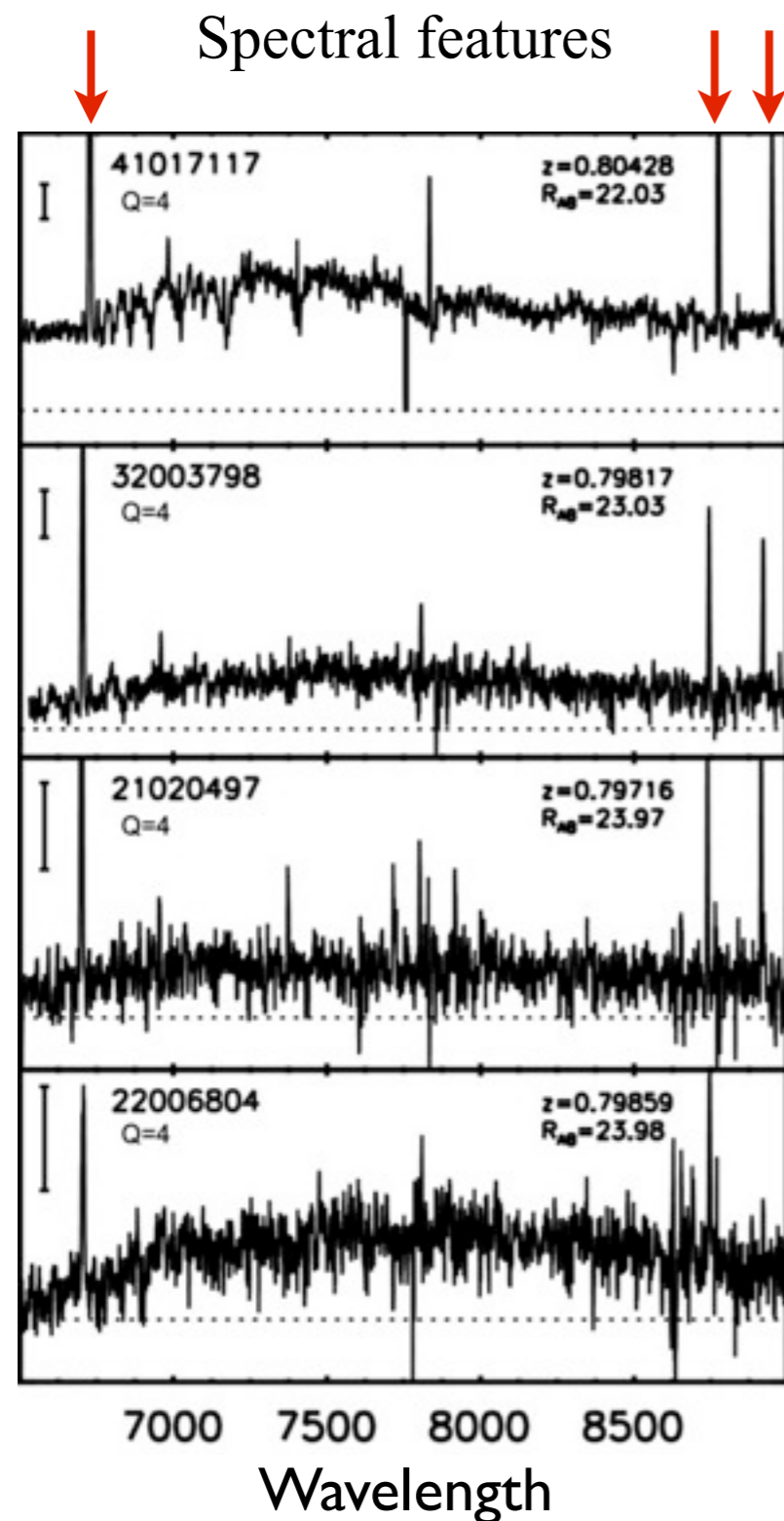


# **Challenges in Redshift Measurement and Use for Large Astronomical Datasets - Spectroscopic and Photometric**

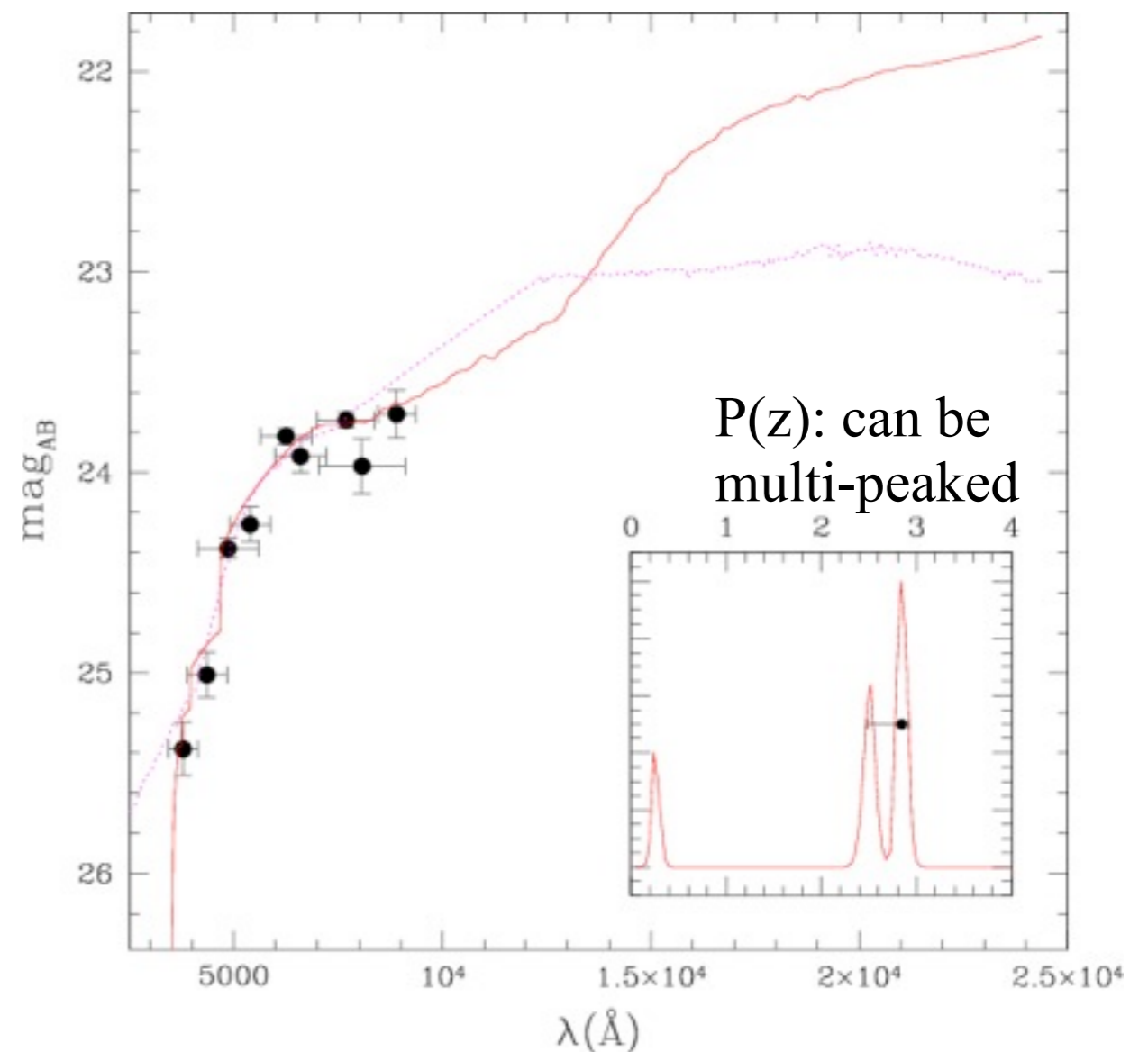
**Benjamin Weiner**  
**Steward Observatory**

