Modelling photo-z in current and future imaging dark energy experiments

Will Hartley, UCL

w. DES (esp. Christopher Bonnett, Michael Troxel,
Adam Amara, Boris Leistedt)

Overview

→ Introduction – the principle of photo-z

→ The basic failure mode of template / model photo-z

→ The danger of ad hoc, photometric corrections

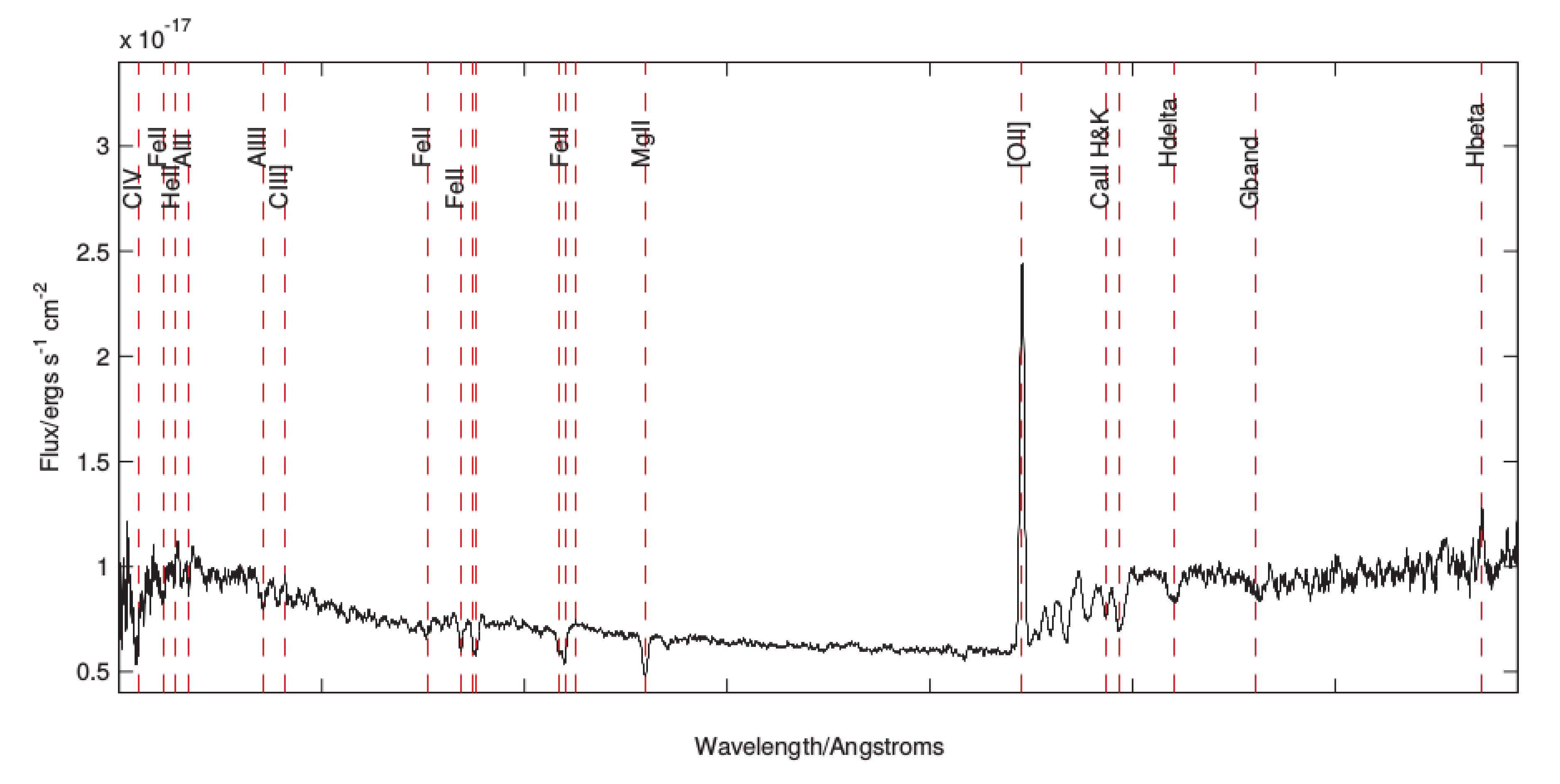
→ Where we are now – BPZ in the DES

→ Next steps – using galaxy evolution to inform:
- template improvements
- better priors

→ Case study - RedMaGiC

Obtaining distances to galaxies...

