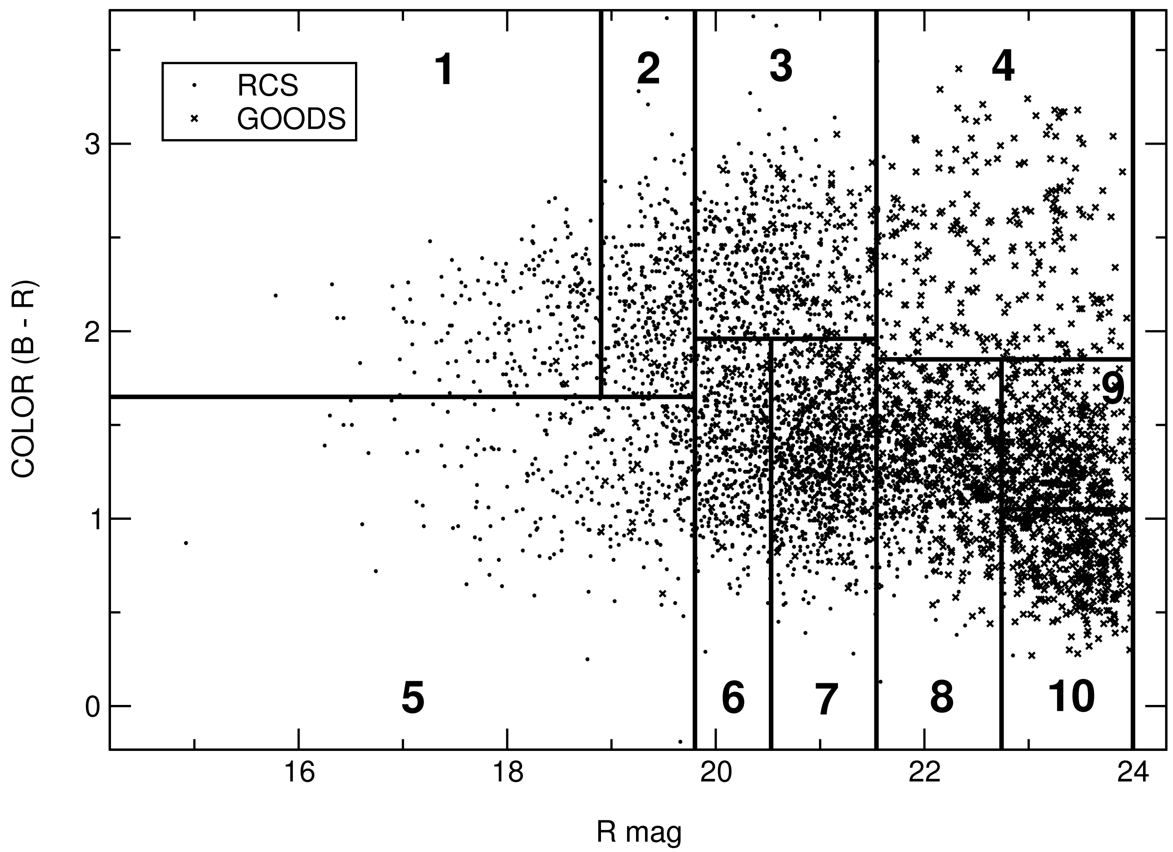
- Artificial neural networks: Collister & Lahav 2004, PASP, 116, 345
- Boosted decision trees: Gerdes et al. 2010, ApJ, 715, 823
- Ensemble lerning: Way & Srivastava, 2006, ApJ, 647, 102
- Gaussian process regression: Bonfield et al. 2010, MNRAS, 405, 987
- Kernel regression: Wang et al. 2007, MNRAS, 382, 1601
- Polynomial fitting: Hsieh et al. 2005, ApJSS, 158, 161
- Random forest: Carliles et al. 2010, ApJ, 712, 511
- Spectral connectivity: Freeman et al. 2009, MNRAS 398, 2012
- Support vector machines: Wadadekar 2005, PASP, 117, 79

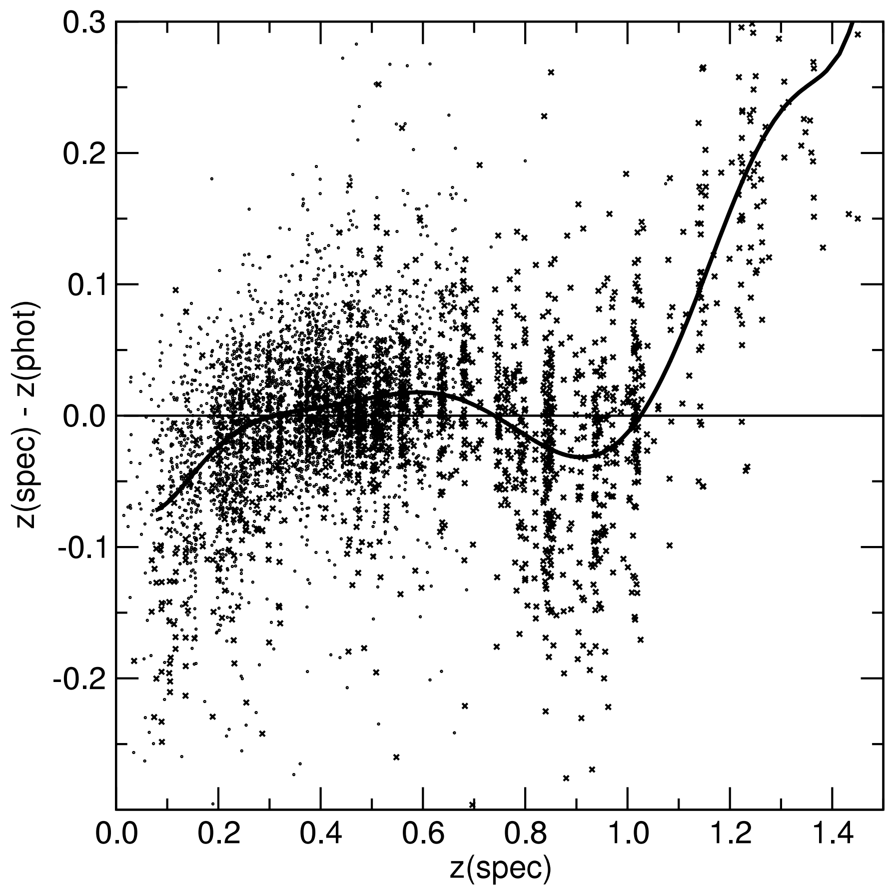
Polinomial fitting

$$\begin{aligned}\text{redshift} = \text{constant} &+ a_0B^2 + a_1V^2 + a_2R_c^2 + a_3z'^2 \\ &+ b_0B + b_1V + b_2R_c + b_3z' + c_0BV \\ &+ c_1BR_c + c_2Bz' + c_3VR_c + c_4Vz' + c_5R_cz'.\end{aligned}$$

kd-tree to divide training set into cells

Hsieh et al. 2005, ApJSS 158, 161





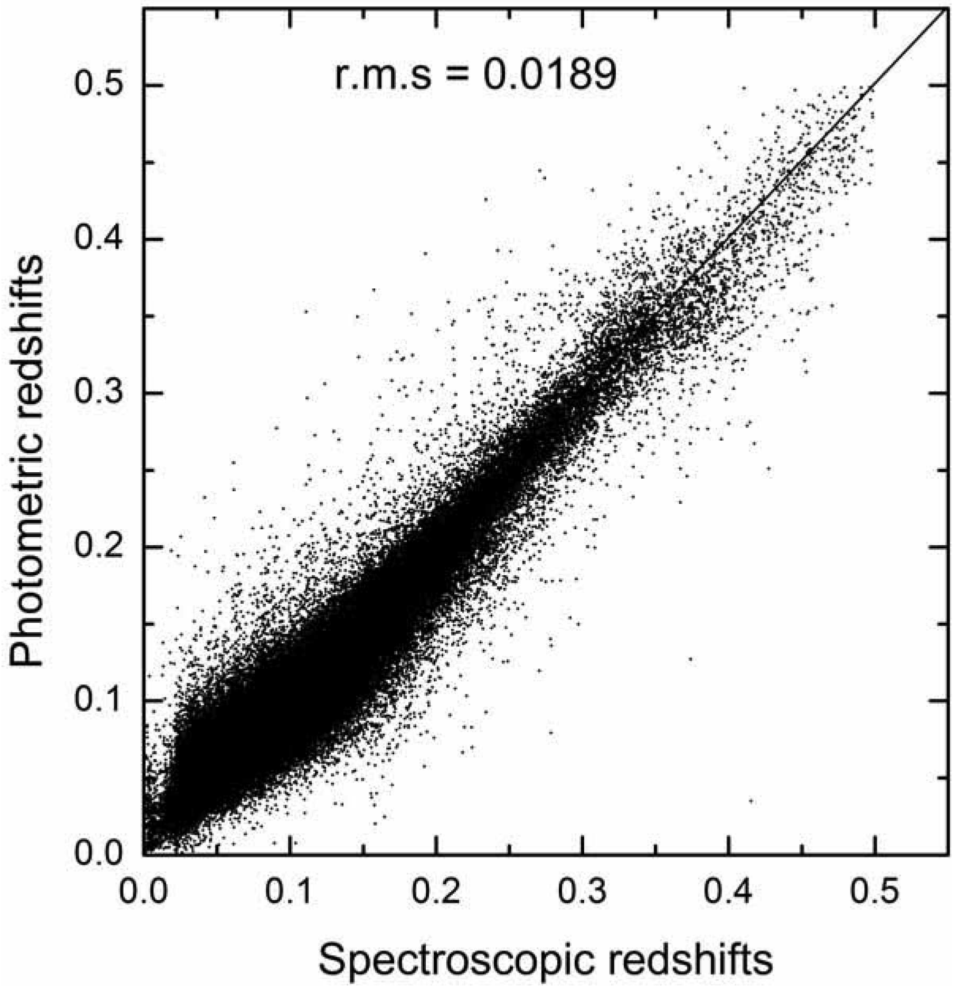
Kernel regression

$$y_{\mathrm{q}} = \frac{\sum_{i=1}^{N} K\left(\frac{D(\boldsymbol{x_i}, \boldsymbol{x_q})}{h}\right) y_i}{\sum_{i=1}^{N} K\left(\frac{D(\boldsymbol{x_i}, \boldsymbol{x_q})}{h}\right)}$$

x_i : data vector of training set (colors, etc.)
x_q : data vector of object q
y_i : spectroscopic redshift vector of training set
y_q : photometric redshift of object q

D:distance (euclidian)
K: kernel function (gaussian)
h:bandwidth (from 10-fold cross-validation)

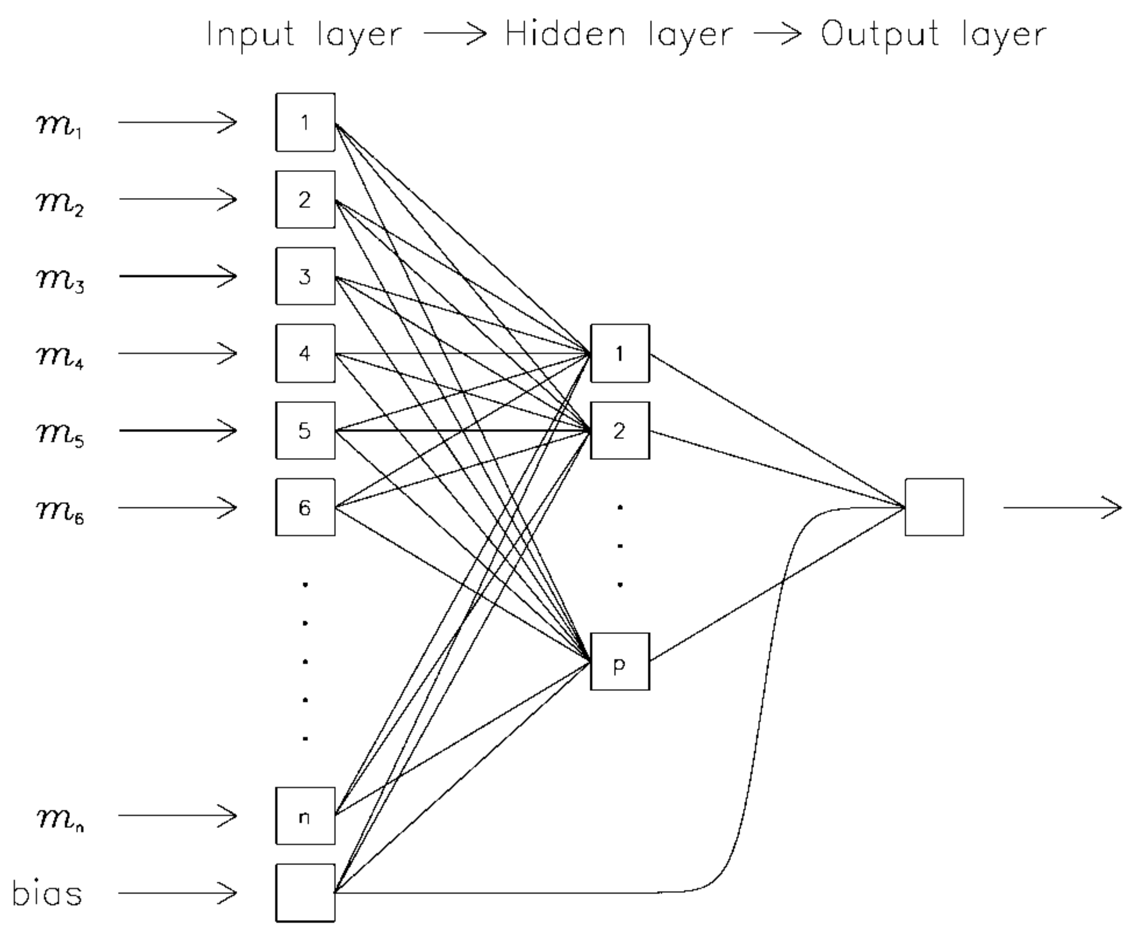
$$\mathrm{CV}(h) = \frac{1}{M}\left[\frac{1}{k_1}\sum_{i=0}^{k_1}(y_{1i} - \widehat{y}_{1i})^2 + \frac{1}{k_2}\sum_{i=0}^{k_2}(y_{2i} - \widehat{y}_{2i})^2 + \ldots + \frac{1}{k_M}\sum_{i=0}^{k_M}(y_{Mi} - \widehat{y}_{Mi})^2\right],$$



Wang et al. 2007, MNRAS, 382, 1601

ANNz

w_{ij} connection weights



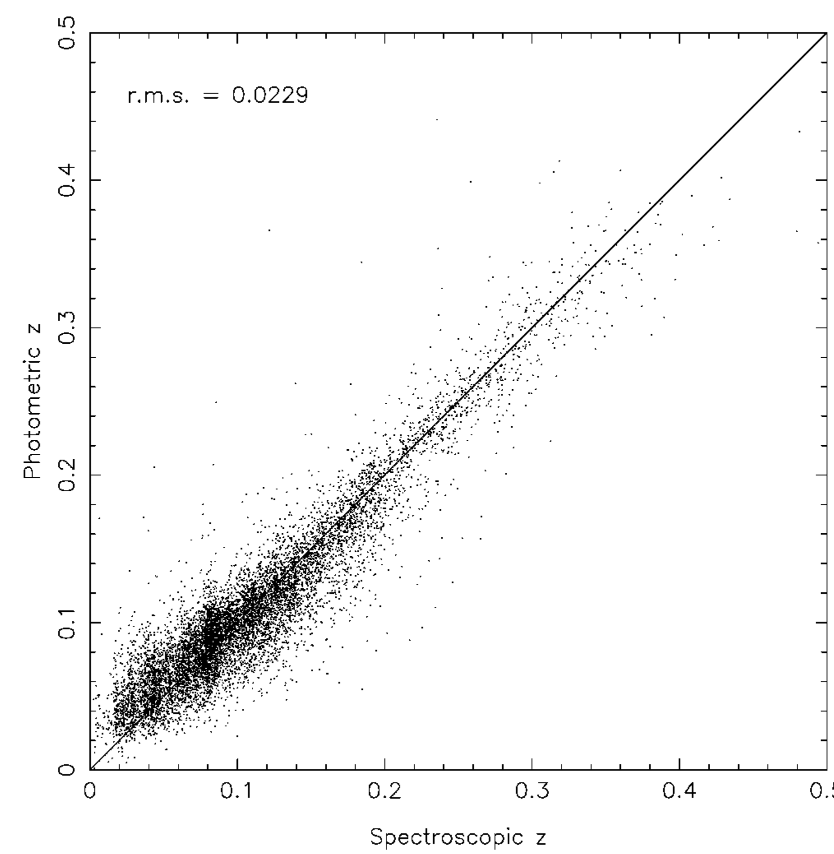
$$u_j = \sum_i w_{ij} g_i(u_i)$$

Activation function:

$$g_j(u_j) = \frac{1}{1+\exp(-u_j)}$$

Minimize cost function to determine the weights:

$$E = \sum_k [z_{\text{phot}}(\boldsymbol{w}, \boldsymbol{m}_k) - z_{\text{spec},k}]^2 + \beta \sum_{i,j} w_{ij}^2$$



Artificial neural networks

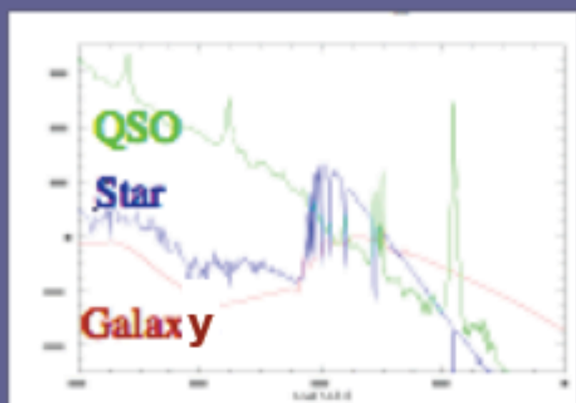
Collister and Lahav, 2004, PASP, 116, 345