Measuring galaxy redshift is fundamental for inferring luminosity,  
but also clustering, gravitational lensing potential, probes of cosmology ...



DEEP2 spectra - Newman et al 2013

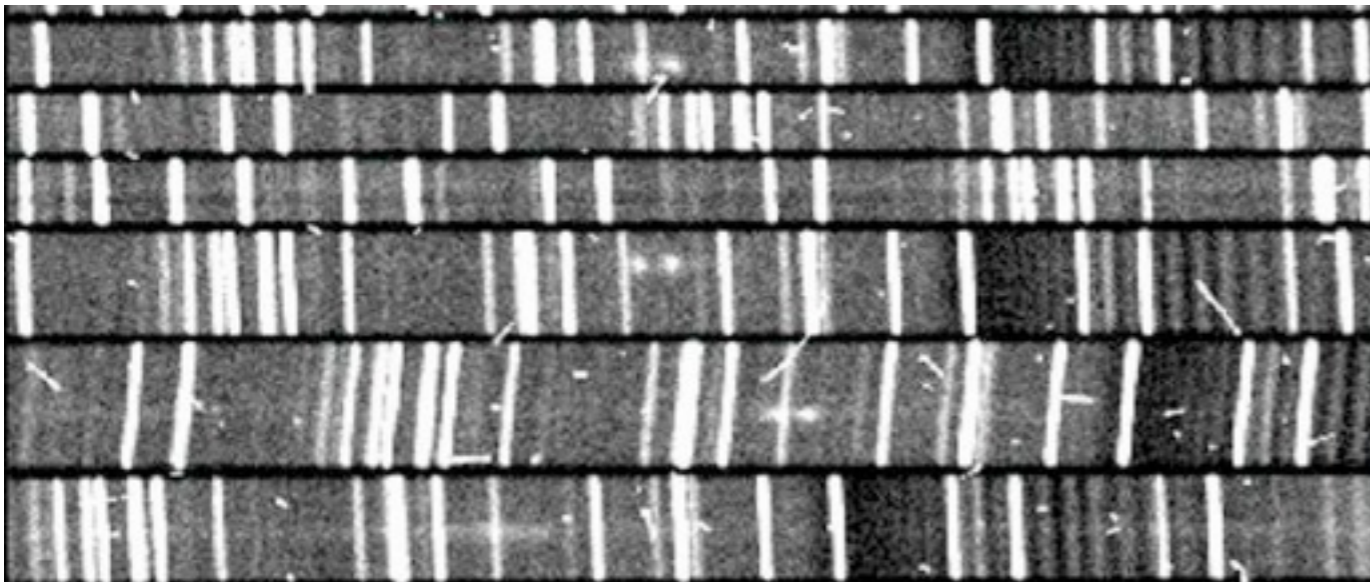
Photometric redshift: fit galaxy models to  
multi-band fluxes, yield a “P(z)”