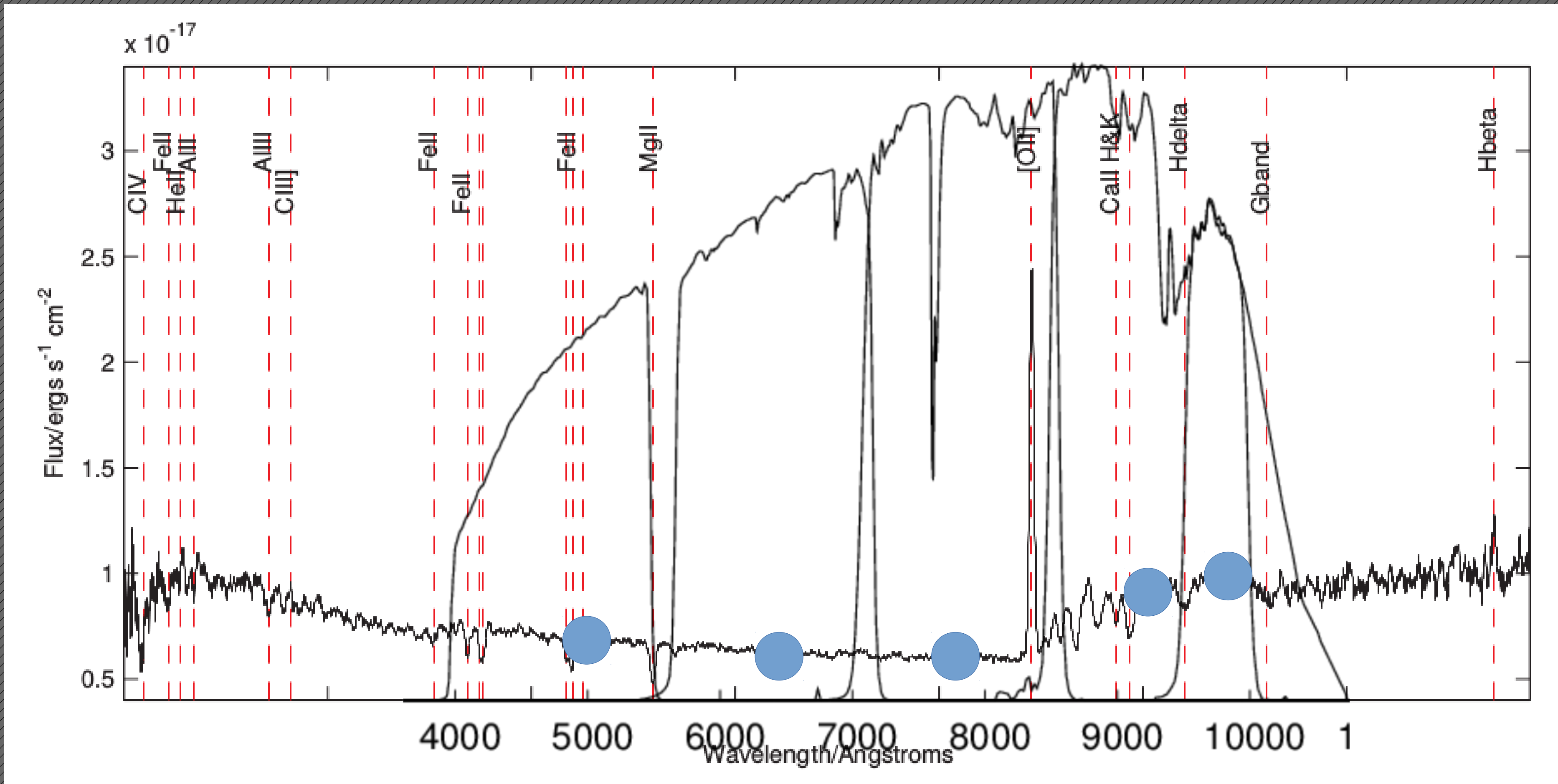
Bradshaw, Almaini, Hartley et al. (2013)



Emission / absorption features allow the accurate determination of redshift

Obtaining distances to galaxies...