A list of template fitting codes

- EAZY: Brammer et al. 2008, ApJ, 686, 1503
- GOODZ: Dahlen et al., 2010, ApJ, 724, 425
- Hyperz: Bolzonella et al. 2000, A&A, 363, 476
- ZEBRA: Feldmann et al. 2006, MNRAS, 372, 565
- Le Phare: Ilbert et al. 2006, A&A, 457, 841
- BPZ: Benitez 2000, ApJ, 536, 571

Template fitting

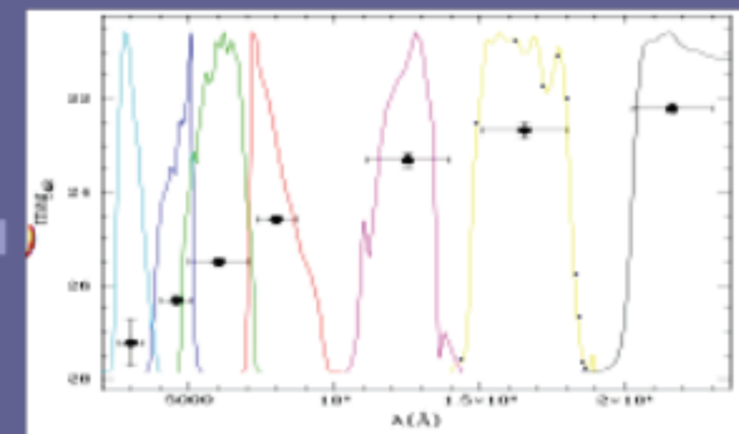
SED library



+Extinction
IGM

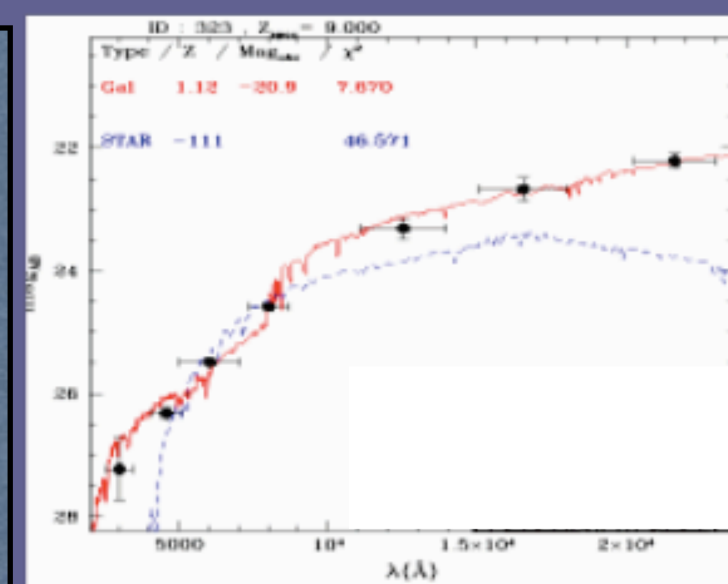
Observed flux +error to
theoretical flux

Set of filters



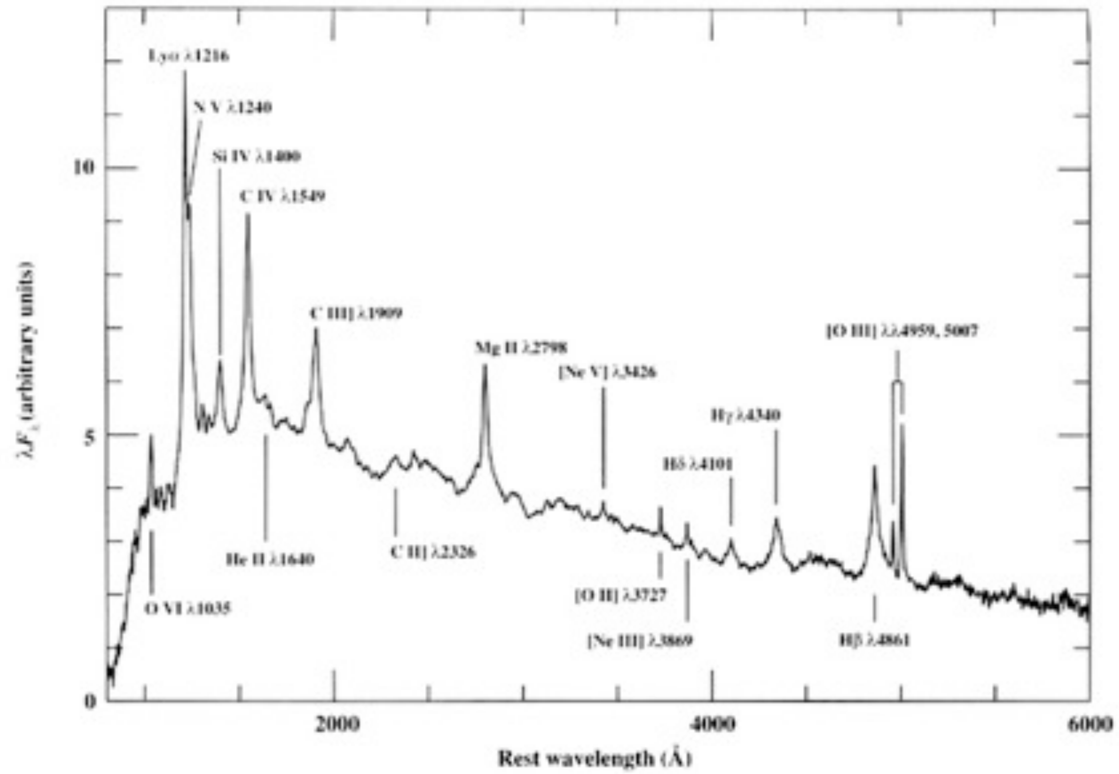
$$\chi^2(z,T,A) = \sum_{f=1}^{N_f} \left(\frac{F^f_{\text{obs}} - A \times F^f_{\text{pred}}(z,T)}{\sigma^f_{\text{obs}}} \right)^2$$

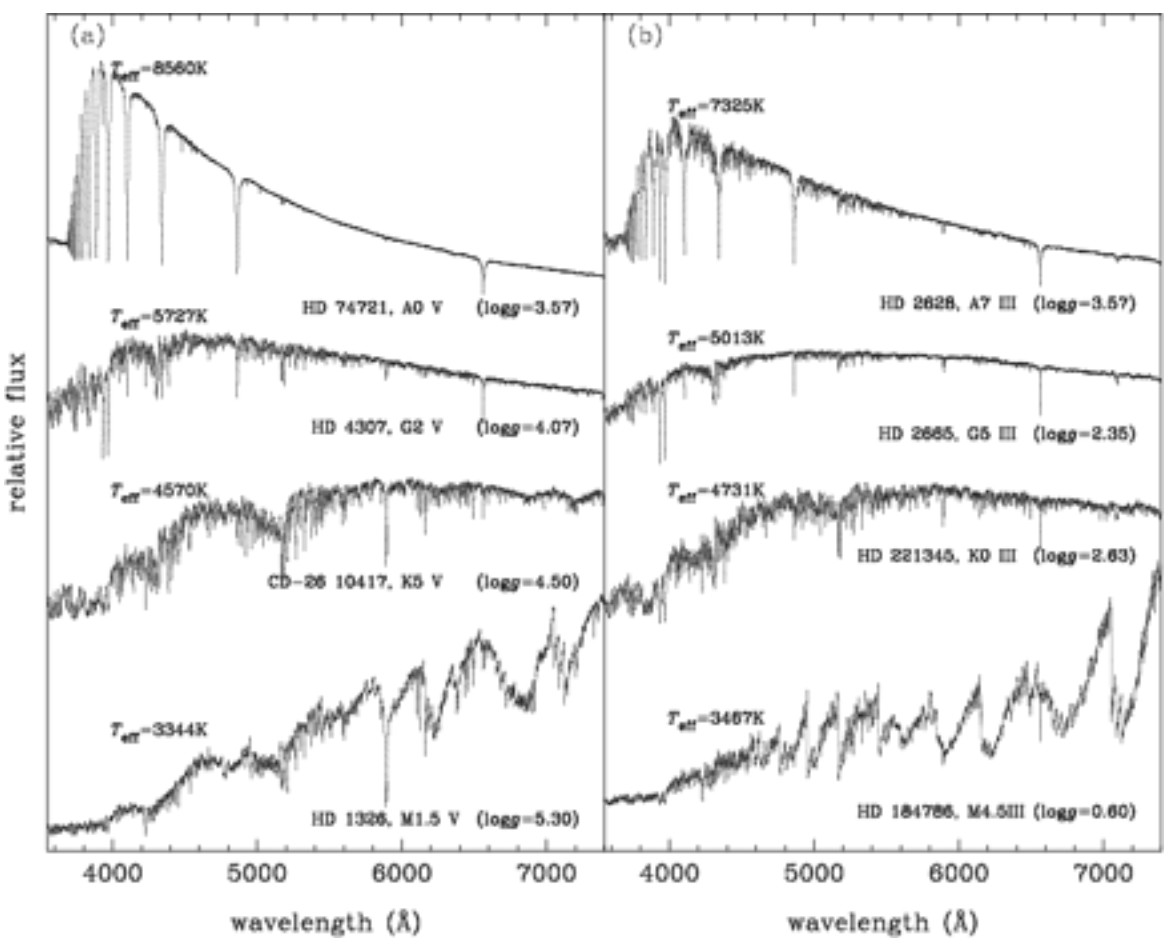
Minimize χ^2 to get best-fitting
redshift and template

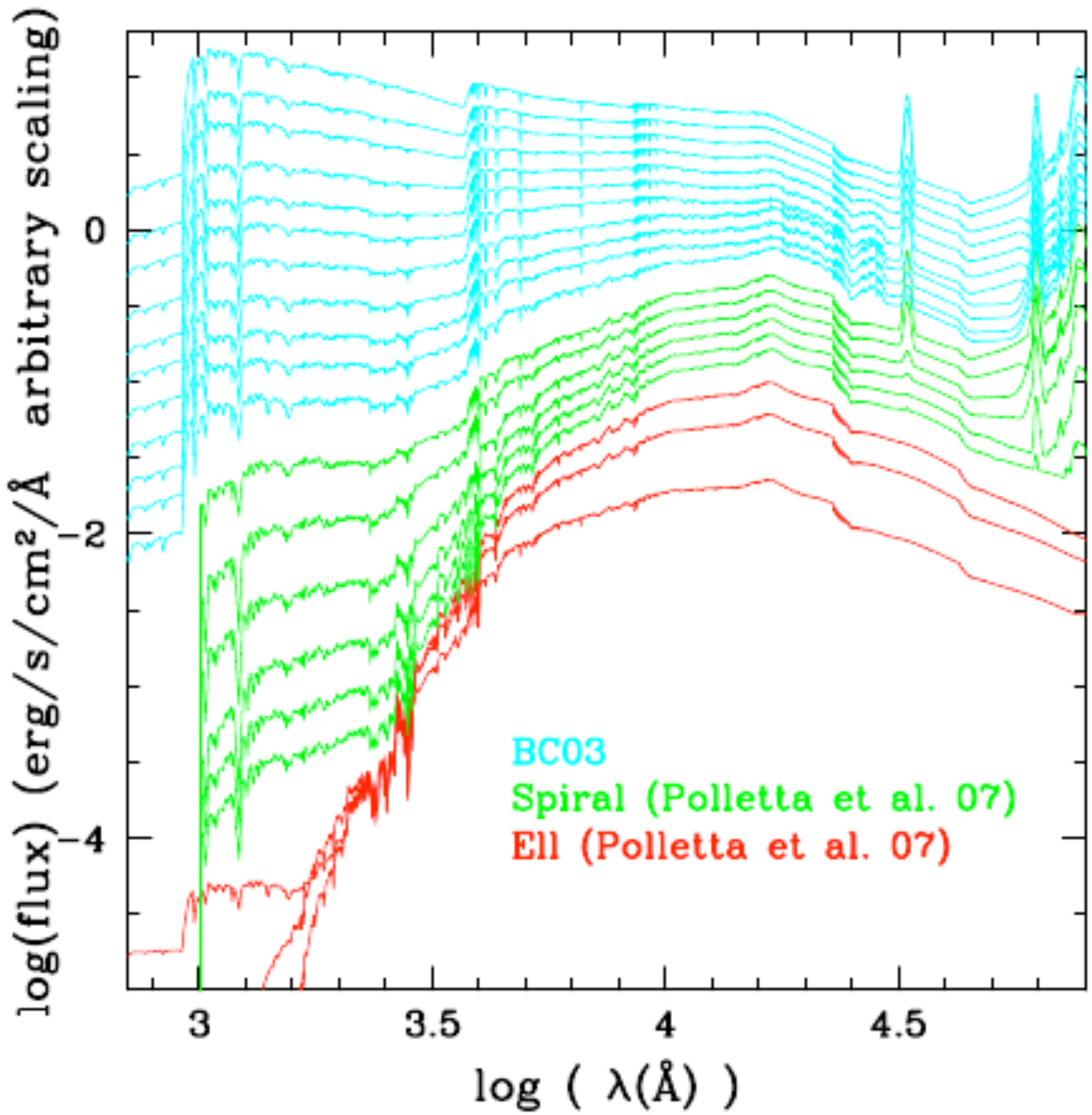


Le PHARE, Arnoux & Ilbert

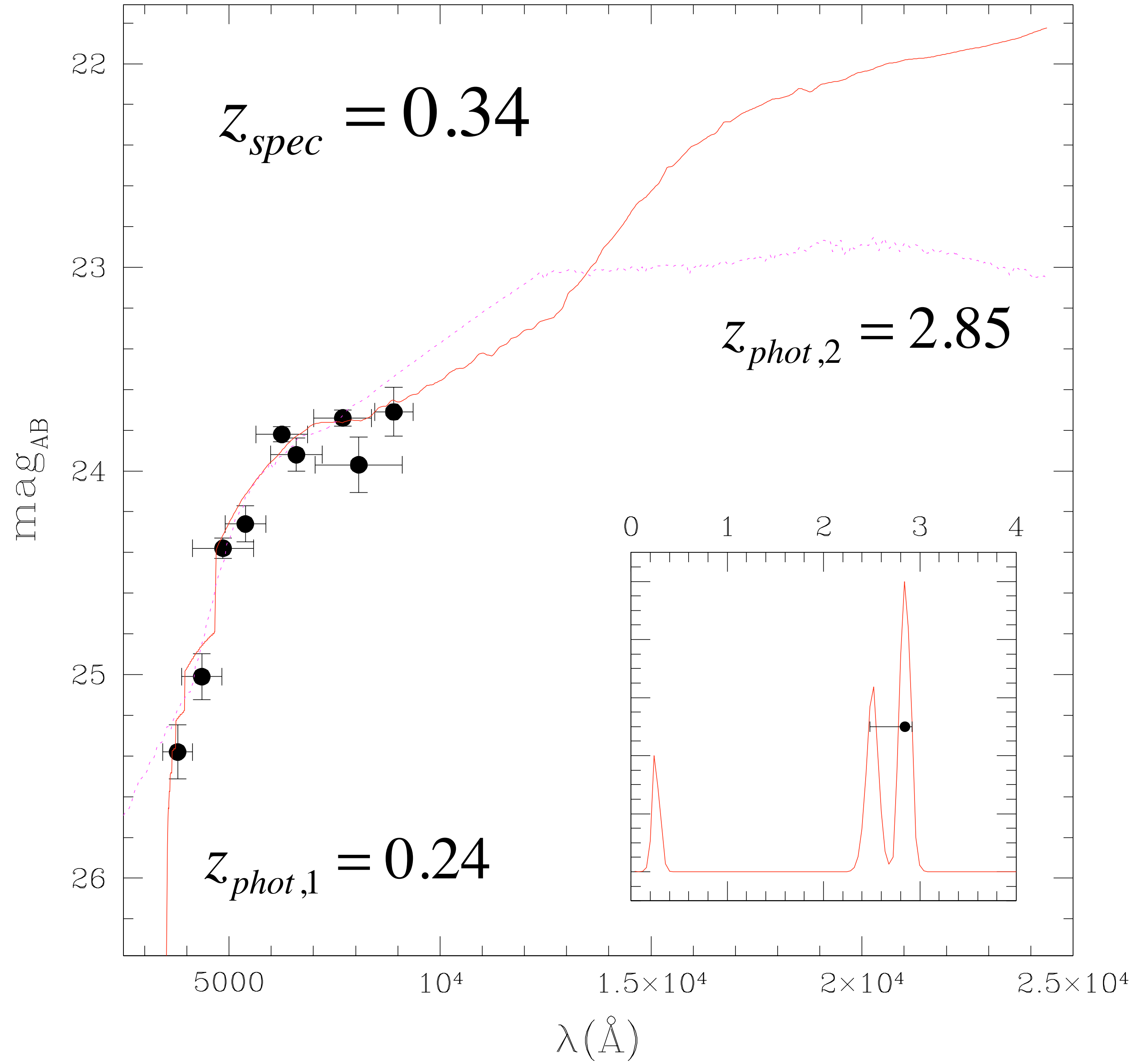
Template library







Degenaracies



Balmer or Lyα break?