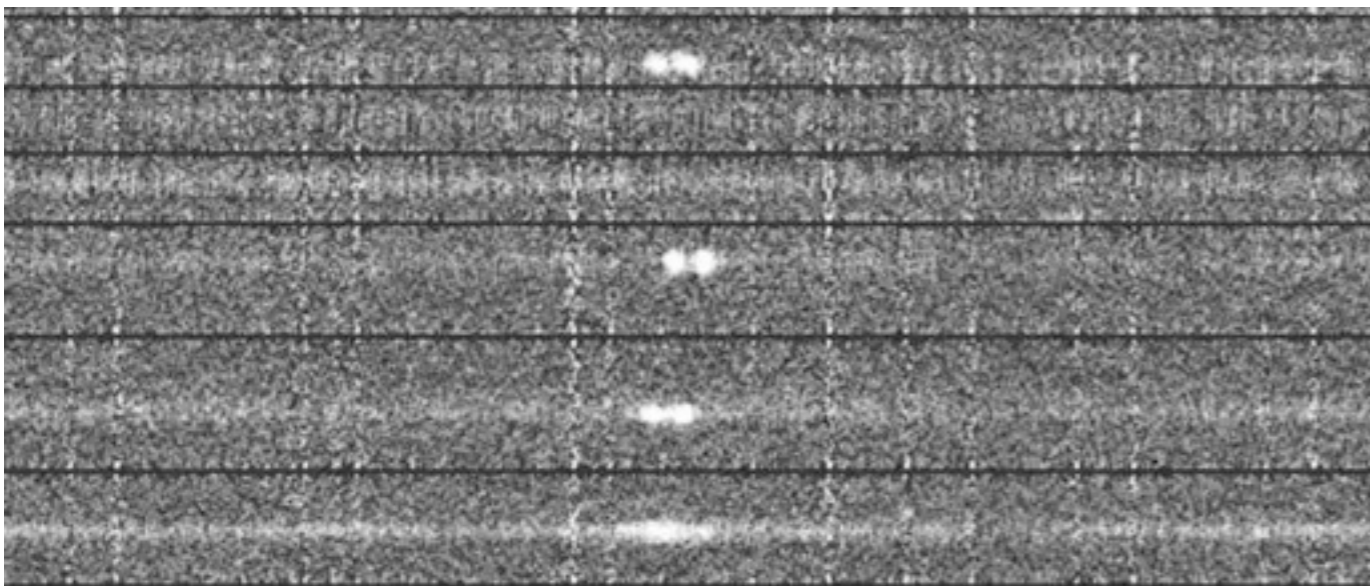
CFHT-LS - Ilbert et al 2006

# Faint Galaxy Spectroscopy from DEEP2

## The Good

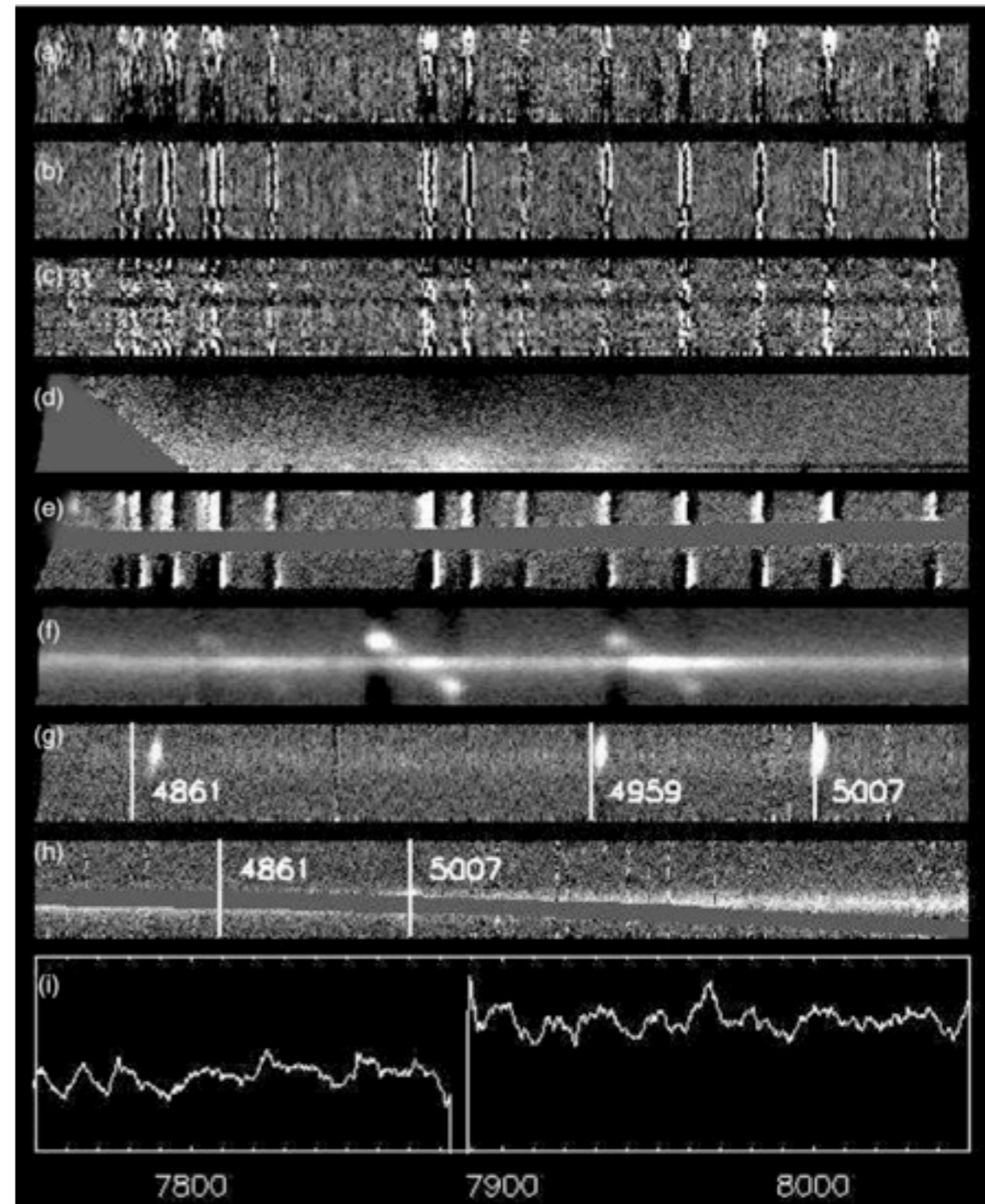


↓ Flatfield, sky subtraction,  
combined exposures, etc



Wavelength →

## The Bad and Ugly

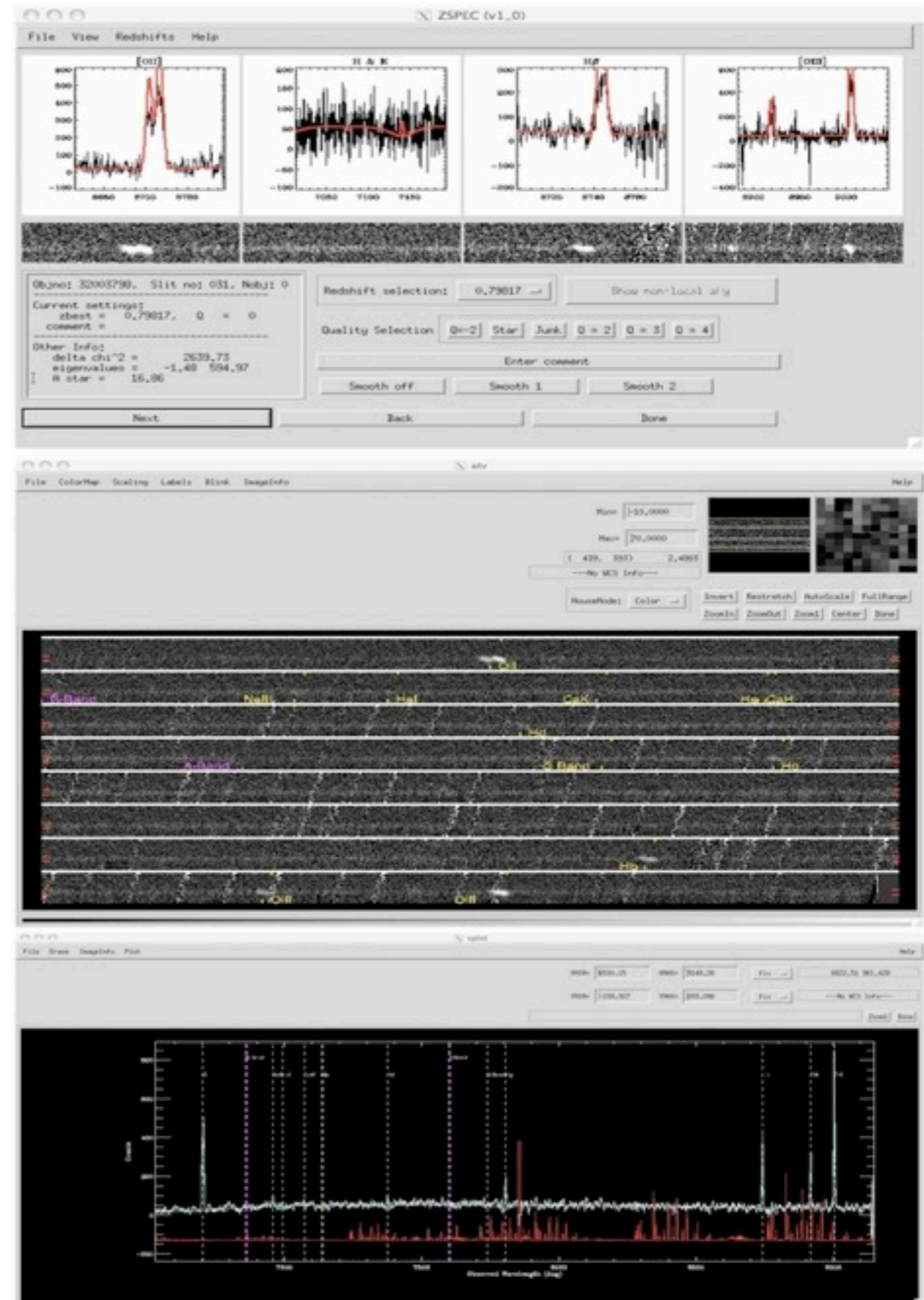


Newman et al 2013

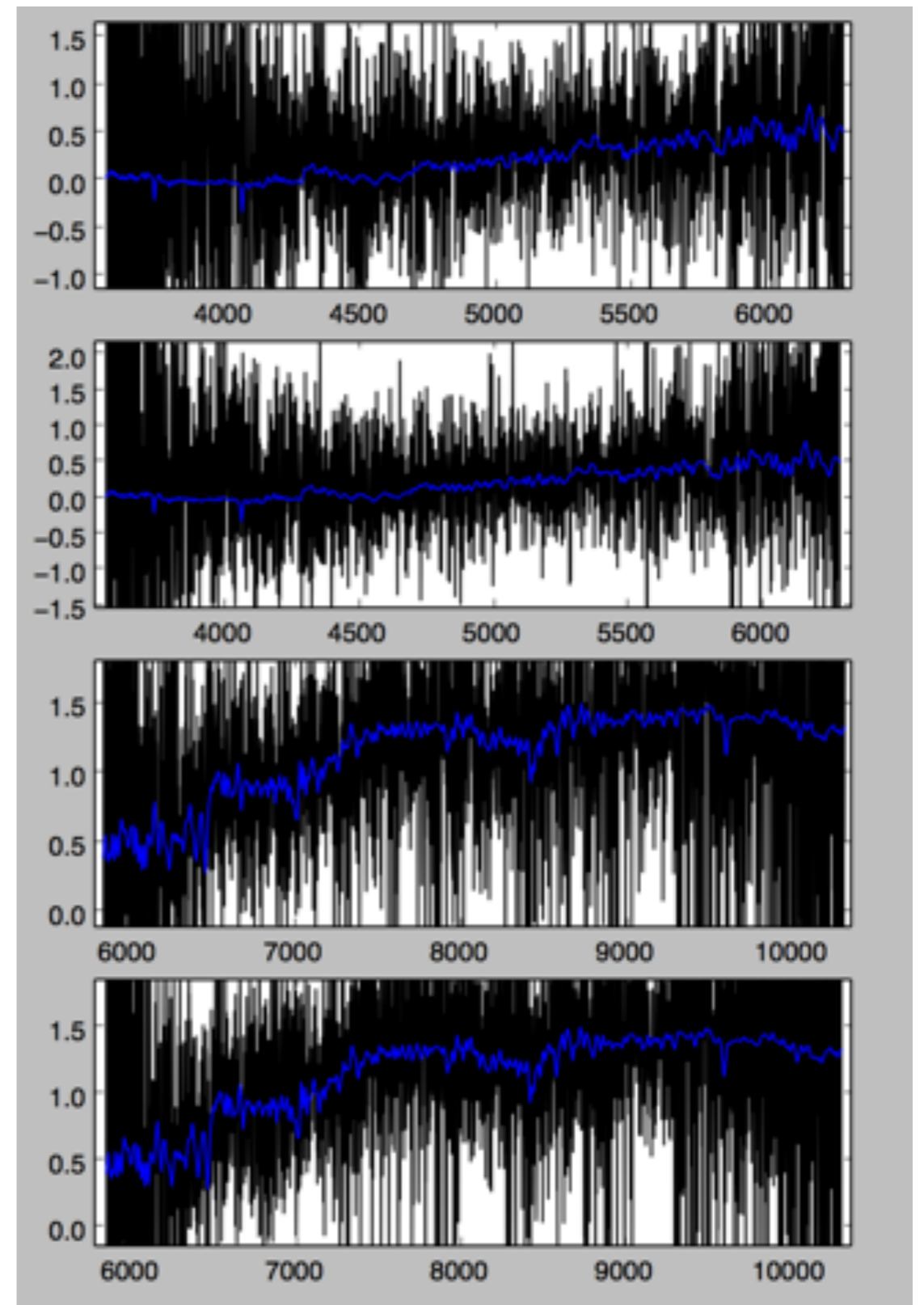
# Measurement: Automated Quality Classification: People

- Redshifts are measured in SDSS and DEEP2 by shifting spectrum and fitting linear combination of a set of templates (physical or PCA components) to the spectrum. Sharp local minima in the run of  $\chi^2(z)$  are candidate redshifts.
- SDSS ( $z \sim 0.15$ ) and BOSS (red galaxies at  $z \sim 0.5$ ) automatically determine redshift quality. In DEEP2 ( $z \sim 0.7-1.4$ ) we could not develop  $\chi^2$  or other criteria to separate good/bad redshifts. We checked  $\sim 50,000$  spectra yielding  $\sim 35,000$  galaxies, using an interactive IDL gui “zspec”. It was a PITA, slow, and won’t scale up.
- Harder in DEEP2 due to lower S/N of spectra; less wavelength coverage; features in red with sky residuals; emission line galaxies.
- Expect similar problems in future very large redshift surveys for cosmology: eBOSS, MS-DESI, Euclid, WFIRST.