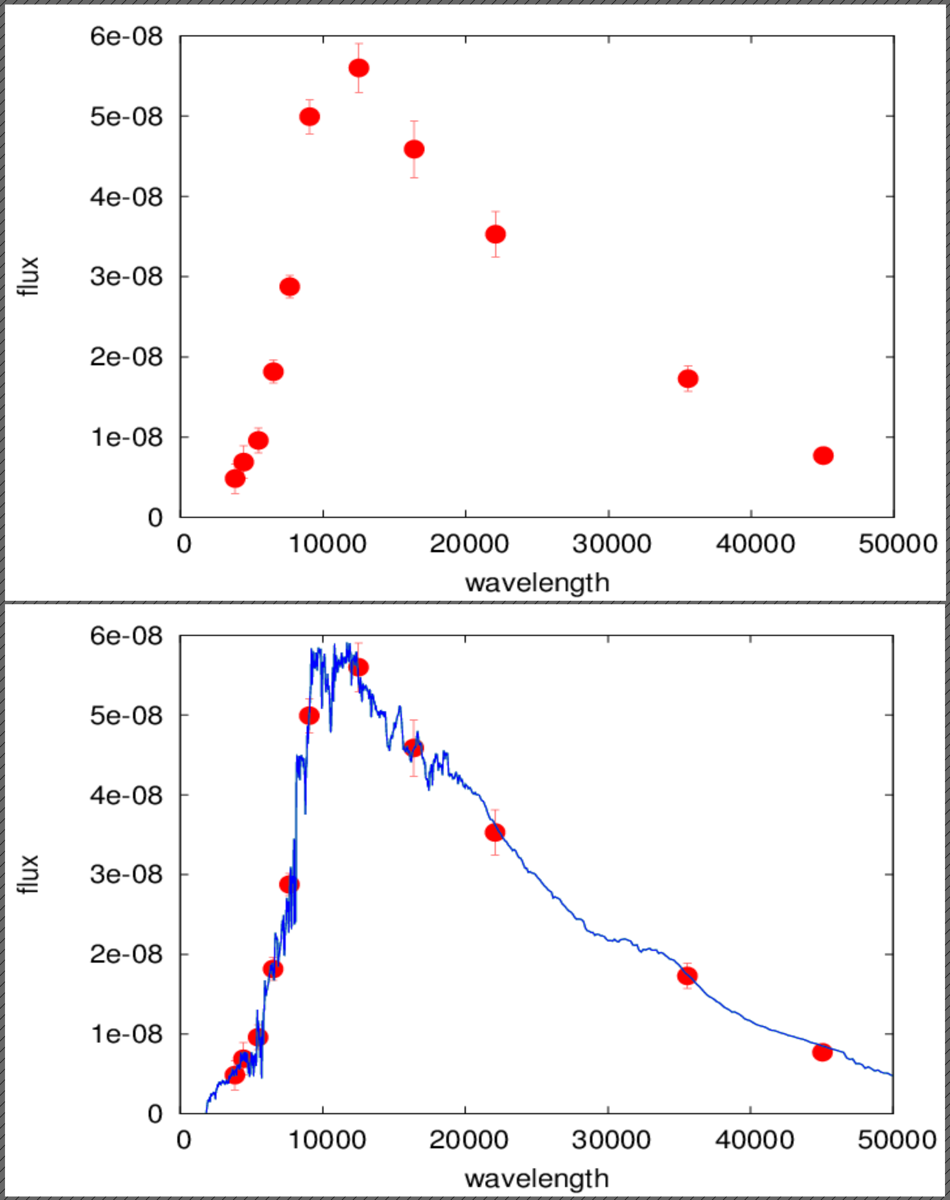
Bradshaw, Almaini, Hartley et al. (2013)



Where only broad band photometry is available, we have a poorly sampled estimate of the continuum shape – which can be used to get an approximate redshift.