Solution:
more filters and
probability
distribution

Ilbert et al. 2006, A&A, 457, 841

Bayesian z estimation

With one template T, color vector C

and object magnitude m_0 :

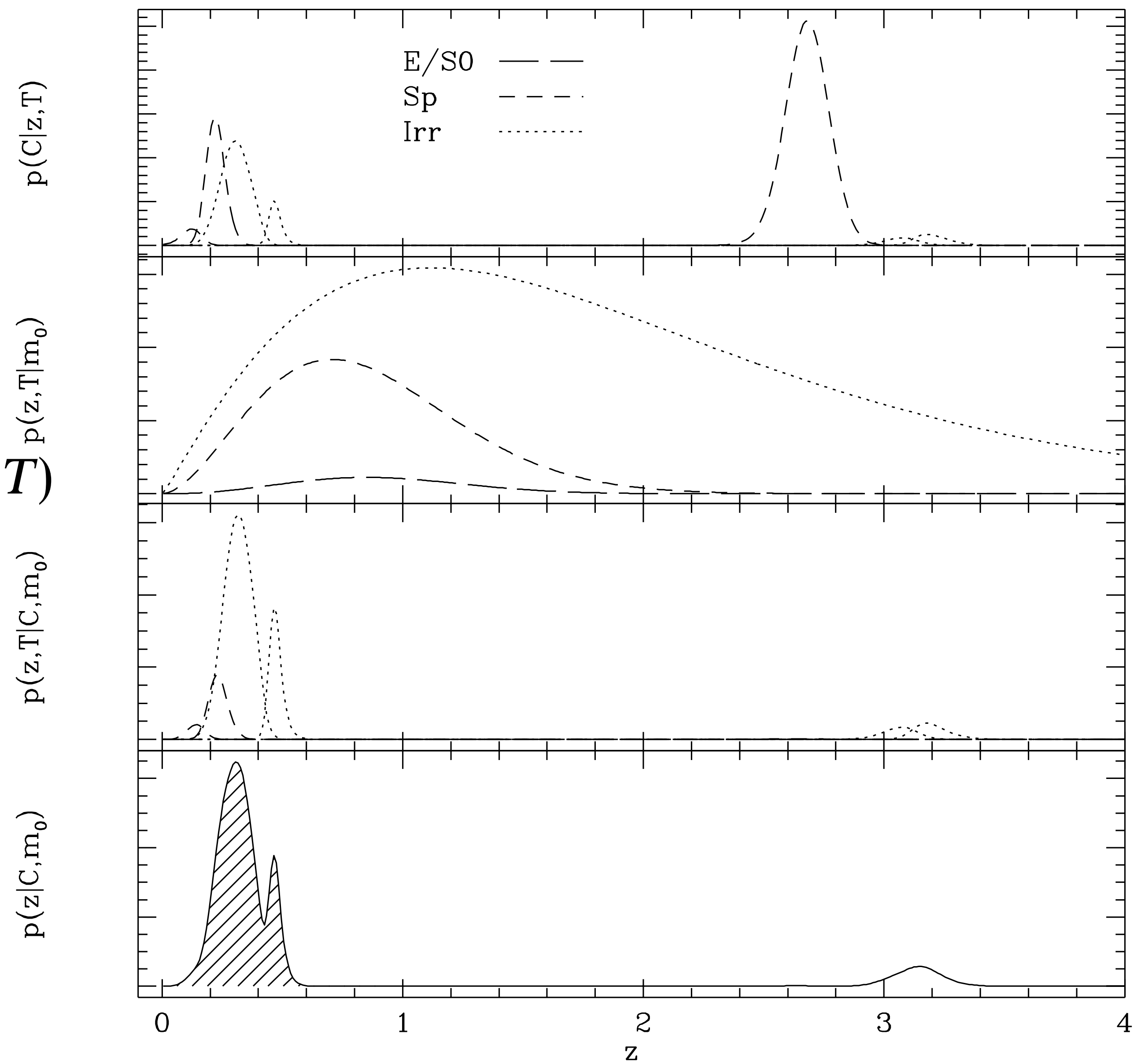
$$p(z \mid C, m_0) = \frac{p(z \mid m_0) p(C \mid z)}{p(C)} \propto p(z \mid m_0) p(C \mid z)$$

With a set of templates:

$$p(z \mid C, m_0) = \sum_T p(z, T \mid C, m_0) \propto \sum_T p(z, T \mid m_0) p(C \mid z, T)$$

$$p(C \mid z, T) \propto F_{TT}(z)^{-1/2} \exp\left[-\frac{\chi^2(z, T, a_m)}{2}\right]$$

$$\chi^2(z, T, a) = \sum_\alpha \frac{(f_\alpha - a f_{T\alpha})^2}{\sigma_{f_\alpha}^2}$$



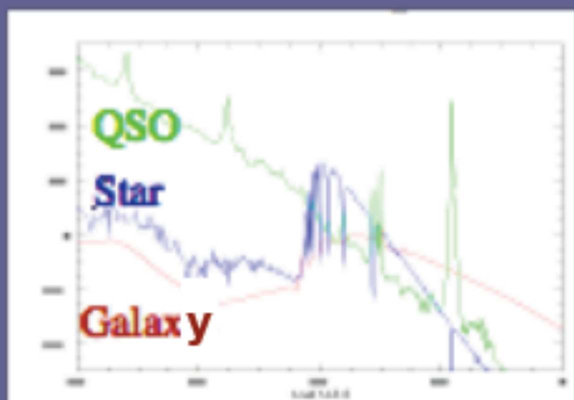
Priors (for example, VVDS redshift distribution):

$$p(z, T \mid m_0) = p(T \mid m_0) p(z \mid T, m_0)$$

Benitez, 2000, ApJ, 536,571

Template fitting with PDF

SED library

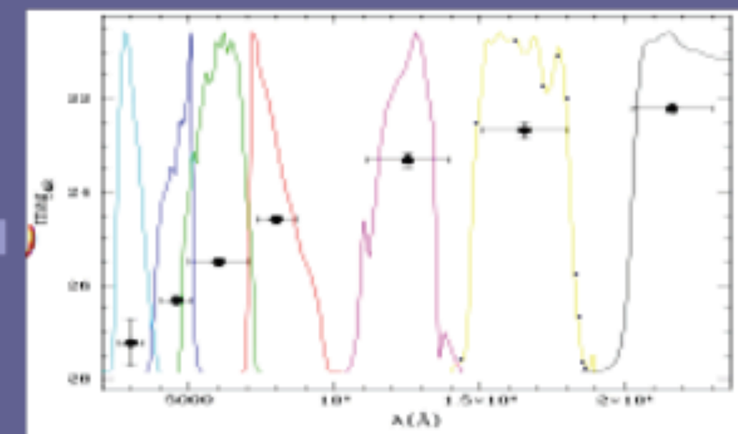


+Extinction
IGM

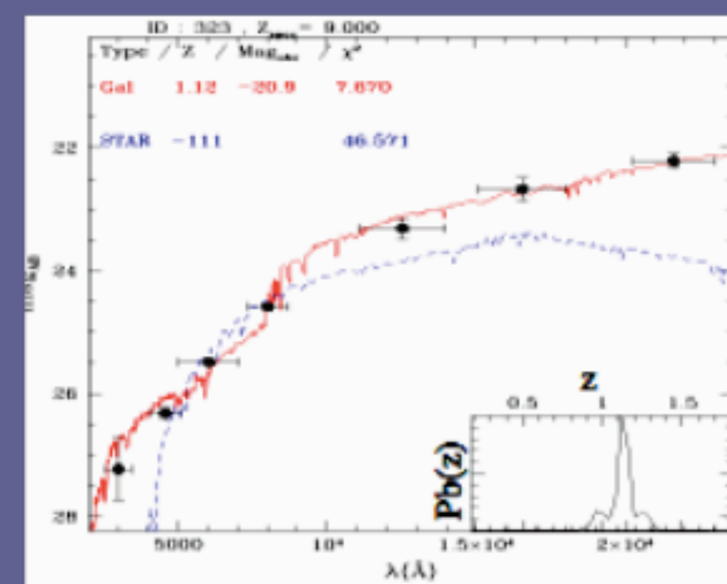
$$\chi^2(z,T,A)=\sum_{f=1}^{N_f}\left(\frac{F^f_{\mathrm{obs}}-A\times F^f_{\mathrm{pred}}(z,T)}{\sigma^f_{\mathrm{obs}}}\right)^2$$

Observed flux +error to theoretical flux

Set of filters



- Best redshift, second solution
- Returns a χ^2 for galaxy, star and QSO library
- Uncertainties, PDF(z)
- Type, E(B-V)
- Absolute magnitudes
- Physical parameters: mass,mean age,SFR,etc.



Le PHARE, Arnoux & Ilbert

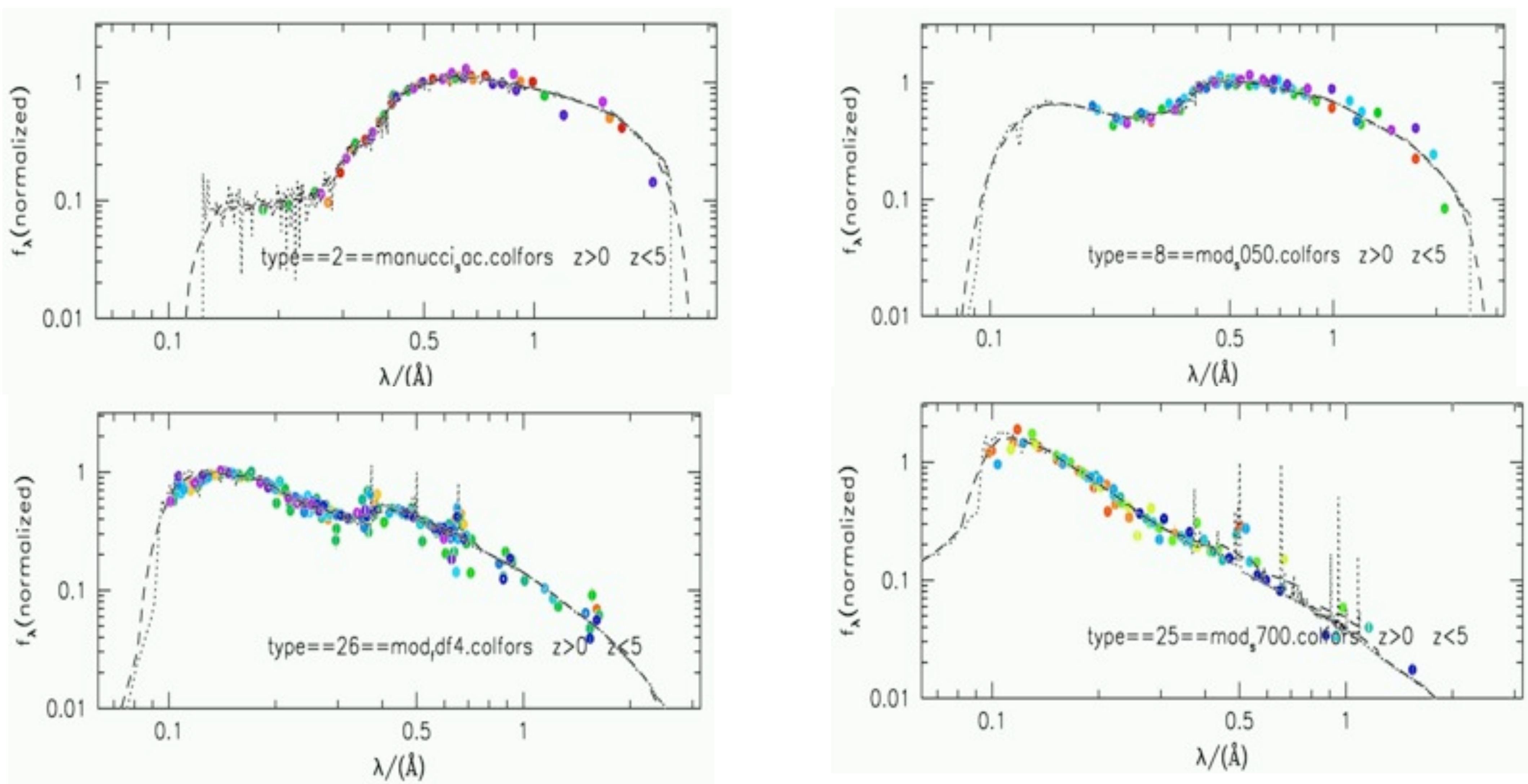
Some more references

- PHAT: Hildebrandt et al. 2010, A&A, 523, A31
- Bayesian studies:

  Wolf, 2009, MNRAS, 397, 520 on QSO
  Edmondson et al. 2006, MNRAS, 371, 1693, lensing
  Budavari, 2009, ApJ, 695, 747, general framework

PhotoZ

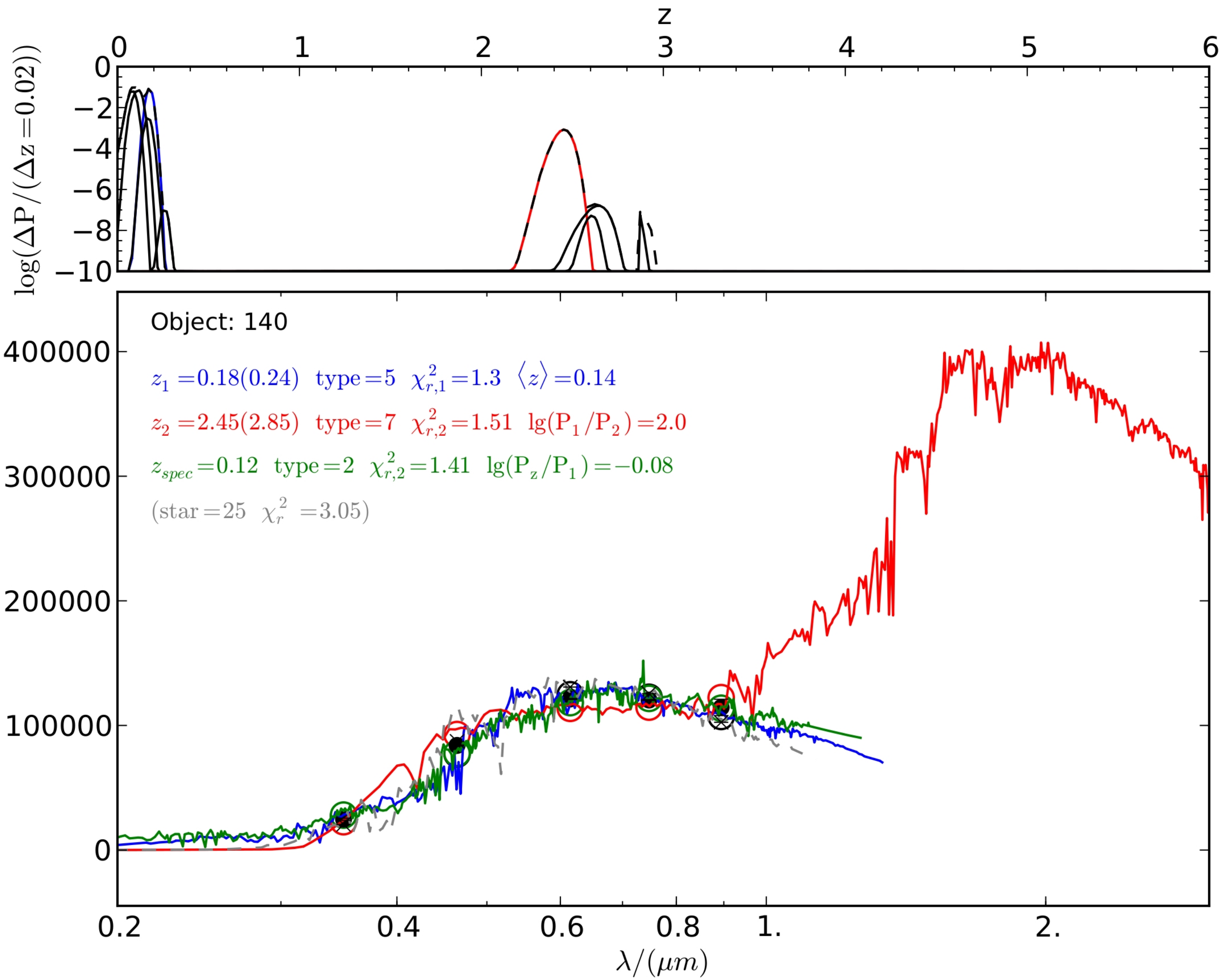
'Local' implementation of Bayesian photometric redshift fitting. Semi-empirical SEDs derived from broad-band fluxes of galaxies with spectroscopic z by fitting them with SEDs of Bruzual&Charlot, Maraston and spectra from FDF, Kinney&Calzetti, Mannucci. Stellar and QSO SEDs also fitted to check if the objects are really galaxies.