## DEEP2 gui: Newman et al 2013



- Redshift fitting is being improved for eBOSS, the emission line galaxy part of SDSS-IV, with: better templates; non-negative fitting for physical fits; fitting in data domain rather than rebinning spectra.  
(Adam Bolton and eBOSS team).
- There may be gaps in what future dark energy surveys want and what the community wants. Wrong or mis-identified redshifts are a noise/bias term in a large scale clustering/BAO survey, but they are a **systematic and asymmetrical error in galaxy properties such as the luminosity function** (or just about any other property).
- Spectroscopic redshift quality assessment should also be of interest as a **machine learning problem**, since it can be formulated as supervised or unsupervised classification.
- There is **lots of public data out there** for anyone to work on, including all of DEEP2 and SDSS-1, 2, 3!



Fitting to separate exposures in data domain without rebinning. Courtesy Adam Bolton (Utah) and eBOSS team.

# Photometric Redshifts: what is $P(z)$ and how should it be used?

Standard method of fitting photo- $z$  is to take a large set of templates, synthesize photometry, at each redshift find the  $\chi^2$  of best fitting template to data, and transform this  $\chi^2(z)$  into “ $P(z)$ ”.

This  $P(z)$  is really  $P(\text{data}|z, \text{templates})$ , a **likelihood of the data**, given the model that the galaxy is exactly described by some template of the set.