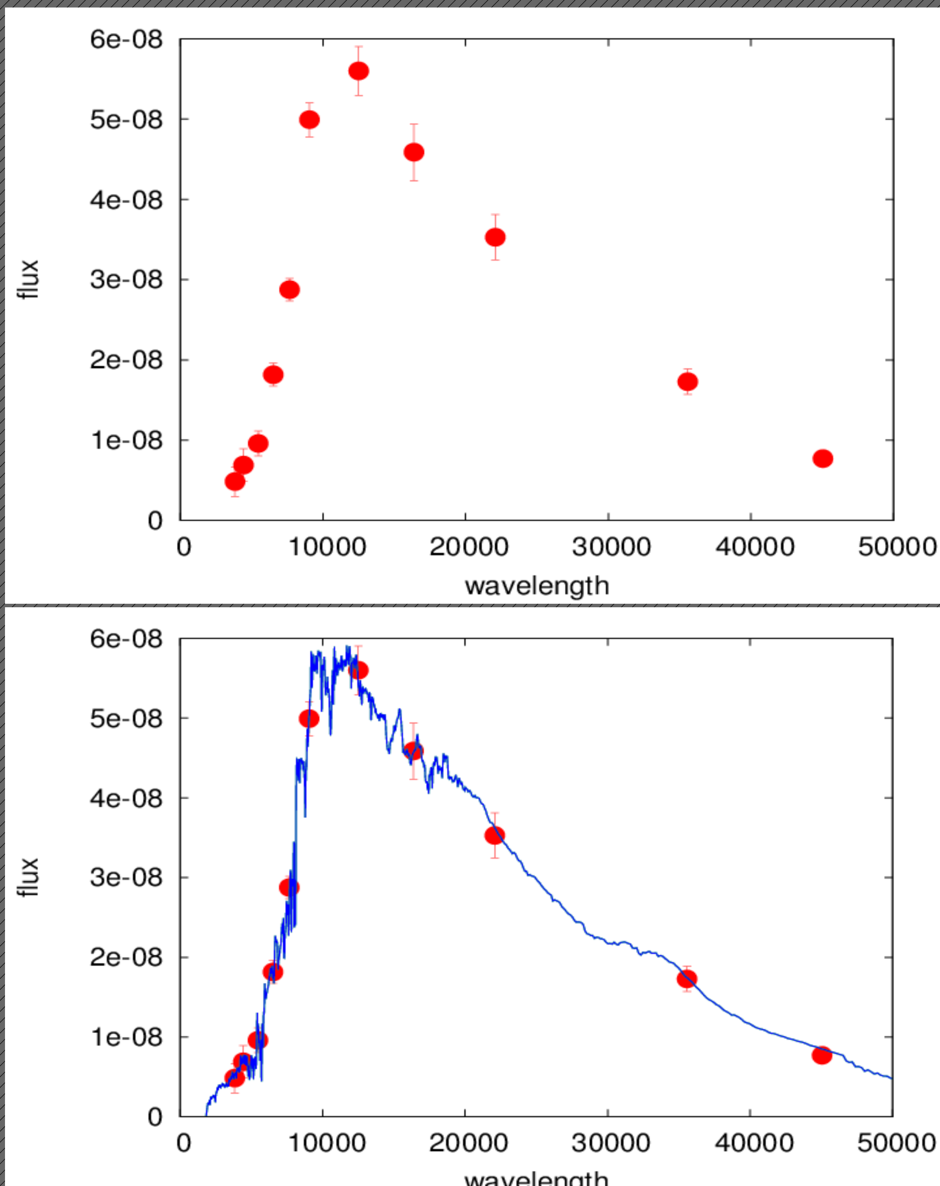
The basic principle of photo-z...

'Template' photo-z:



The basic principle of photo-z...