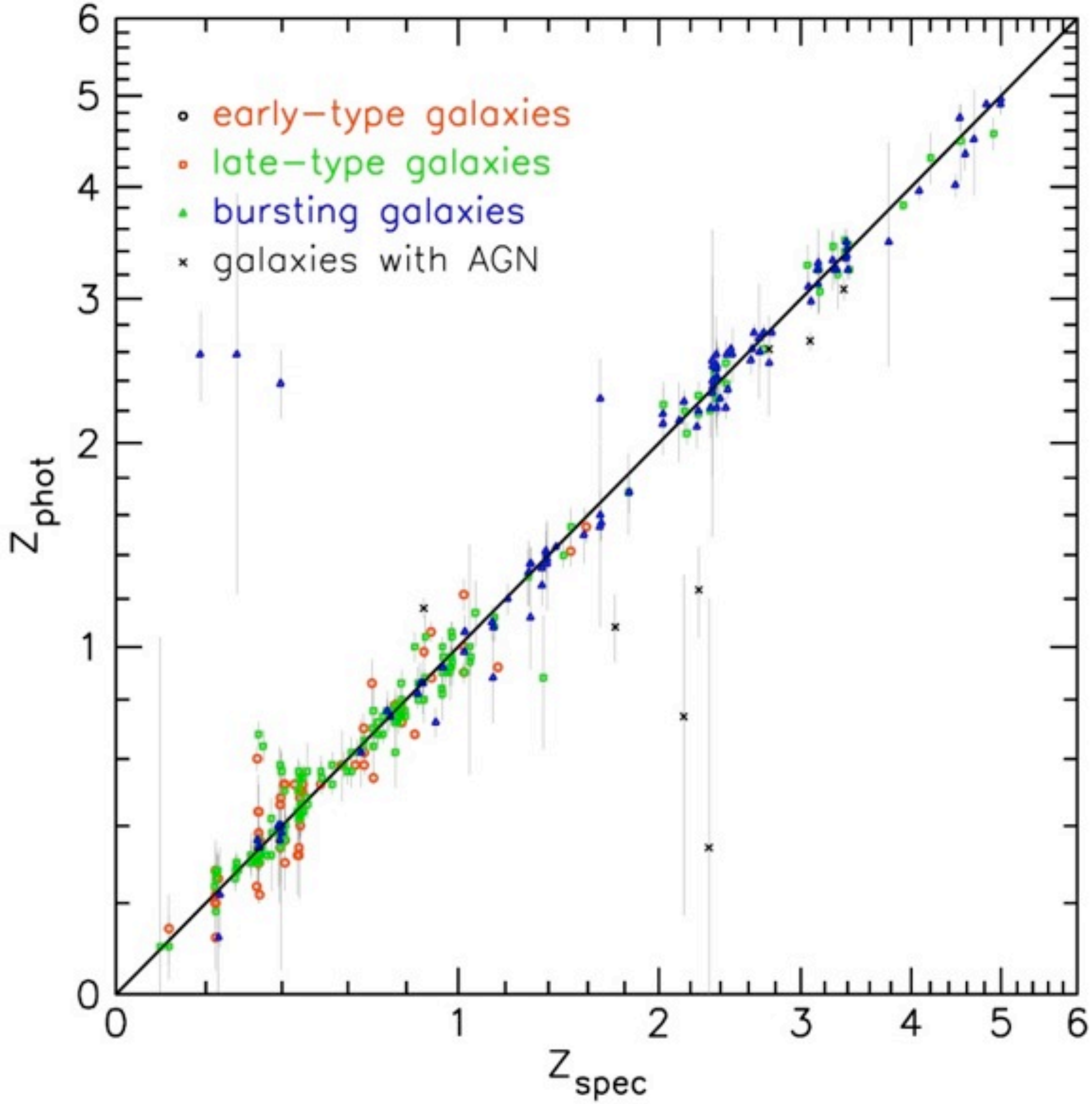
Bender et al. 2001

Proc. of the ESO/ECF/STScI Workshop "Deep Fields", S. Cristiani, A. Renzini, E. Williams Eds., Springer, p. 96

PDF with PhotoZ



Fors Deep Field



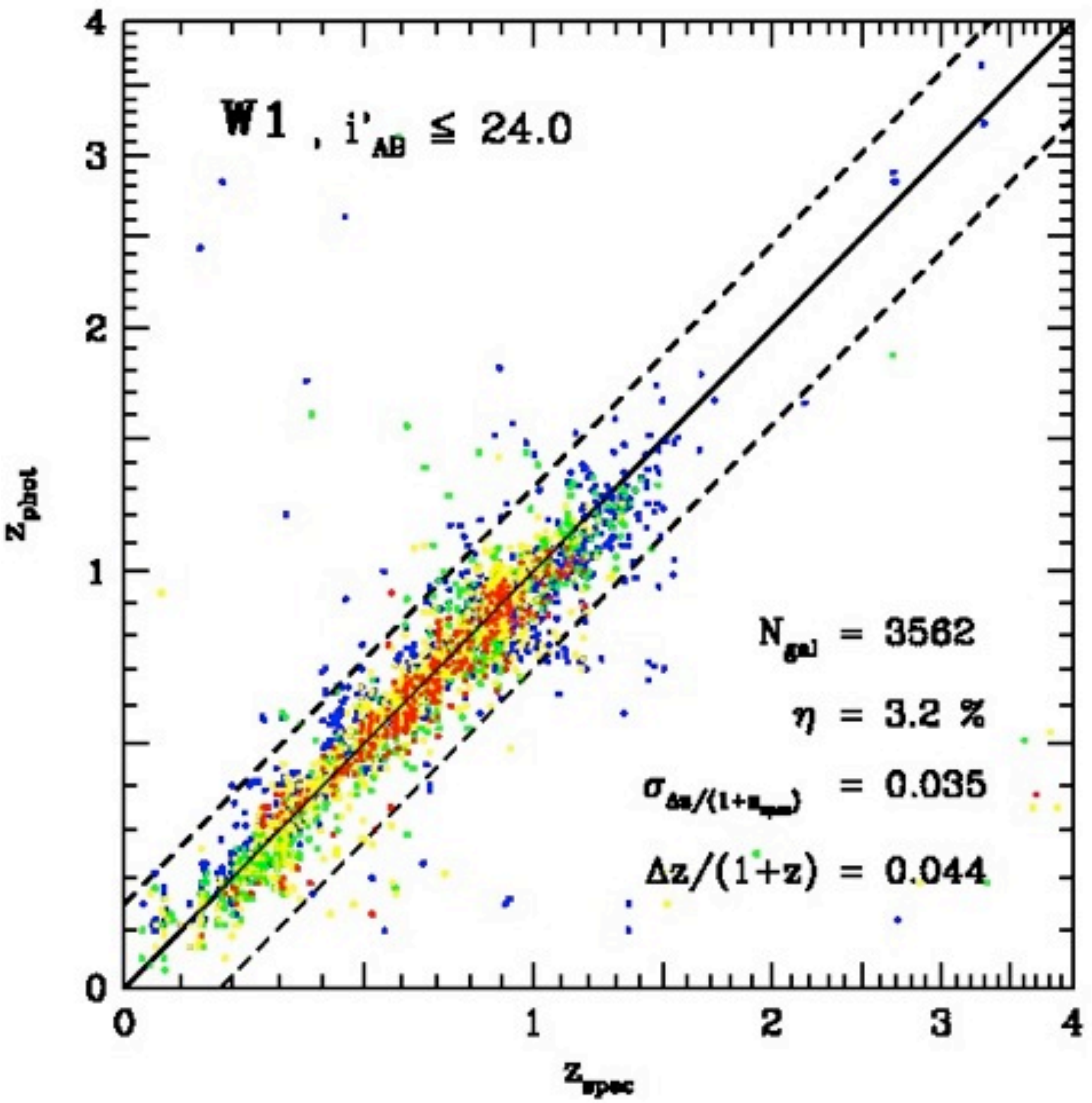
Gabasch et al. 2004, A&A, 421, 41

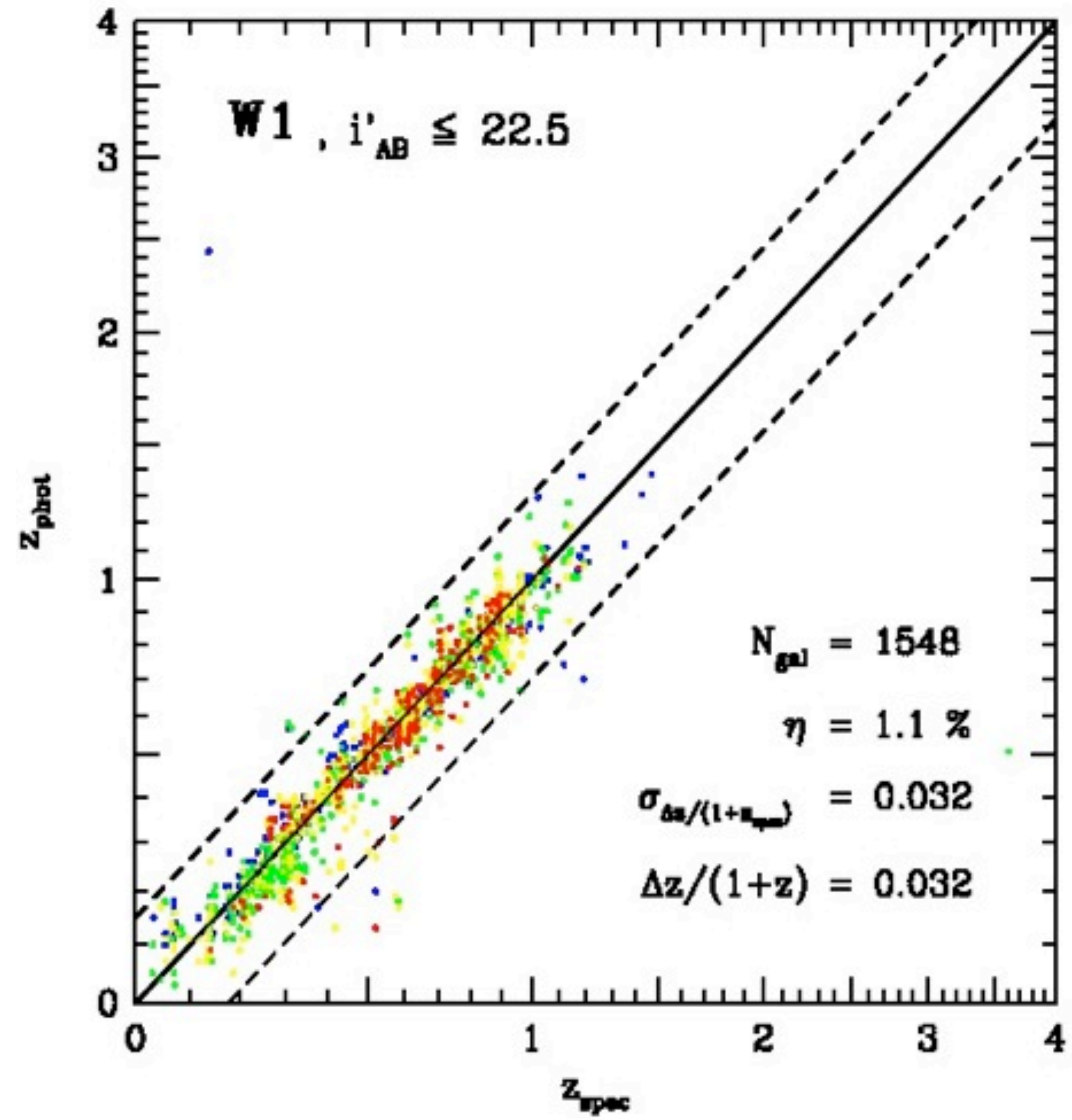
180 galaxies used to derive semi-empirical SEDs

180 galaxies in the control sample

Only ~ 1% catastrophic failures on normal galaxies! (mostly very blue, faint dwarf objects with almost power-law SEDs)

CFHT Legacy Survey



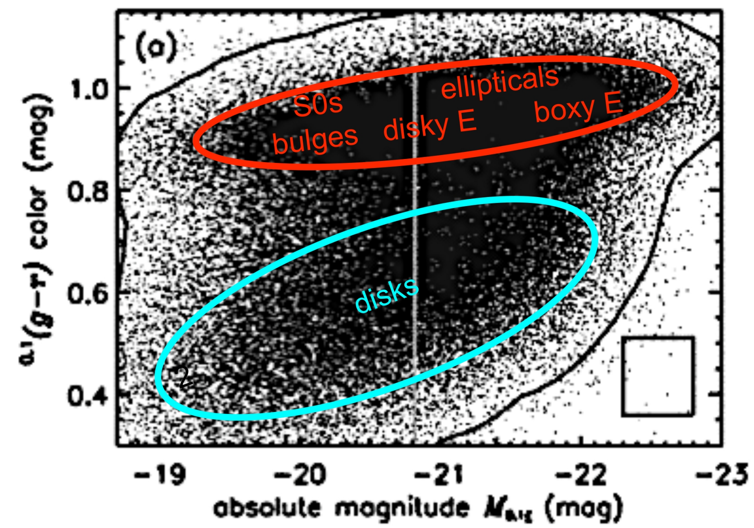


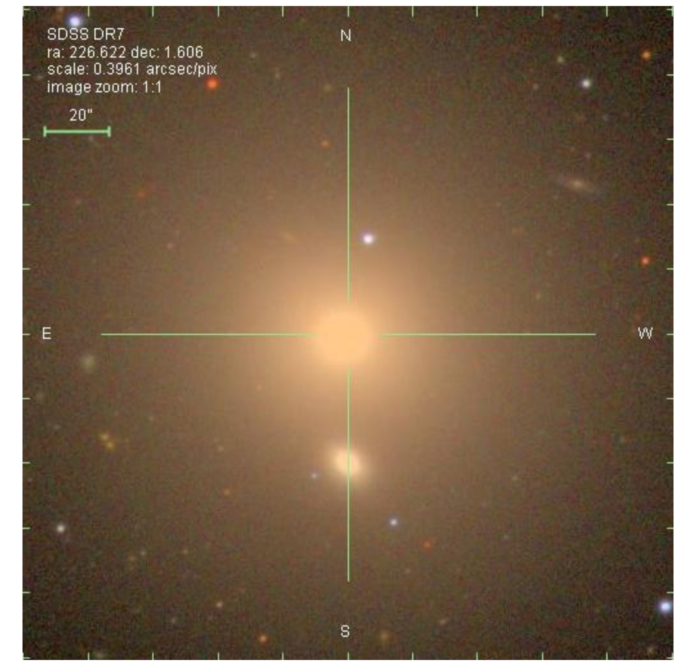
Brimioulle, Seitz et al., in prep.