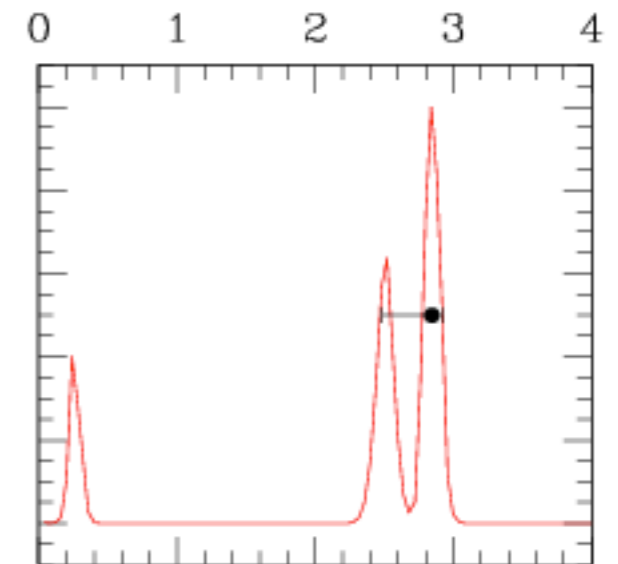
$$P(\text{model}|\text{data}) = \frac{P(\text{data}|\text{model})P(\text{model})}{P(\text{data})}$$

$P(\text{data}|\text{model})$  is likelihood of data given the model  
 $P(\text{model})$  is prior on the model params

Ideally  $P(z) = P(z_{\text{true}}|z_{\text{phot}})$   
which is distinct from  
 $E(z) = P(z_{\text{phot}}|z_{\text{true}})$   
the error distribution

$P(\text{model})$  can include prior beliefs on how likely each template is at the given redshift and luminosity, as in the BPZ method (Benitez 2000).

However, priors on these and everything else that can go wrong (e.g. photometric zero point offsets, galaxies not exactly in the template set) are rarely done in a full Bayesian way.  $P(z)$  should be interpreted cautiously, and catastrophic outliers are common at the 5-10% level.



$P(z)$  from Ilbert+ 2006

# Galaxies are strongly clustered on scales smaller than a good photo-z error bar

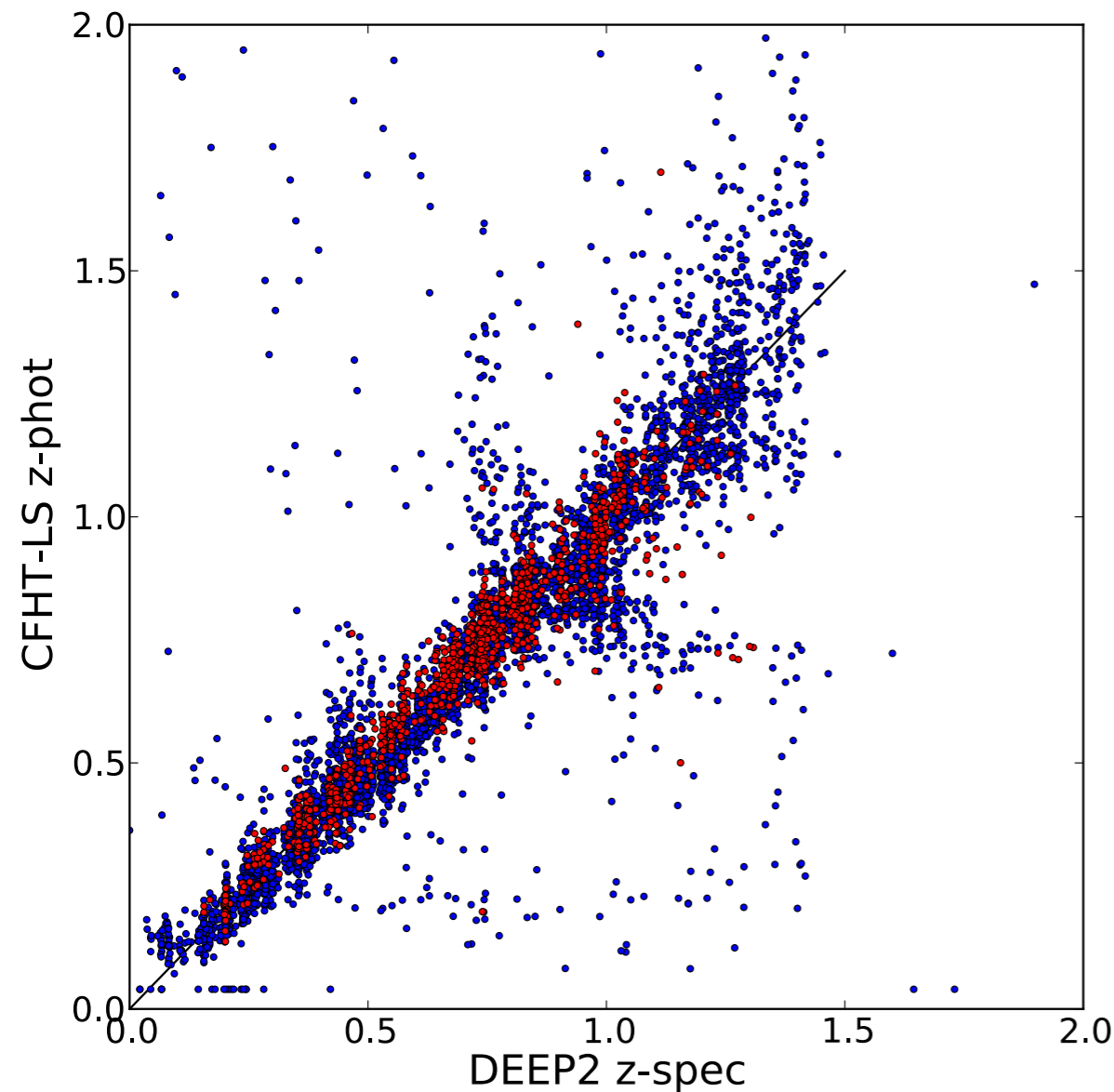
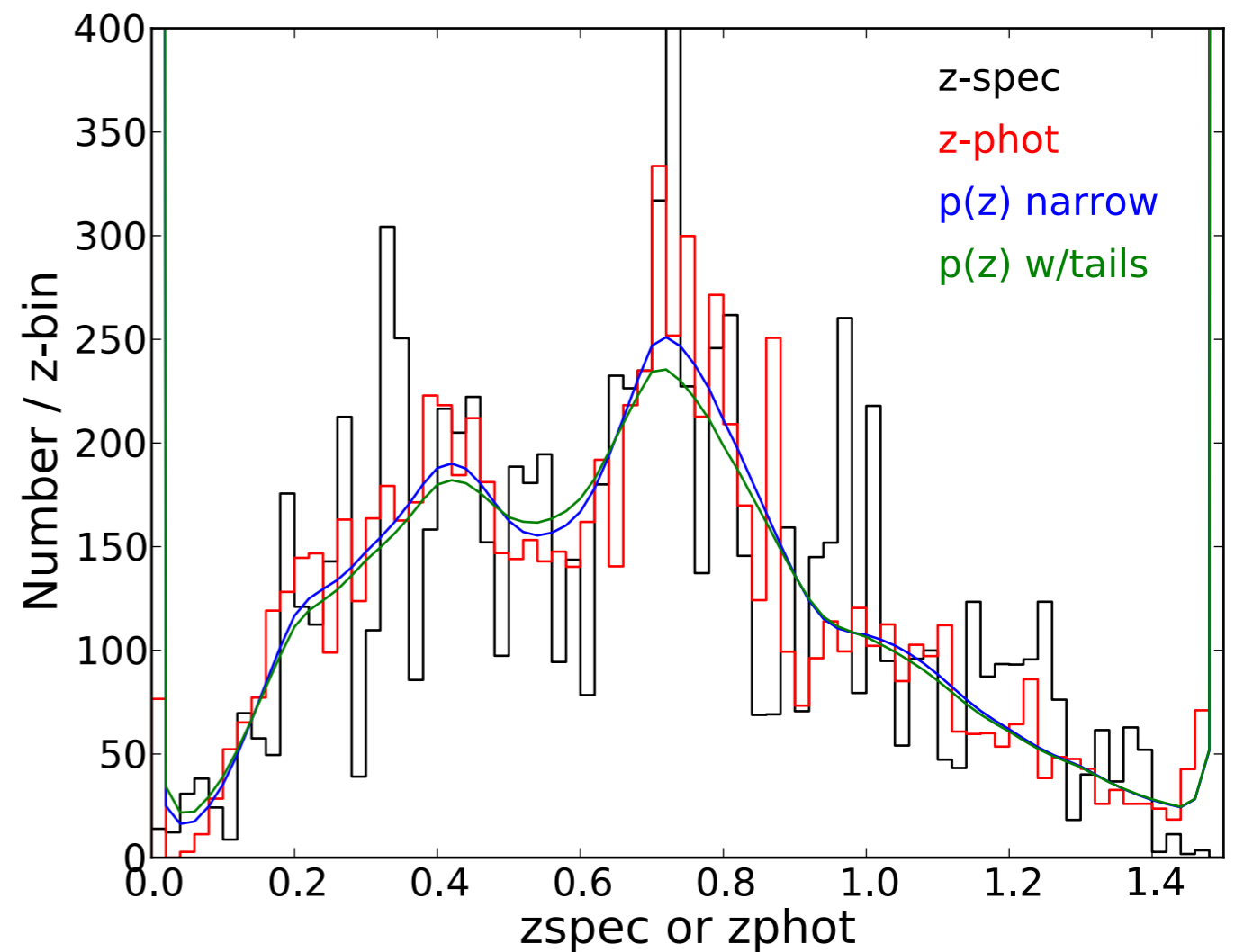


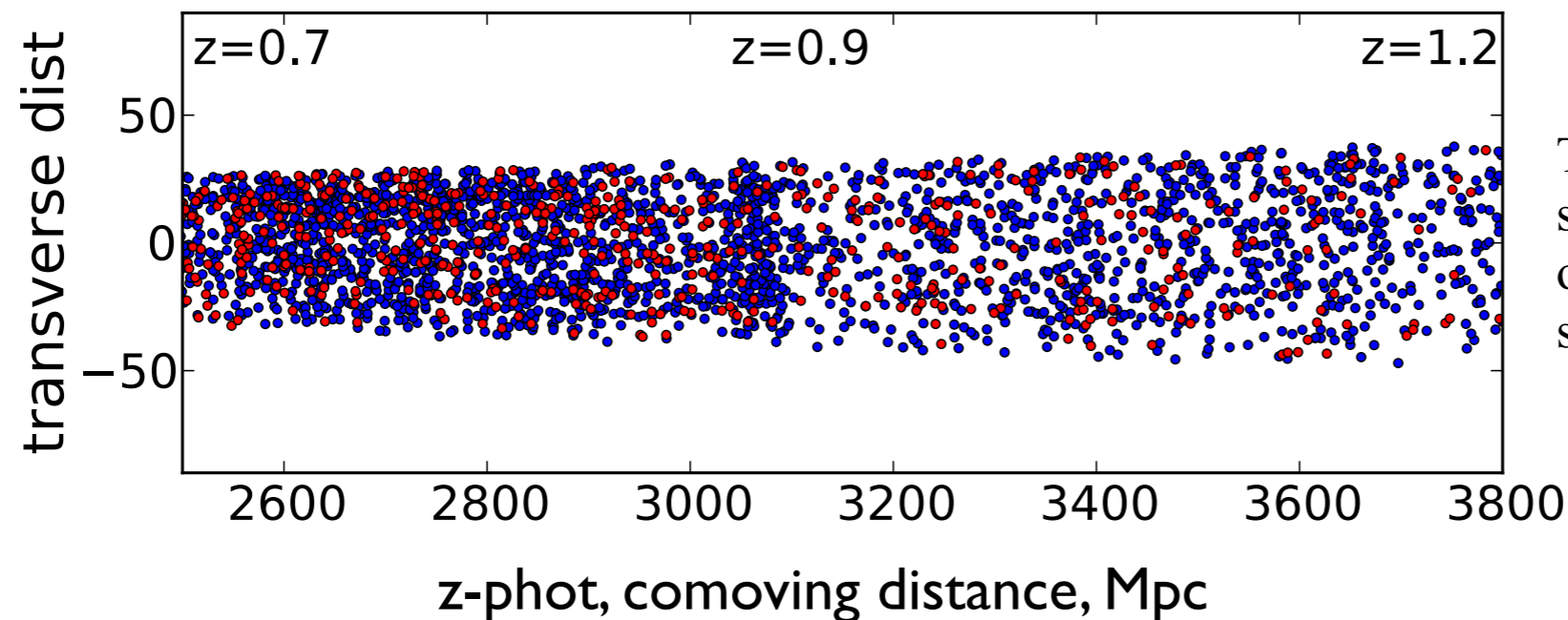
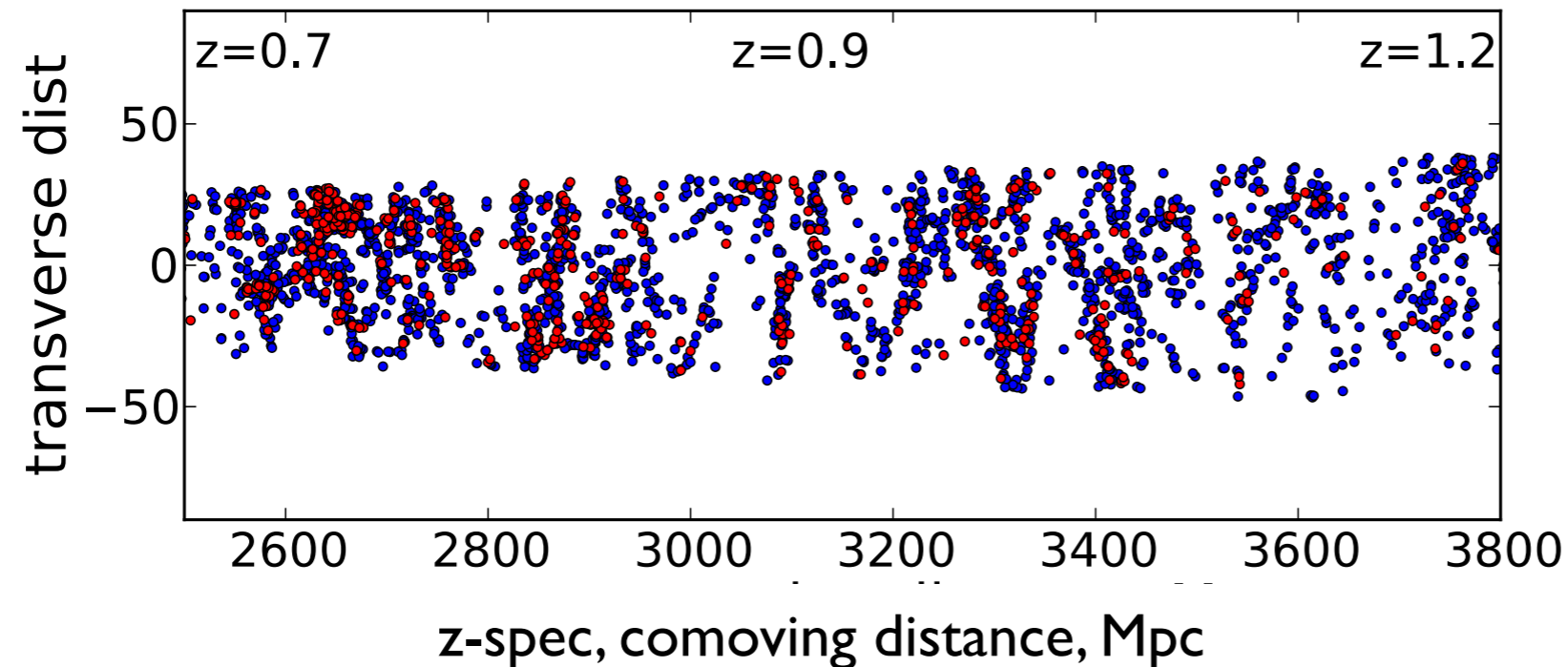
photo-z: Ilbert et al 2006, spec-z: DEEP2