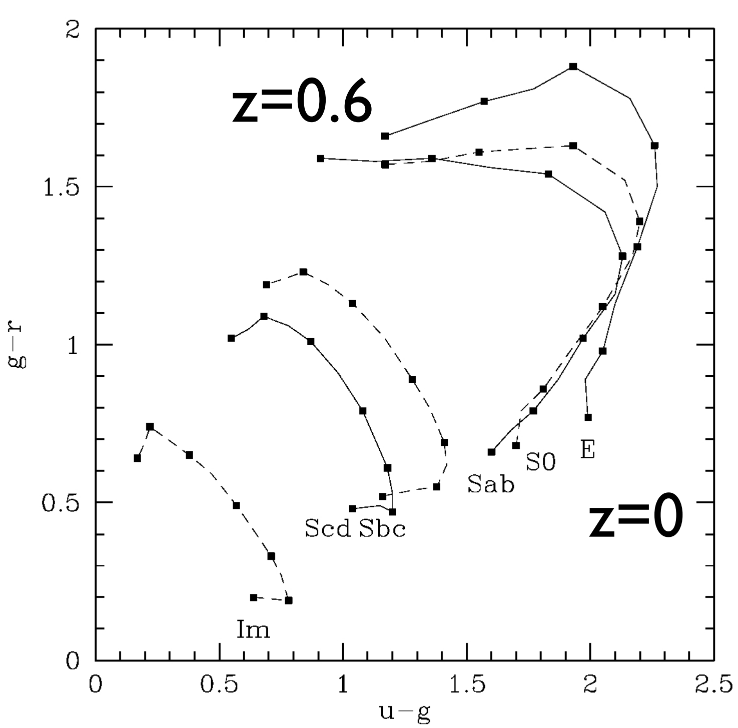
Luminous Red Galaxies

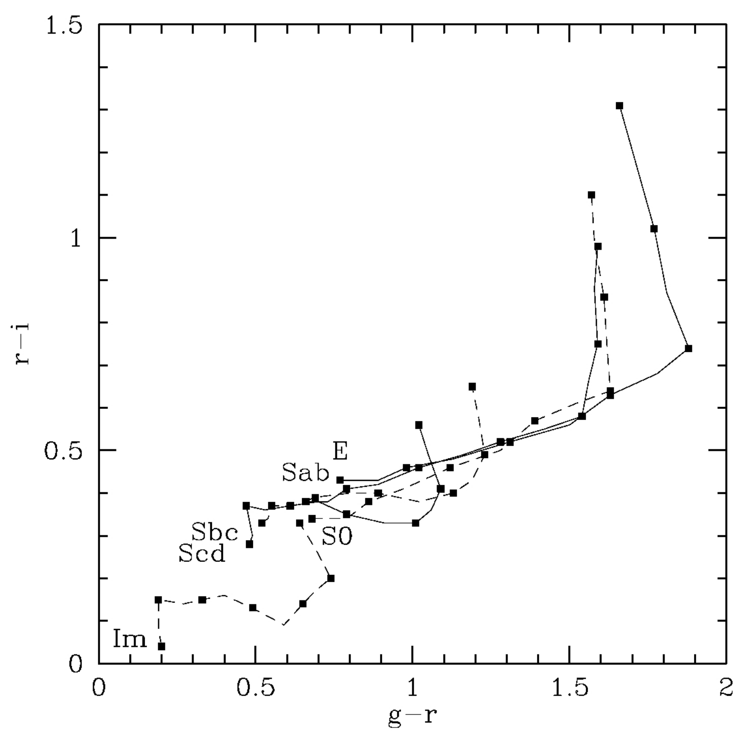
Eisenstein et al. 2001:
up to z~0.4

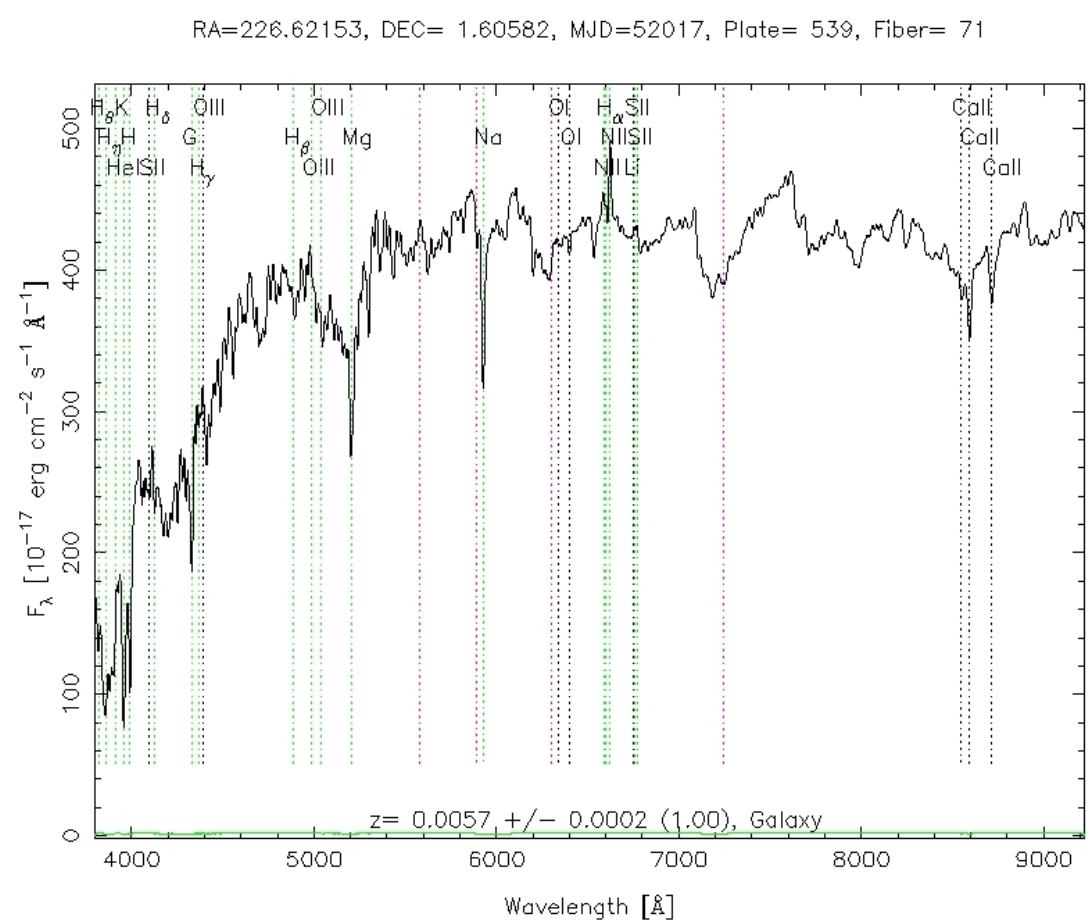
'BOSS' sample:
up to z~0.7







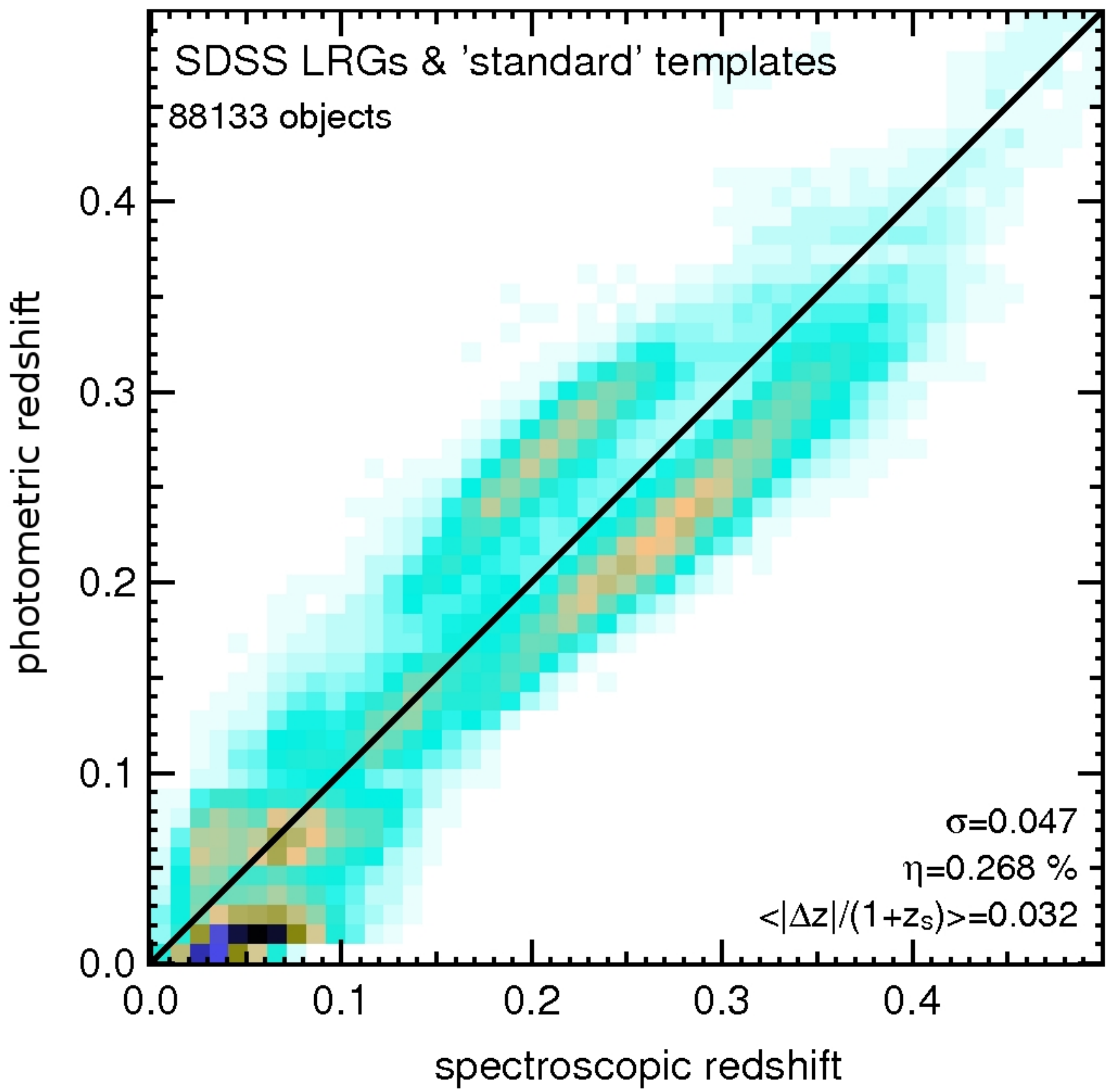




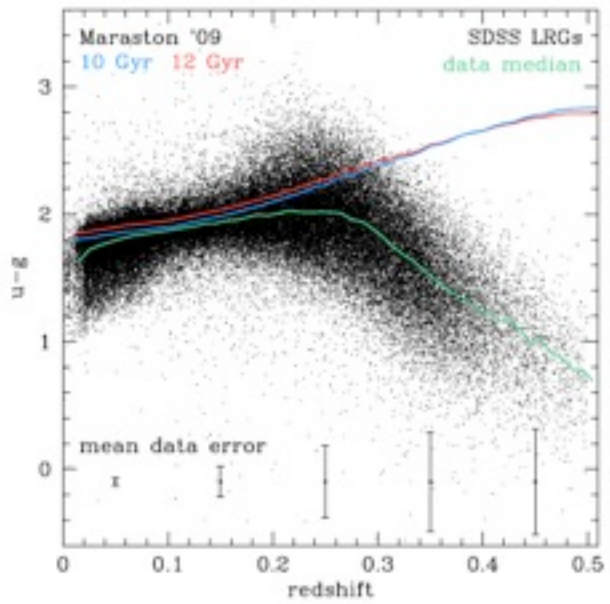
PhotoZ for LRZ

Master Thesis of Natascha Greisel

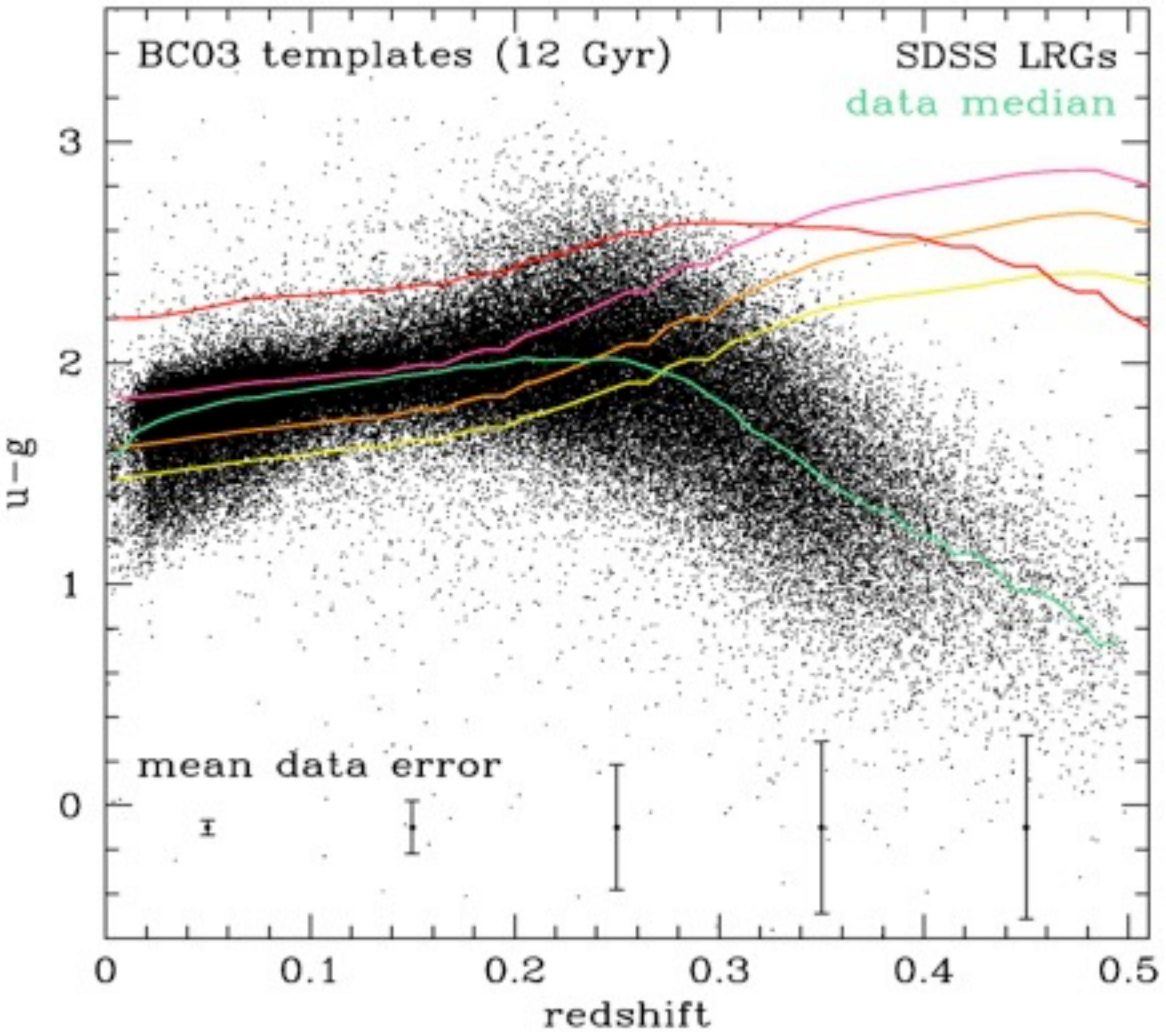
Based on ~90000 SDSS LRG

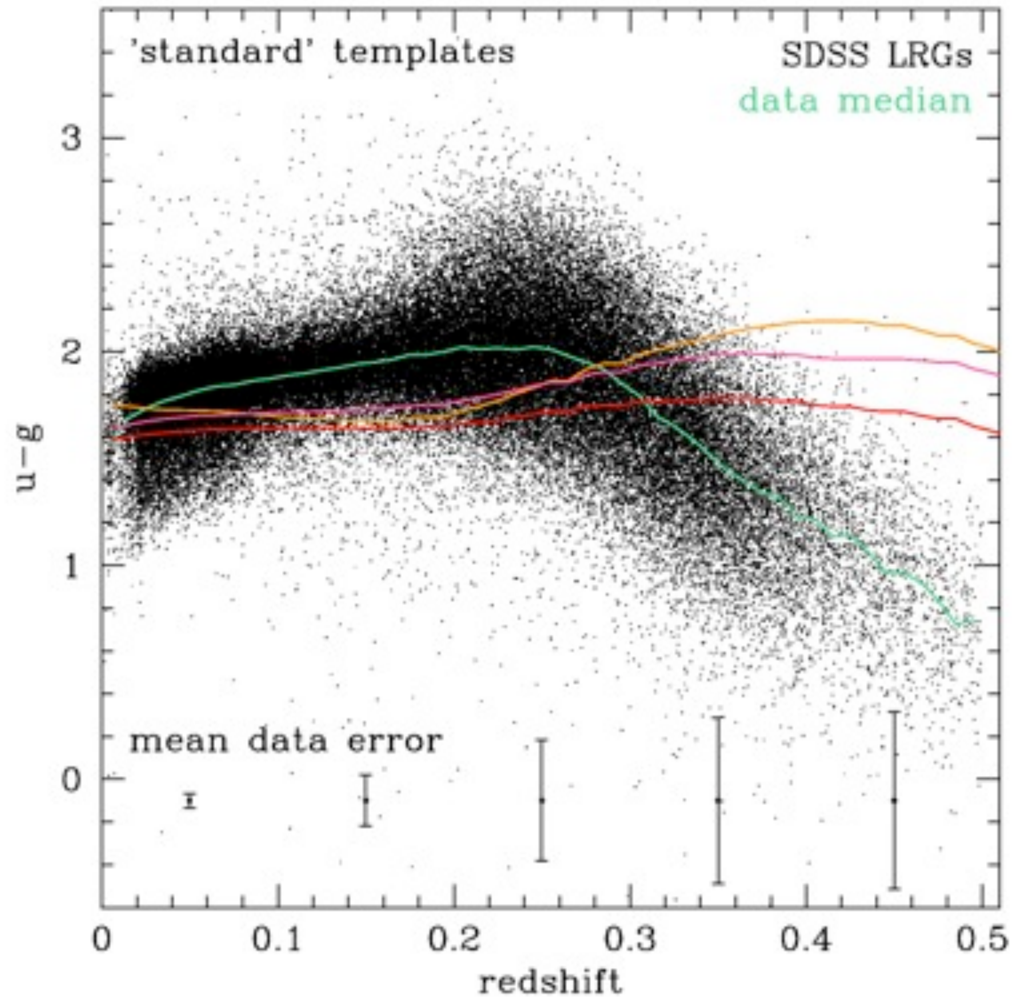


Greisel, Seitz, Bender et al., in prep.

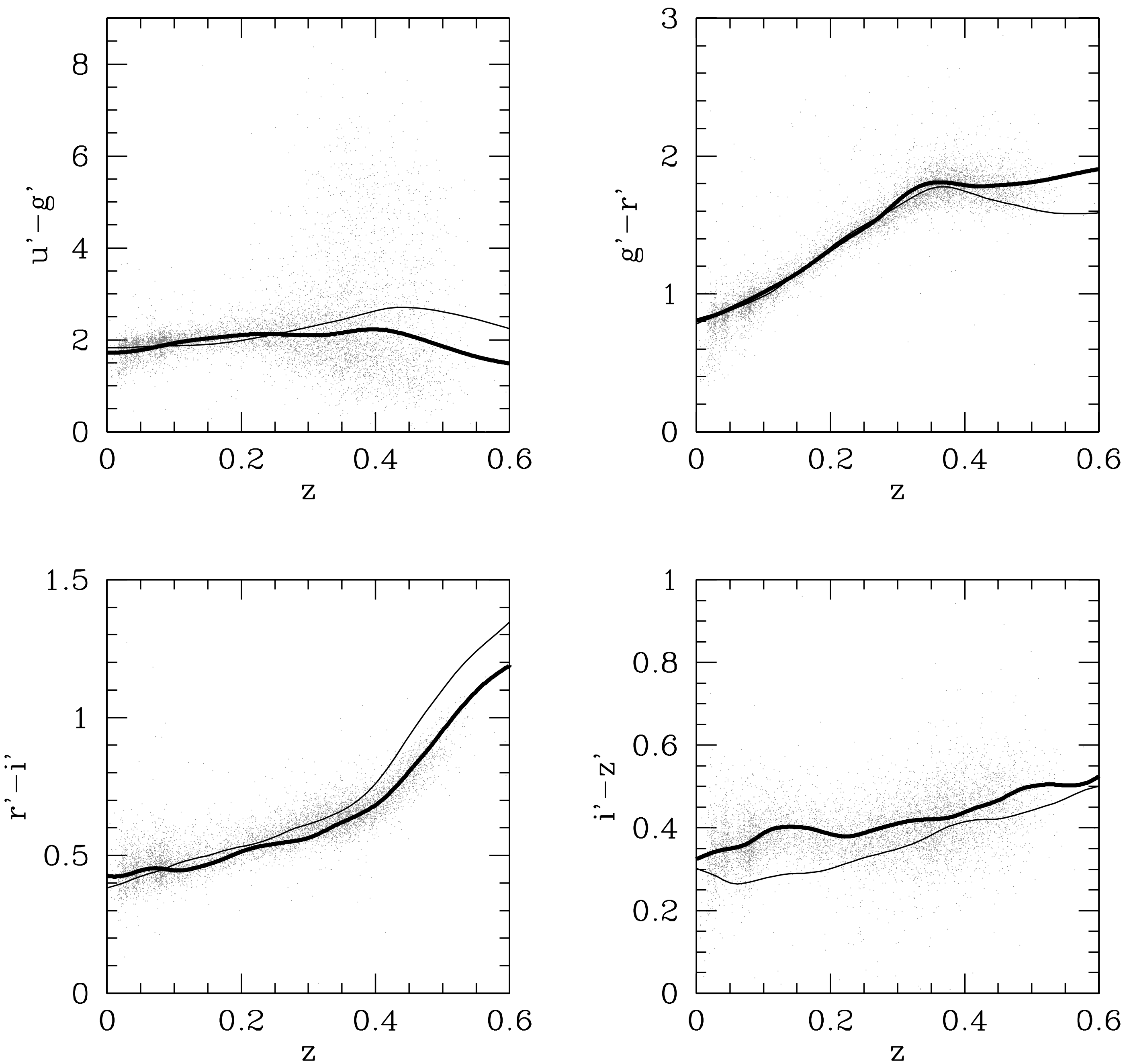


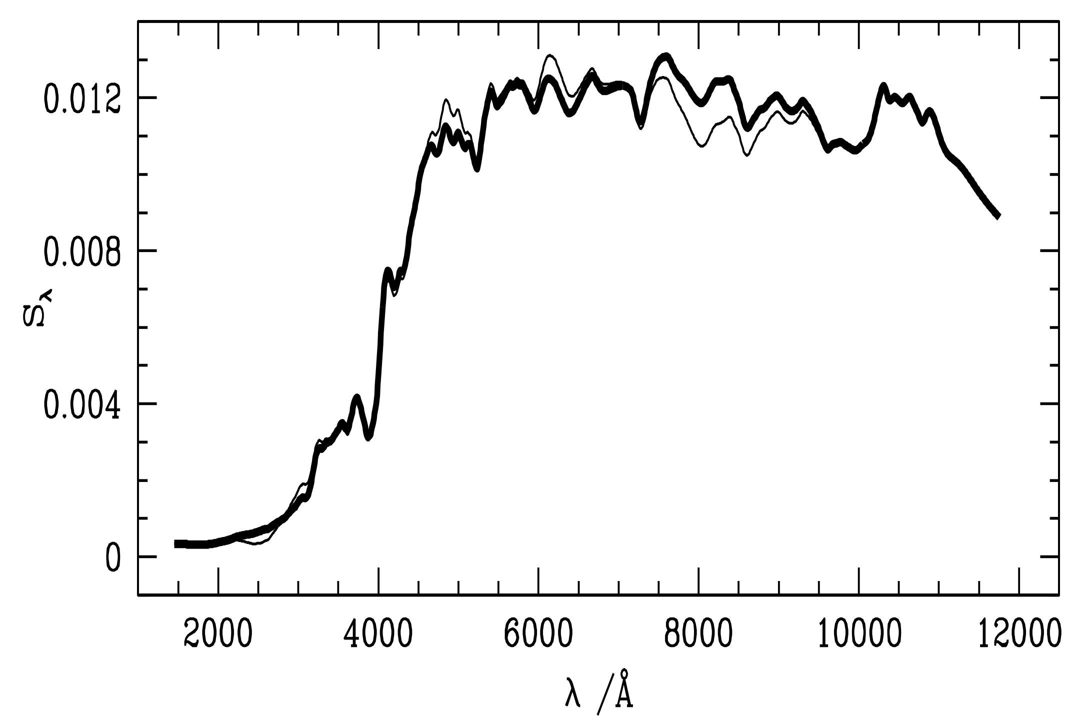
Missing SEDs for red galaxies





Empirical Template ‘correction’