$N(z)$  is smoothed by photo-z. **Summing up or sampling from the  $P(z)$  for each galaxy smooths it **again**** rather than reconstructing the true  $N(z)$ .



Spikes in this plot are large scale structure of the universe, not histogram binning noise.

Photometric redshifts smooth out the clustering of galaxies,  
so need to calculate everything in large redshift bins



The best z-phot is plotted. If we sampled from the  $P(z)$  for each object this would be even more smoothed-out.

# How can we use the extra info in $P(z)$ to make more accurate inferences?

Use case: cluster gravitational lensing

Fit cluster mass to set of measured shears of background field galaxies

Point estimate of best  $z$ :

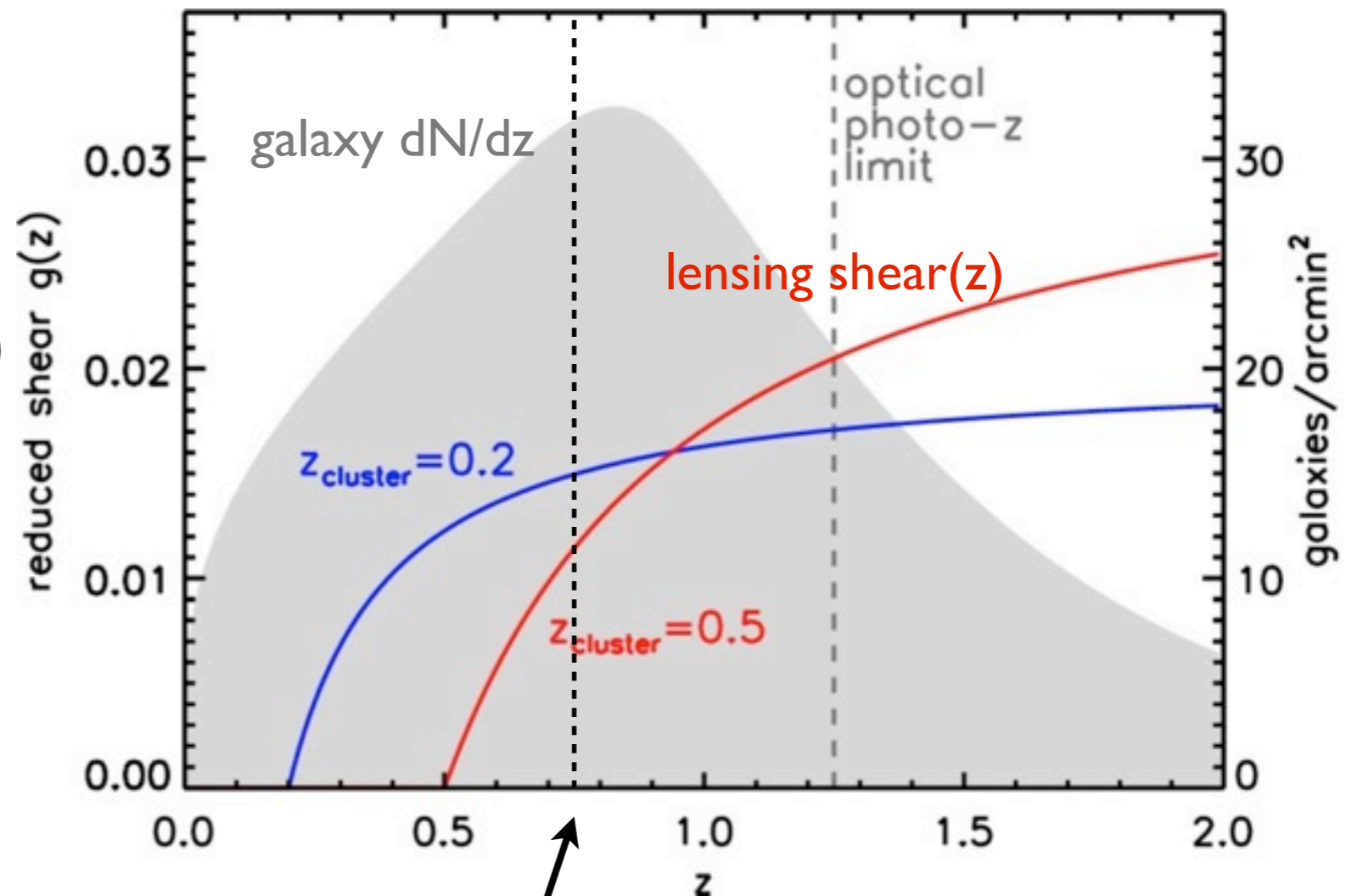
$$P(M_{cl}|shear) = P(M_{cl}|shear(z_{best}))$$

Or marginalize over  $P(z)$ :

$$P(M_{cl}|shear) = \int P(M_{cl}|shear(z))P(z)dz$$

Marginalizing over  $P(z)$  should matter where the shear( $z$ ) is changing rapidly, like just behind the cluster.