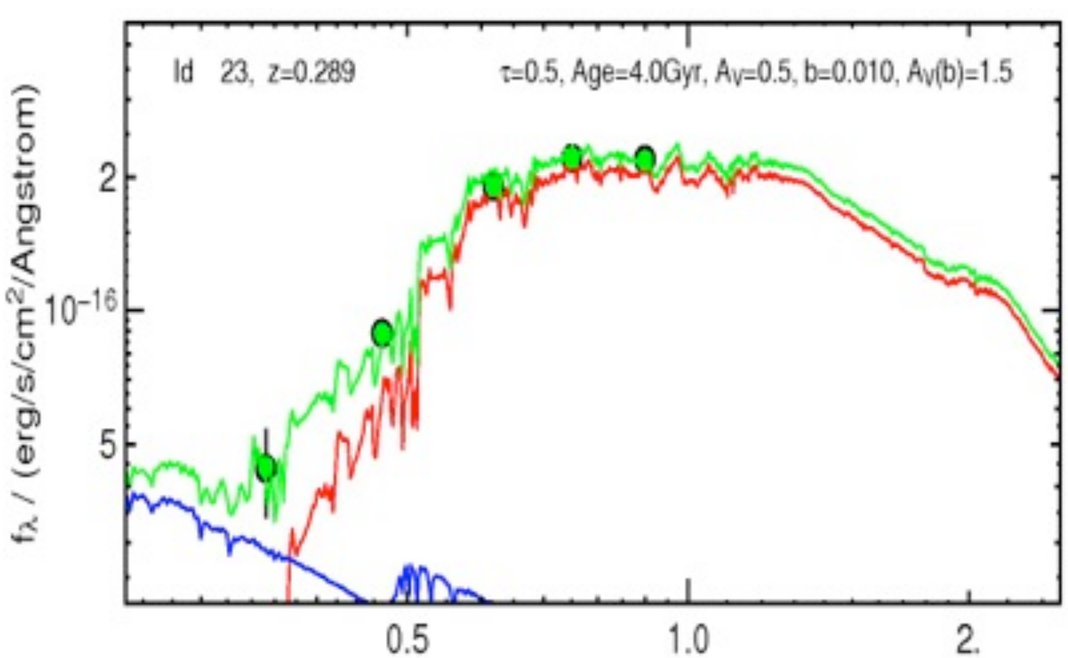


Csabai et al. 2003, ApJ, 125, 580

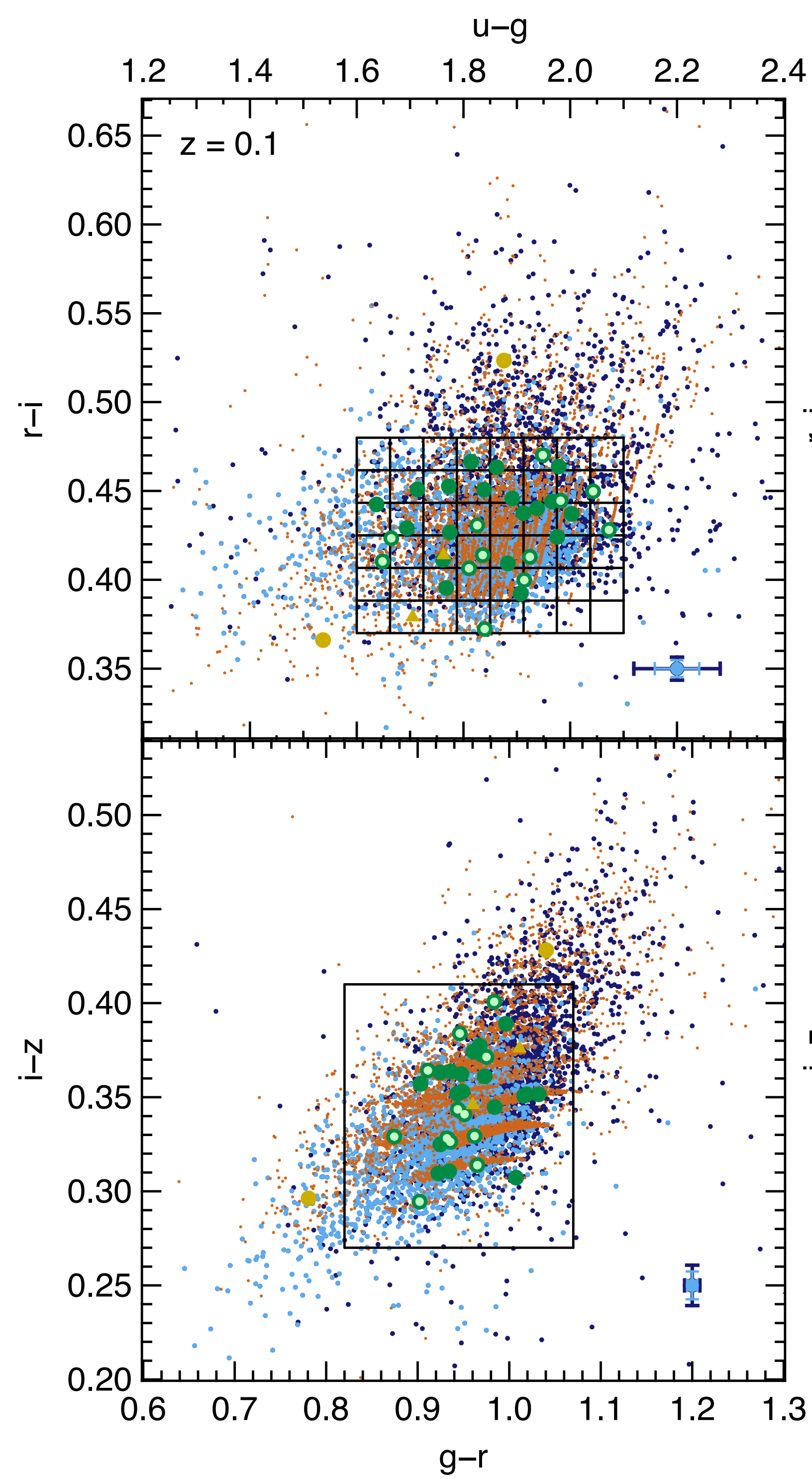
New SEDs to cover color space and z

SEDs constructed using the code of Drory et al. (2004, ApJ, 608, 742) so to sample the color space of LRG as a function of z SEDs are a composition of model SED (Bruzual & Charlot 2003, MNRAS, 344, 1000) and burst spectrum:

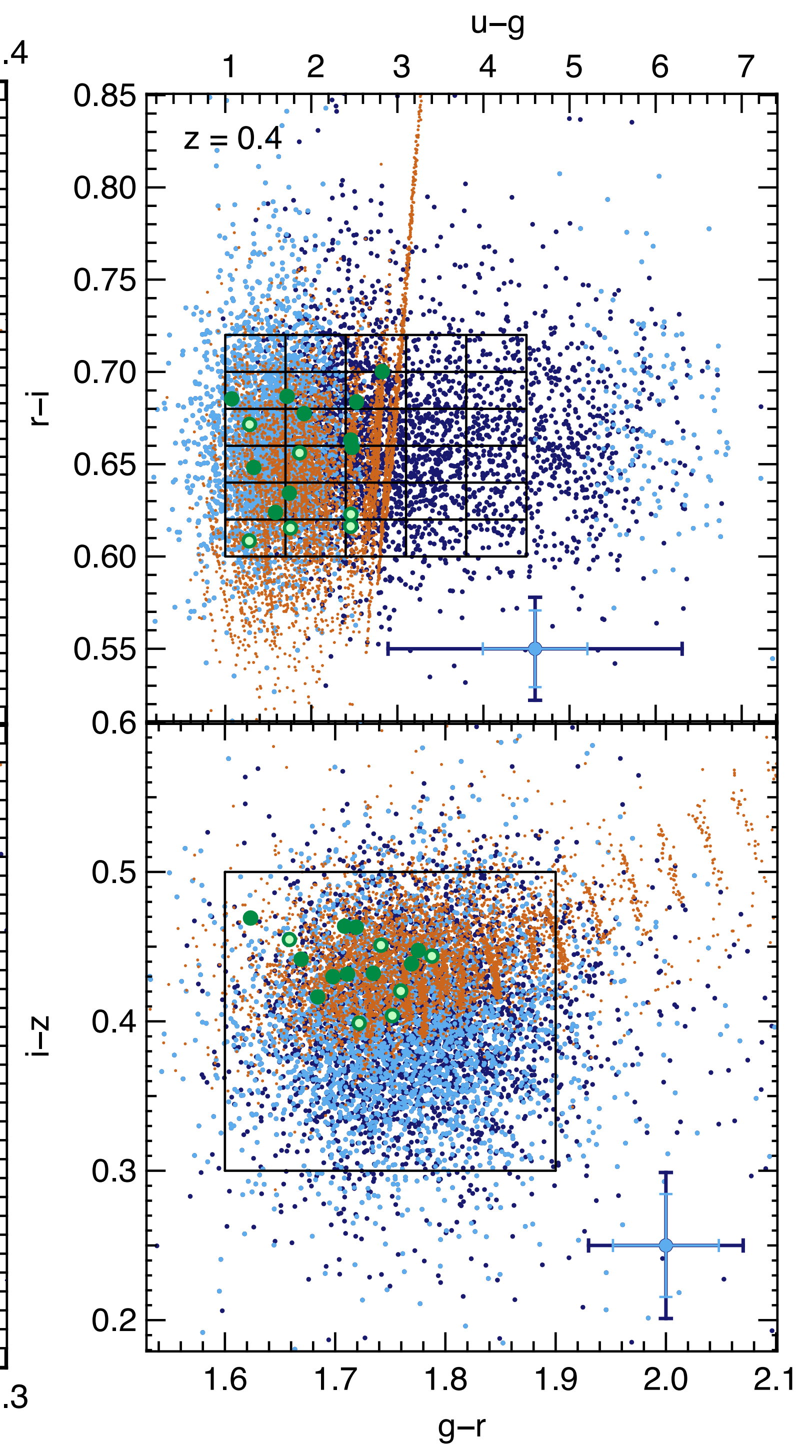
$$SED = \alpha\left(SED_{\mathrm{mod}} + \beta\, SED_{burst}\right)$$



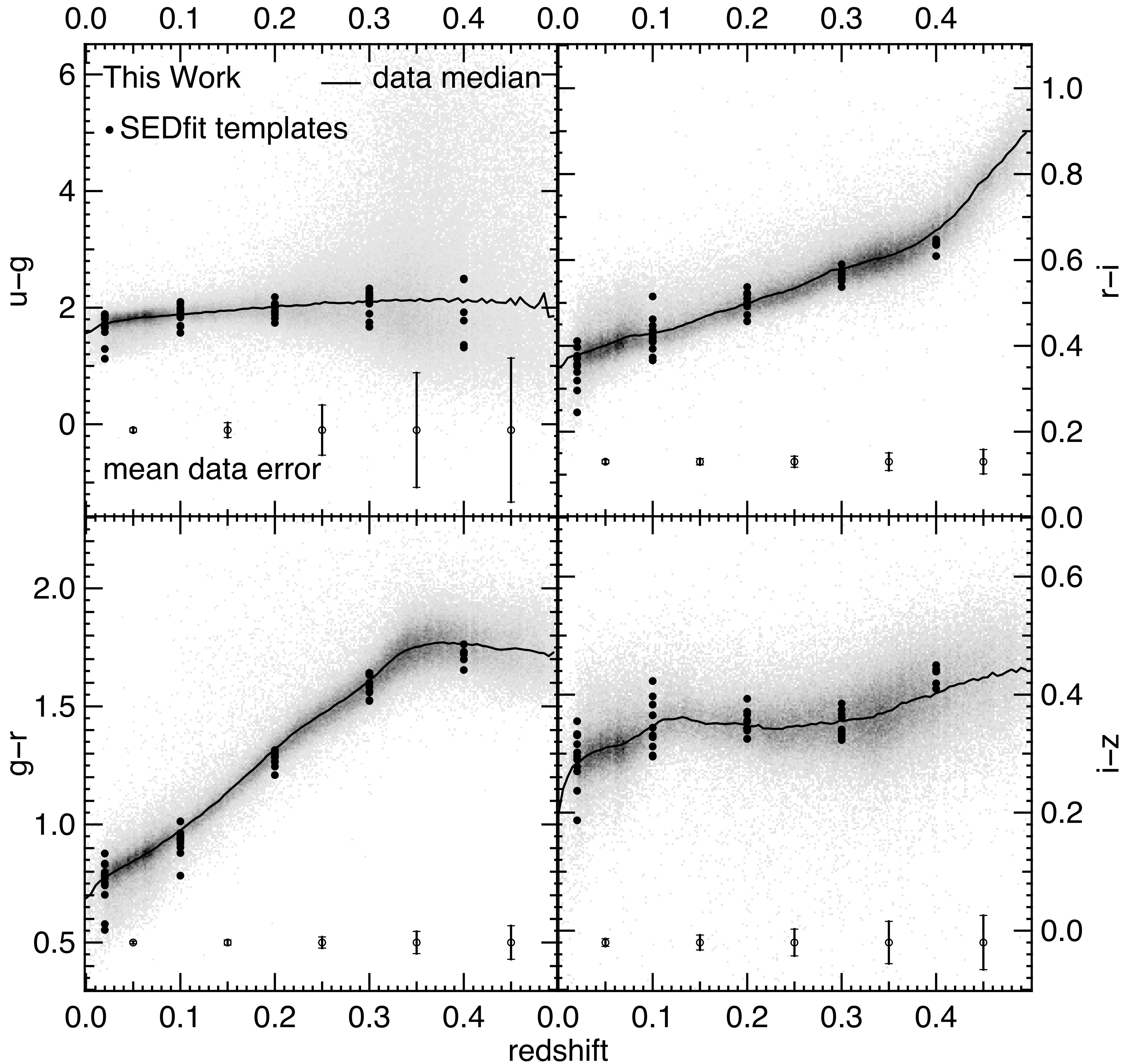
0.08<z<0.12



0.38<z<0.42



Color properties of optimized SEDs



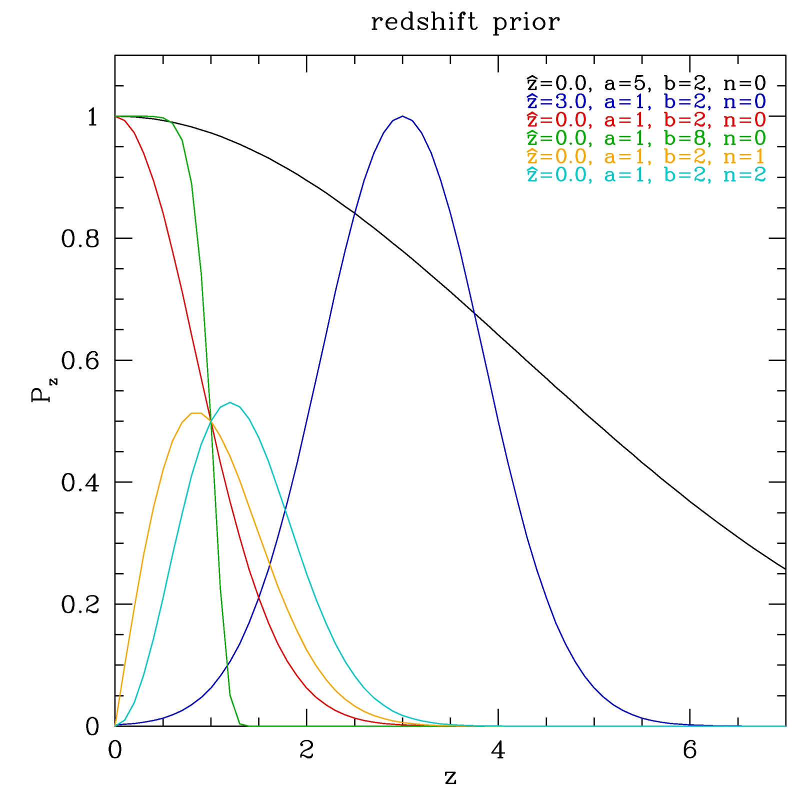
Redshift and magnitude priors

z prior:

$$P_{z,T} = z^n \exp\left[-\ln 2 \frac{(z - z_T)^b}{a^2} \right]$$

$$n = 0, b = 2 \ (Gaussian)$$

$$z_T = 0.02, 0.1, 0.2, 0.3, 0.4; a = 0.2$$

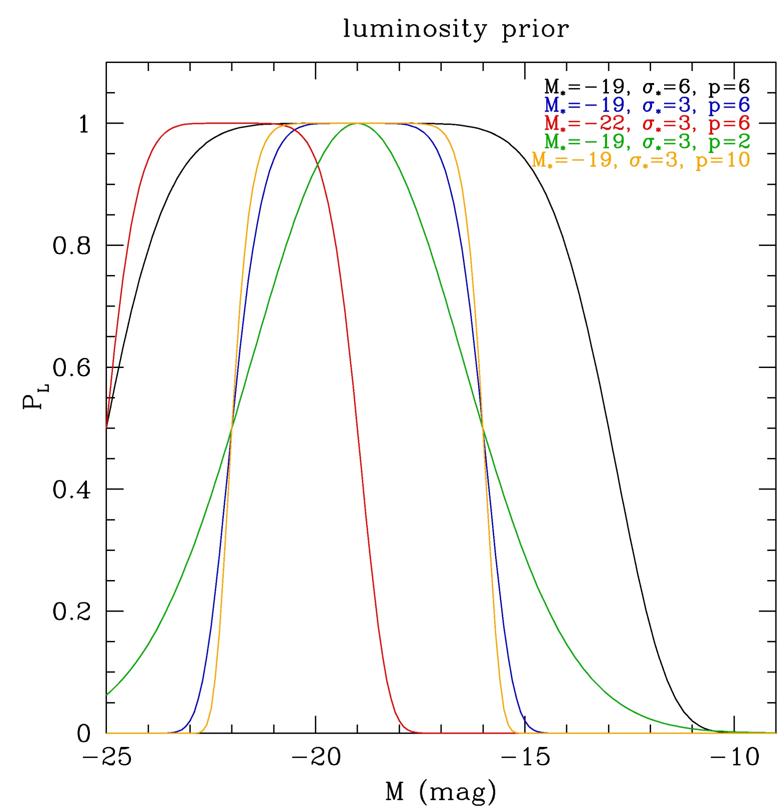


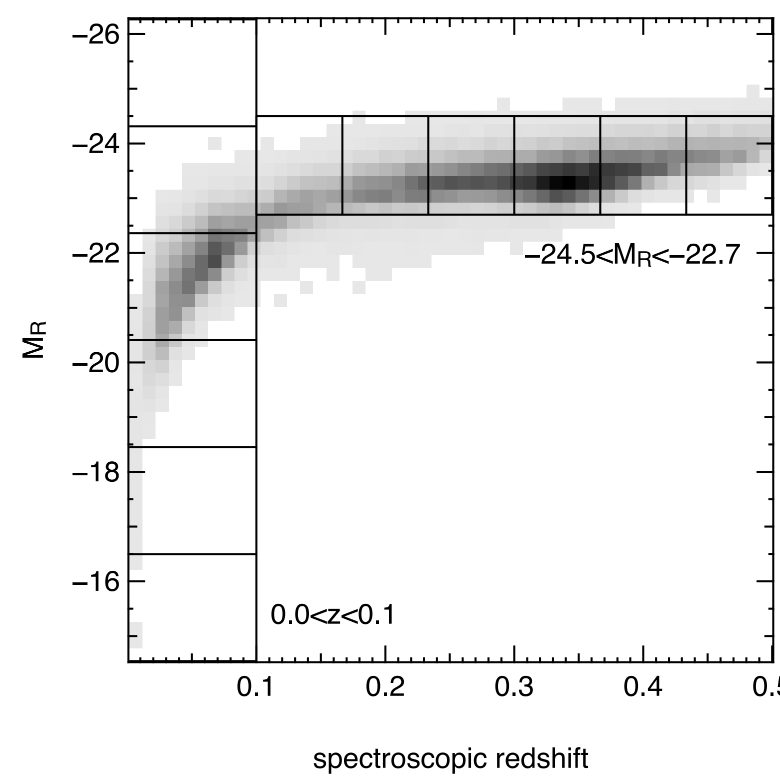
Luminosity prior:

$$P_L = \exp\left[-\ln 2 \frac{(M - M_*)^p}{\sigma_*^2} \right]$$

$$p = 6 \ (Flat\ top)$$

$$M_* = -21, \sigma_* = 3$$

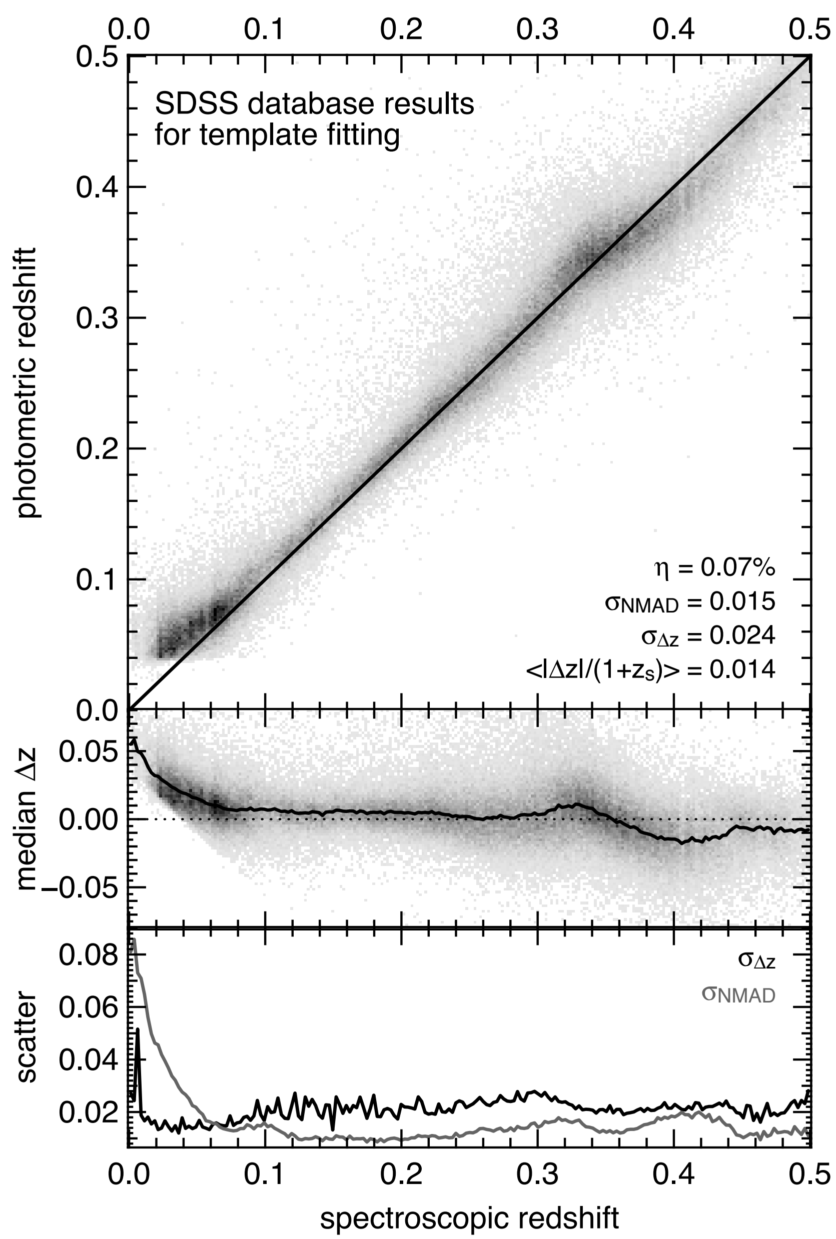


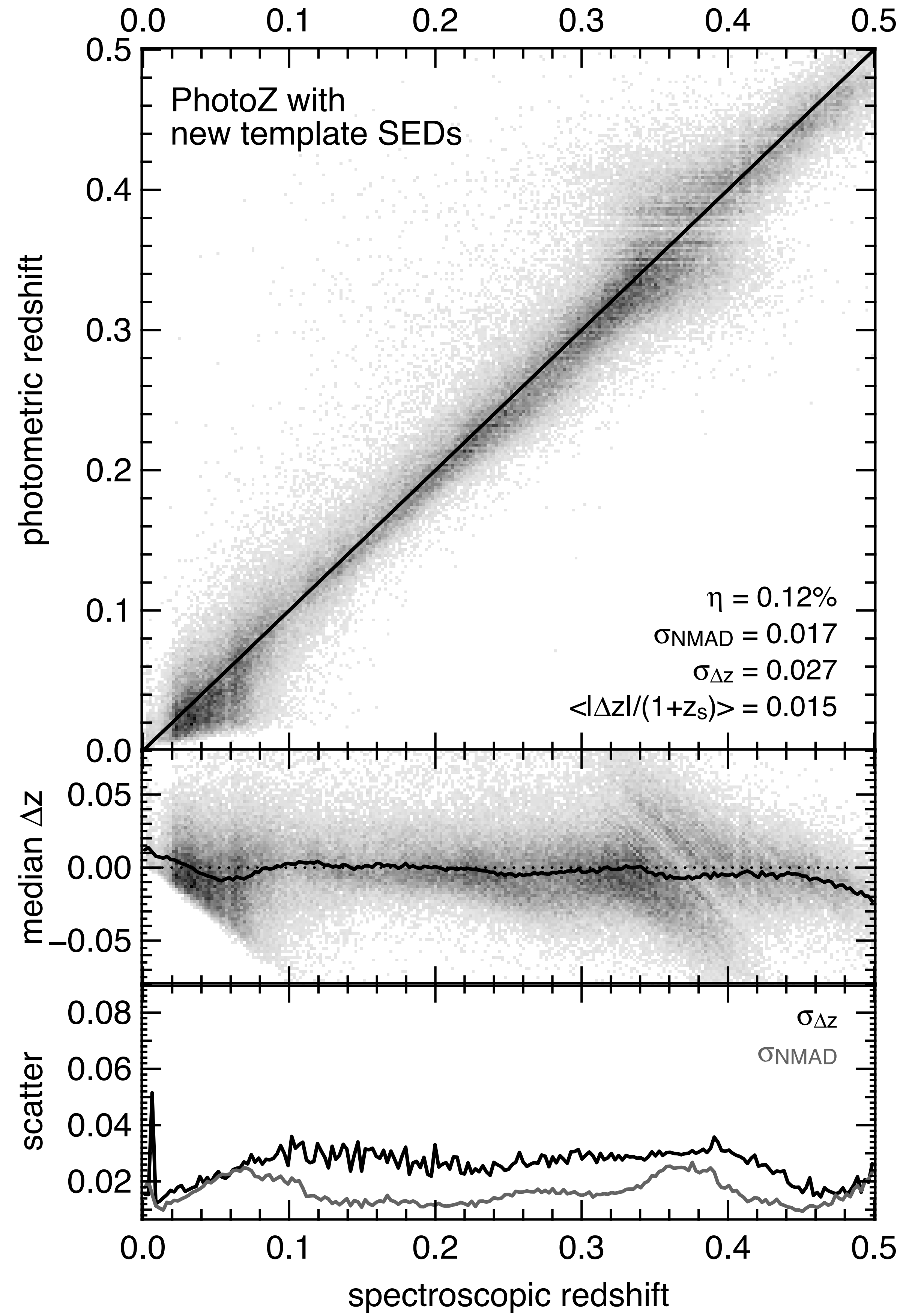


SLOAN LRGs with optimized SEDs

$$\eta : \frac{|\Delta z|}{1+z_{spec}} > 0.15$$

$$\sigma_{\Delta z/(1+z)} = 1.48 Median \left(\frac{|\Delta z|}{1+z_{spec}} \right)_{non-outliers}$$



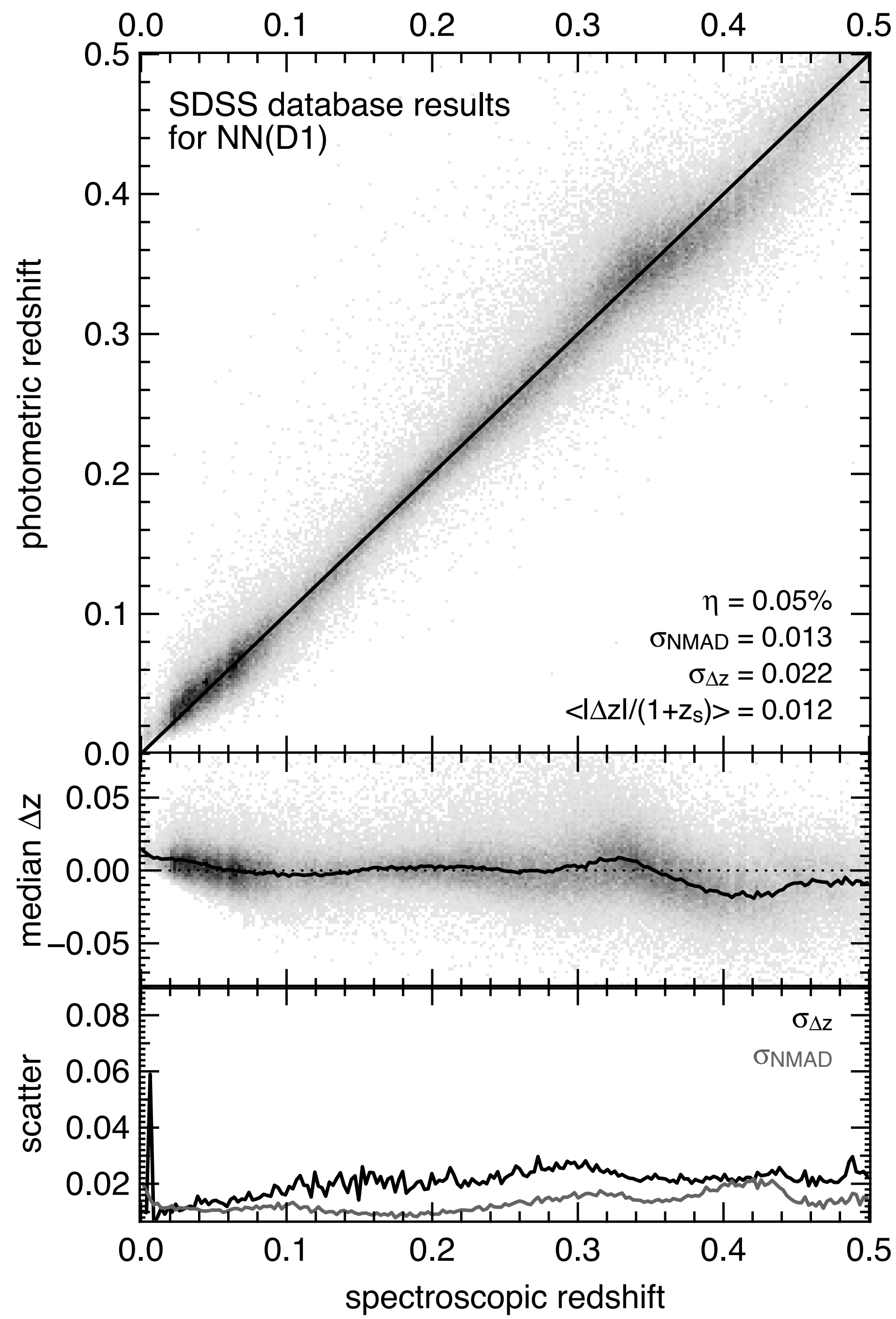


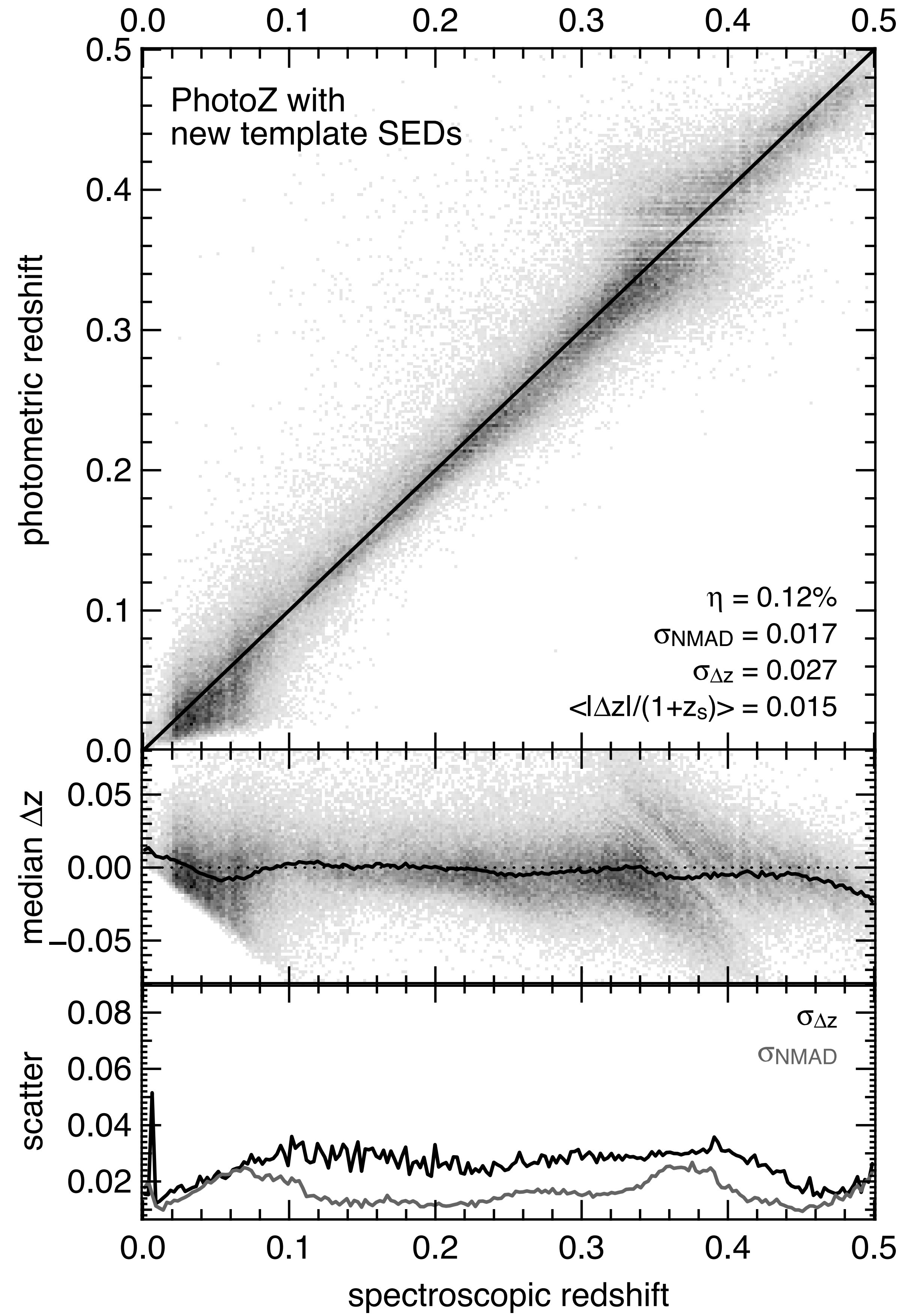
Csabai et al. 2003, ApJ 125, 580

SLOAN LRGs: ANNz results

$$\eta : \frac{|\Delta z|}{1+z_{spec}} > 0.15$$

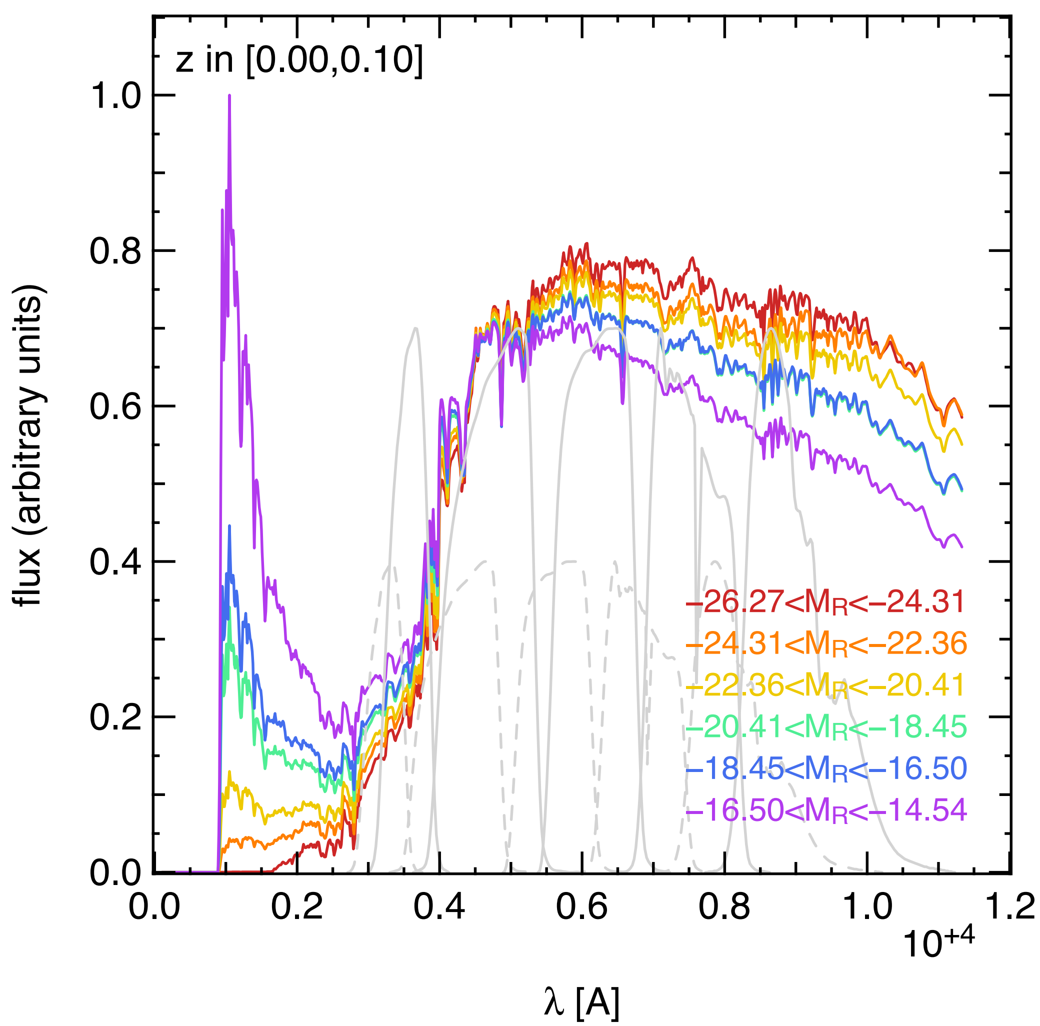
$$\sigma_{\Delta z/(1+z)} = 1.48 Median \left(\frac{|\Delta z|}{1+z_{spec}} \right)_{non-outliers}$$



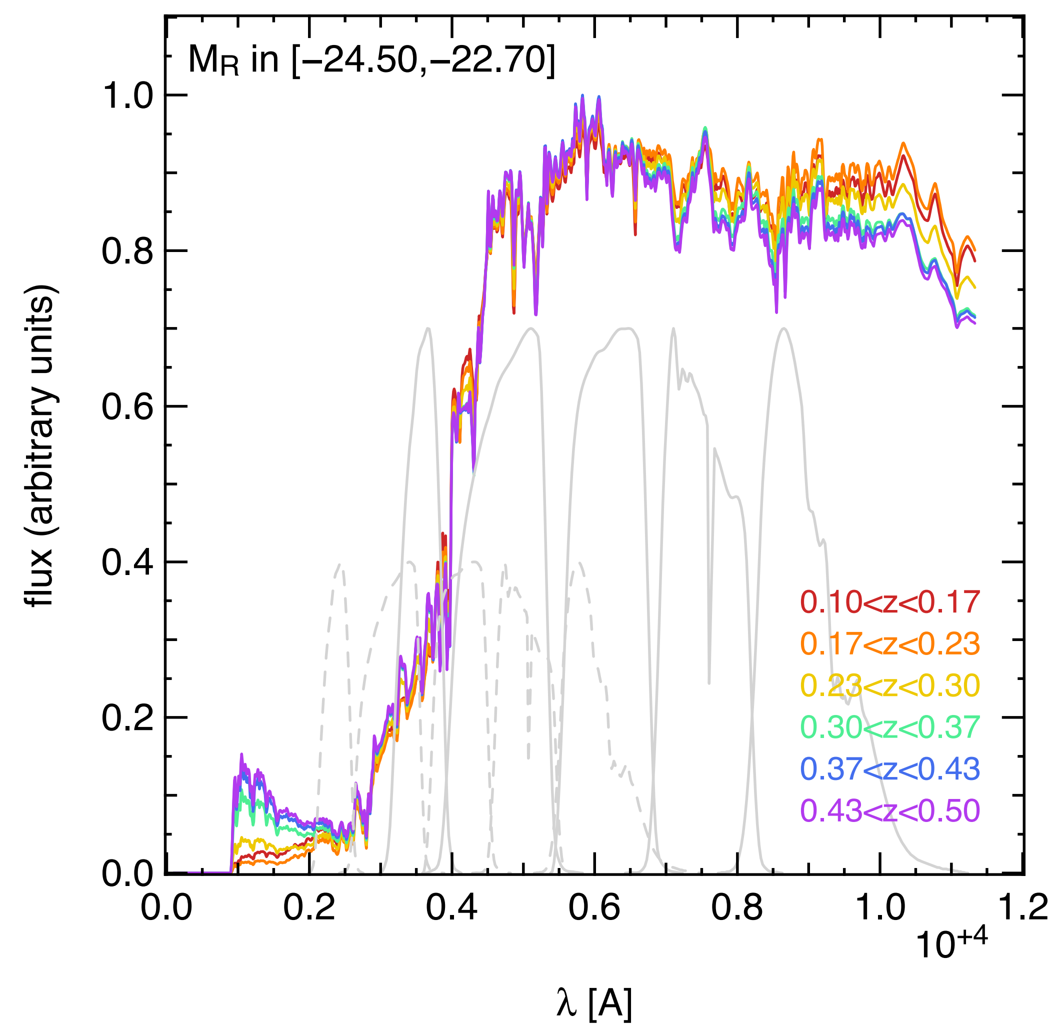


LRGs spectral evolution

Faint local LRGs have residual recent star formation

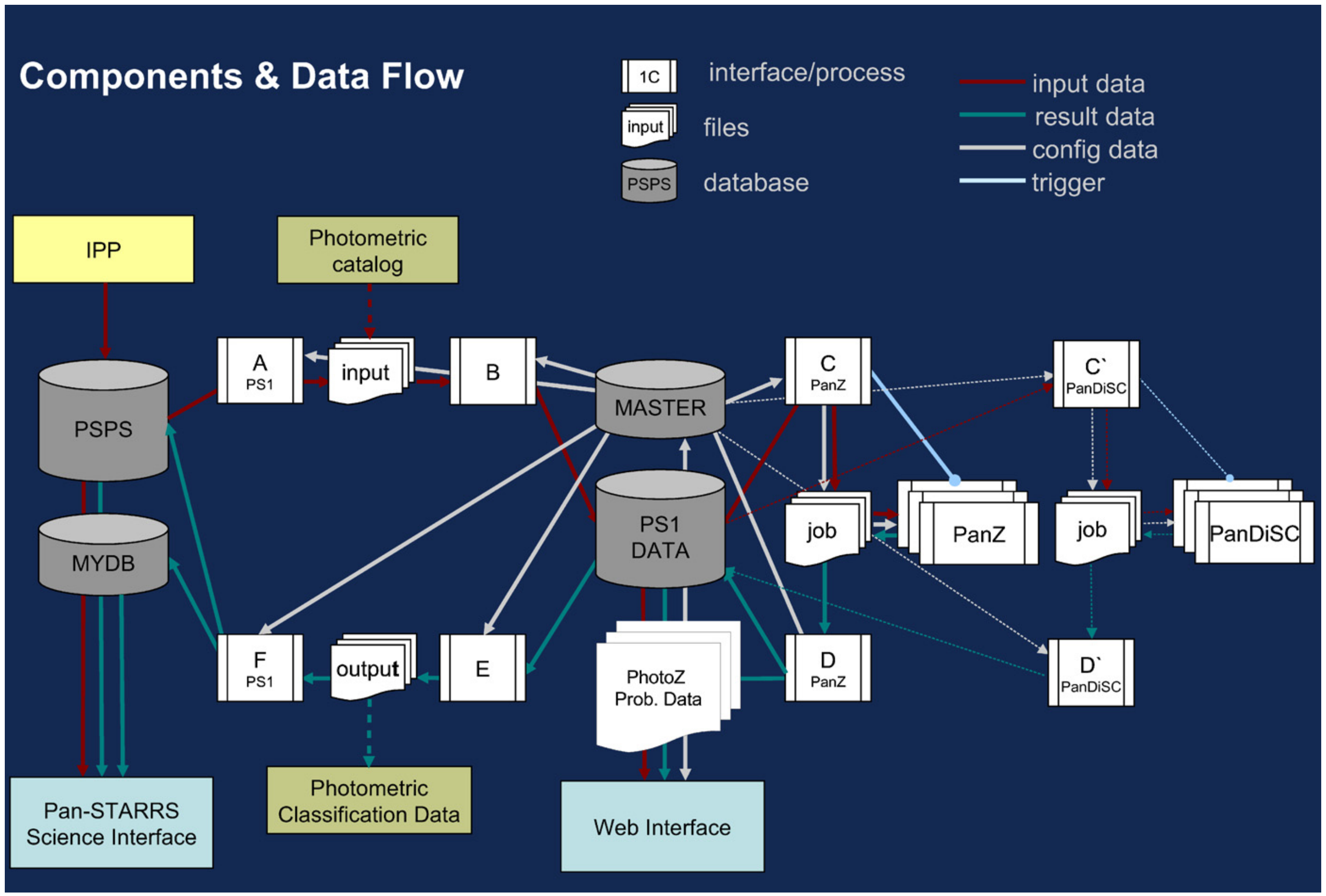


UV flux present at higher z maps into increased NIR flux in local LRGs



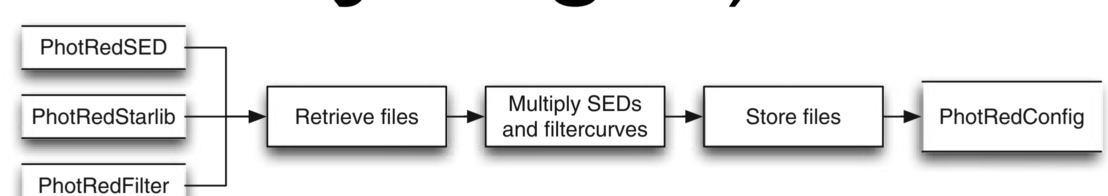
Implementations: PanZ and AstroWise

PanZ & PanDiSC for PS1



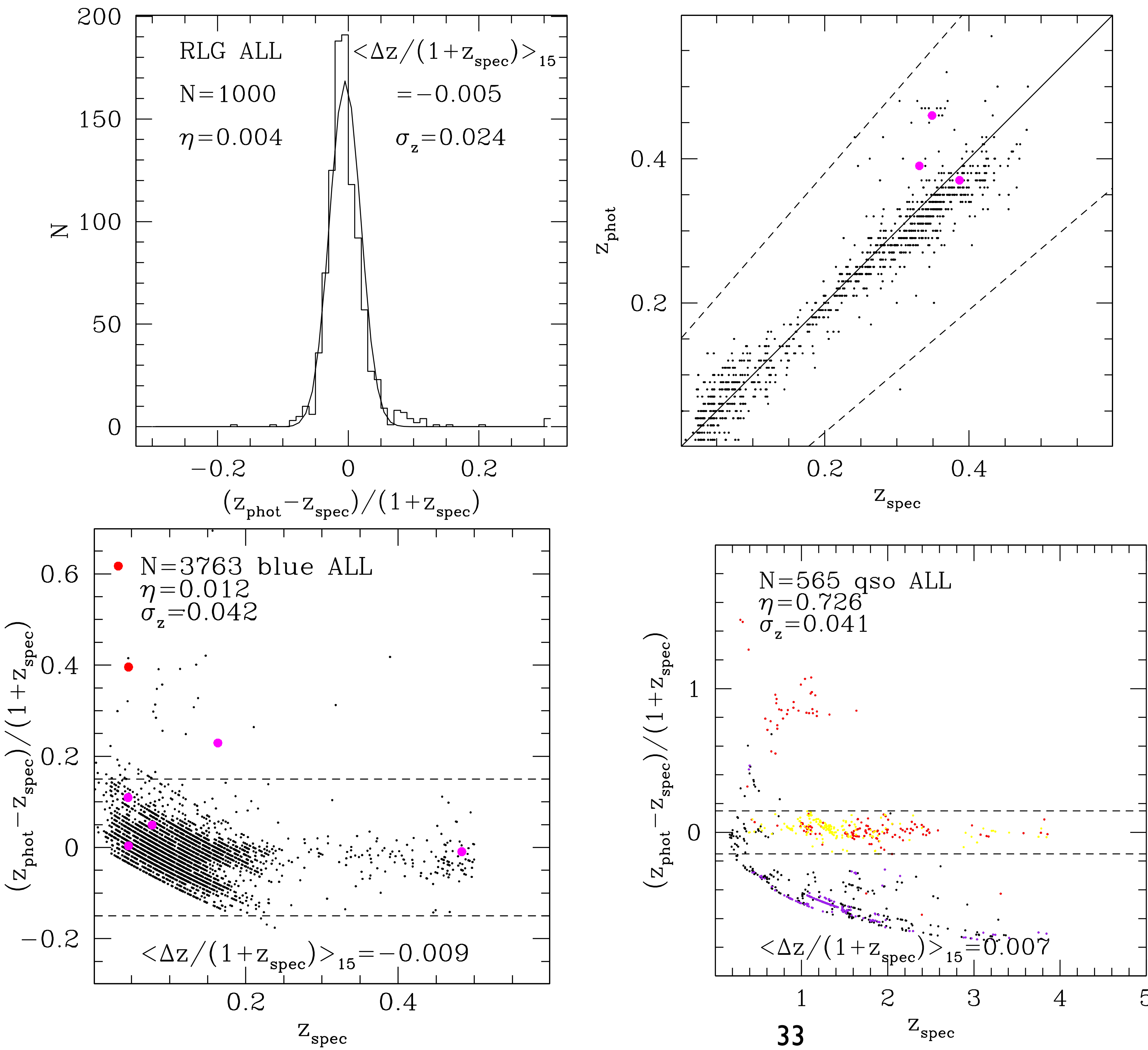
Database-supported automatic production of photometric redshifts (R. Senger)

PhotoZ in AstroWise for KIDS (J. Snigula)



Saglia et al. 2012, ApJ, 746, 128
Saglia et al., 2012, Exp.A. 127

PanStarrs: first results



Saglia et al. 2012, ApJ,746, 128

Excellent LRGs

Reasonable blue galaxies

Bad QSOs...

(see Salvato et al. 2011, ApJ 742, 61)

Conclusions

- Photometric redshifts from broad-medium band photometry deliver precisions better than 2% (for LRGs) with low fractions of catastrophic failures

- Currently running (PanSTARRS1, KIDS) or soon starting (DES) photometric surveys will deliver catalogues with hundred thousands of galaxies

- Several science cases can be served: search for galaxy clusters, confirmation of eRosita extended sources, BAOs, weak lensing tomography, etc.

- EUCLID science case relies on exquisitely accurate photometric redshifts (OU-PHZ)

- Even if empirical methods (i.e. ANNz) are probably superior if extensive spectroscopic training sets are available, template fitting bayesian methods are required to study galaxy properties and their evolution. A combination of both methods will provide the best solution.