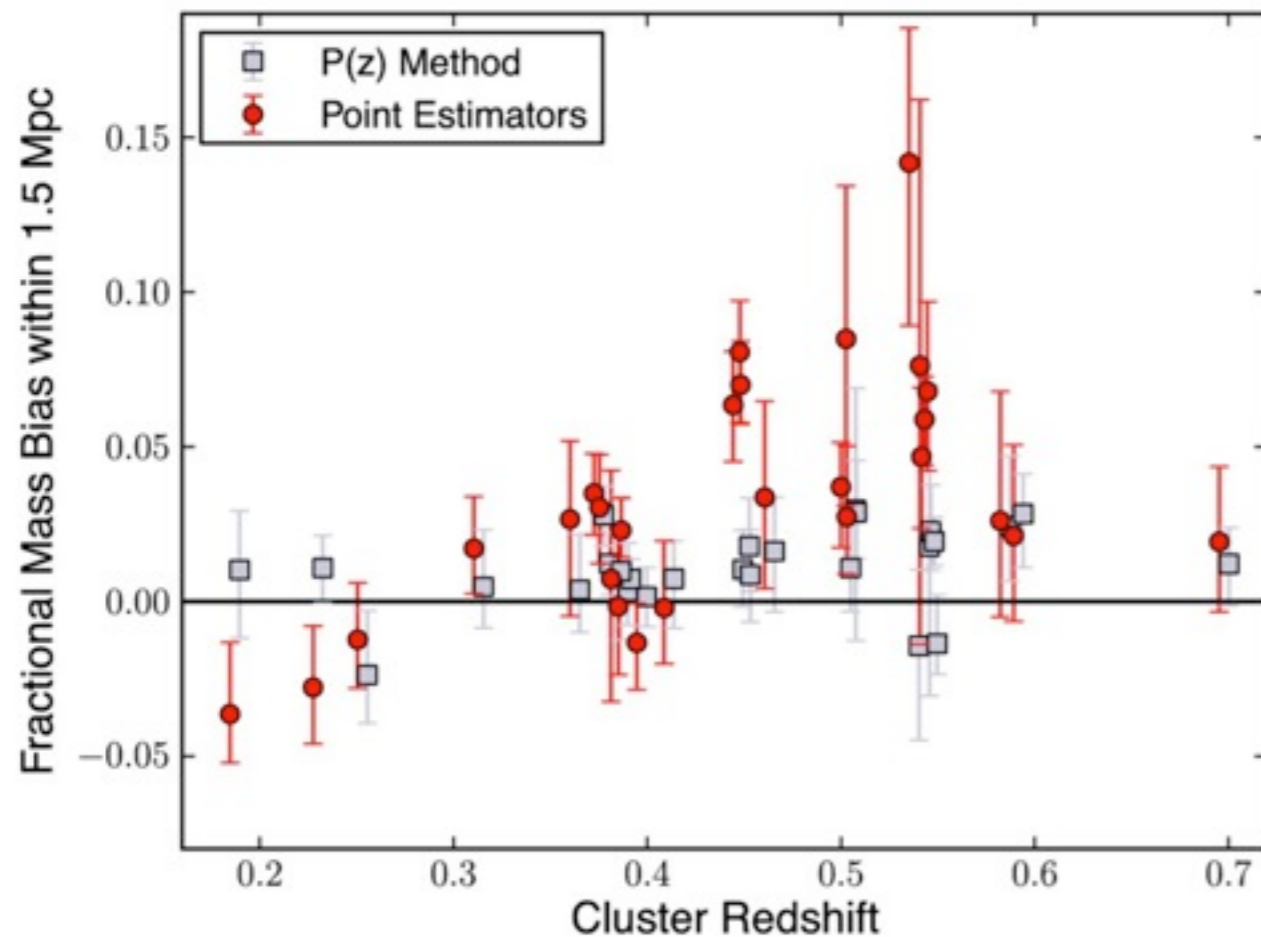
But wait! Marginalizing sounds good, but isn't this summing over  $P(z)$ , and **didn't we just decide that oversmooths the  $N(z)$  distribution?**



Applegate et al 2014

Here, if  $P(z)$  is broad or skewed, it influences the expected shear

# Inference on a skew function like shear or galaxy luminosity can be biased by $P(z)$ width



Applegate et al 2014

Weak gravitational lensing and galaxy angular clustering are both strongly dependent on the redshift distribution  $N(z)$  of the sample.

These are keystones of efforts to measure cosmological parameters and dark energy, so  $<1\%$  effects become critically important.

Applegate 2014 shows that **lensing cluster masses from point estimator  $z_{\text{best}}$  are biased by a few %** for redshifts where the bulk of the galaxies are just behind the cluster, compared to marginalizing over each galaxy's  $P(z)$ . But there is still an irreducible error from photo- $z$  redshift errors.

**But Sheldon (2012) has argued that fitting the overall  $N(z)$  to the observables is less biased than using each galaxy's  $P(z)$ , due to finite width and errors in each individual  $P(z)$ .**

By now you should be **TOTALLY CONFUSED.**

# How should users of a large photometric database do inference with photo-zs like good Bayesians?

1.  $P(model|fluxes) \propto P(fluxes|model) \times Prior$

“Model” is something like  $N(z, flux > limit)$  or luminosity function  $\phi(L, z)$ .

The likelihood  $P(fluxes|model)$  means generating all the observed fluxes from your LF and its evolution, for all galaxies including foreground and background. Few people will do this.

2.  $P(model|z_{phot}) \propto P(z_{phot}|z_{true}) \times P(z_{true}|model) \times Prior(model)$

$\uparrow$   
 $E(z_{phot}|z_{true})$

Error distribution, NOT  $P(z)$ :

$P(z_{true}|z_{phot})$

$\uparrow$   
Likelihood of drawing a galaxy at  $z_{true}$  from model

Implication: If you want to do forward modeling, **you need access to the photo-z method to characterize the error distribution  $E(z_{phot}|z_{true})$ . Just having the  $P(z)$  isn't enough** even if the  $P(z)$  are perfectly correct.

Similar to deconvolution and the need to model the PSF. Community use of DES, LSST products will depend on access to methods, not just data products.

**In summary: Lots of opportunity for clever techniques to improve use of both spectroscopic and photometric redshifts.**