'Template' photo-z:



Machine Learning photo-z:

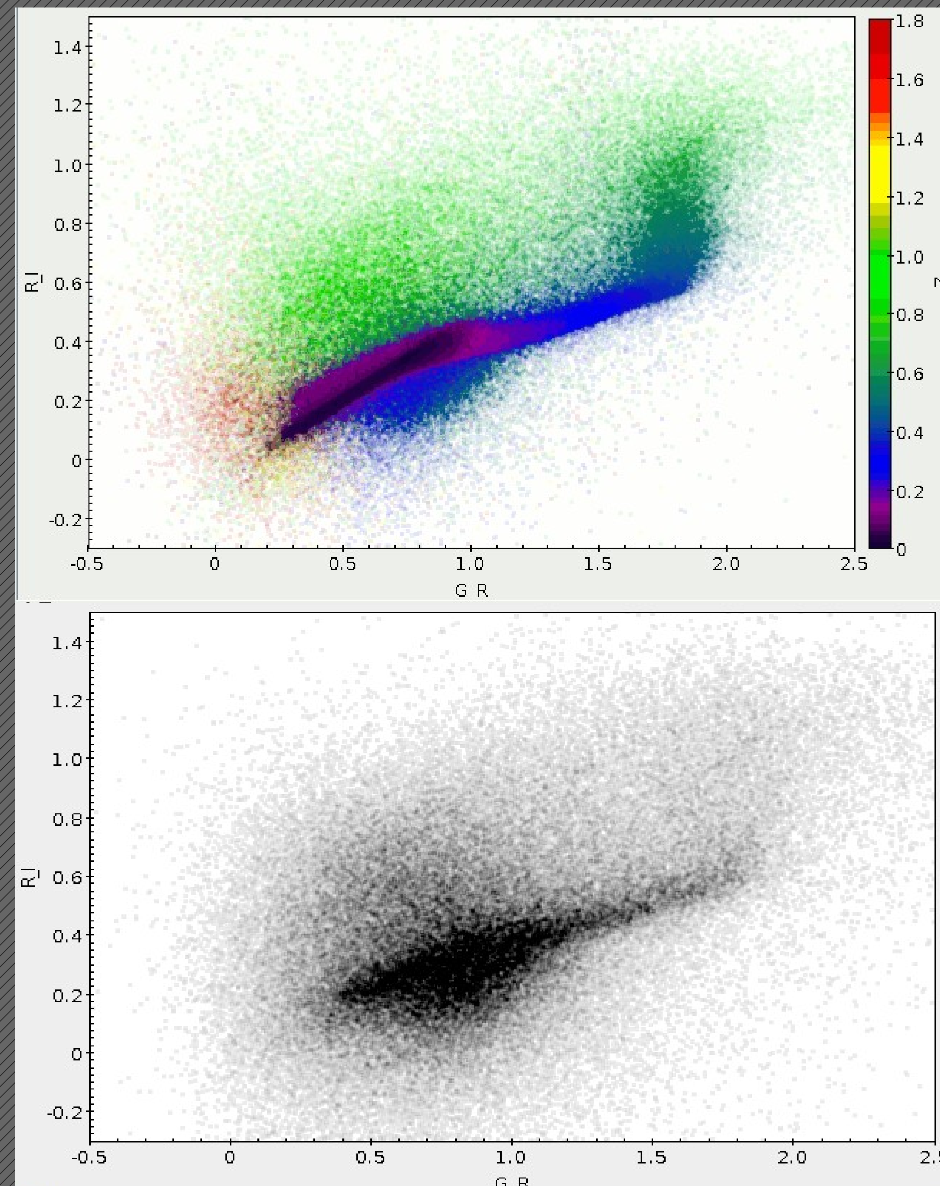


Photo-z in cosmology experiments (stage III, IV) – pushing the limits

DES:

- 5 bands (g, r, i, z, Y)
- $i < 24$ (10σ detection limit)
- target accuracy on bias, $dz/(1+z) \sim 0.01$
- target accuracy on scatter, $dz/(1+z) \sim 0.12$

EUCLID:

- 8* bands (u/g → H)
- vis < 24.5
- target accuracy on bias, $dz/(1+z) \sim 0.002$
- target accuracy on scatter, $dz/(1+z) \sim 0.05$

Deep extragalactic fields manage

$dz/(1+z) \sim 0.03$ (11 bands, UDS)
$dz/(1+z) \sim 0.01$ (~30 bands, COSMOS)

in scatter for bright samples.

But bias is typically not optimised (or even measured).

For DES SV, all of the best-performing codes were based on training (Machine Learning) approaches.

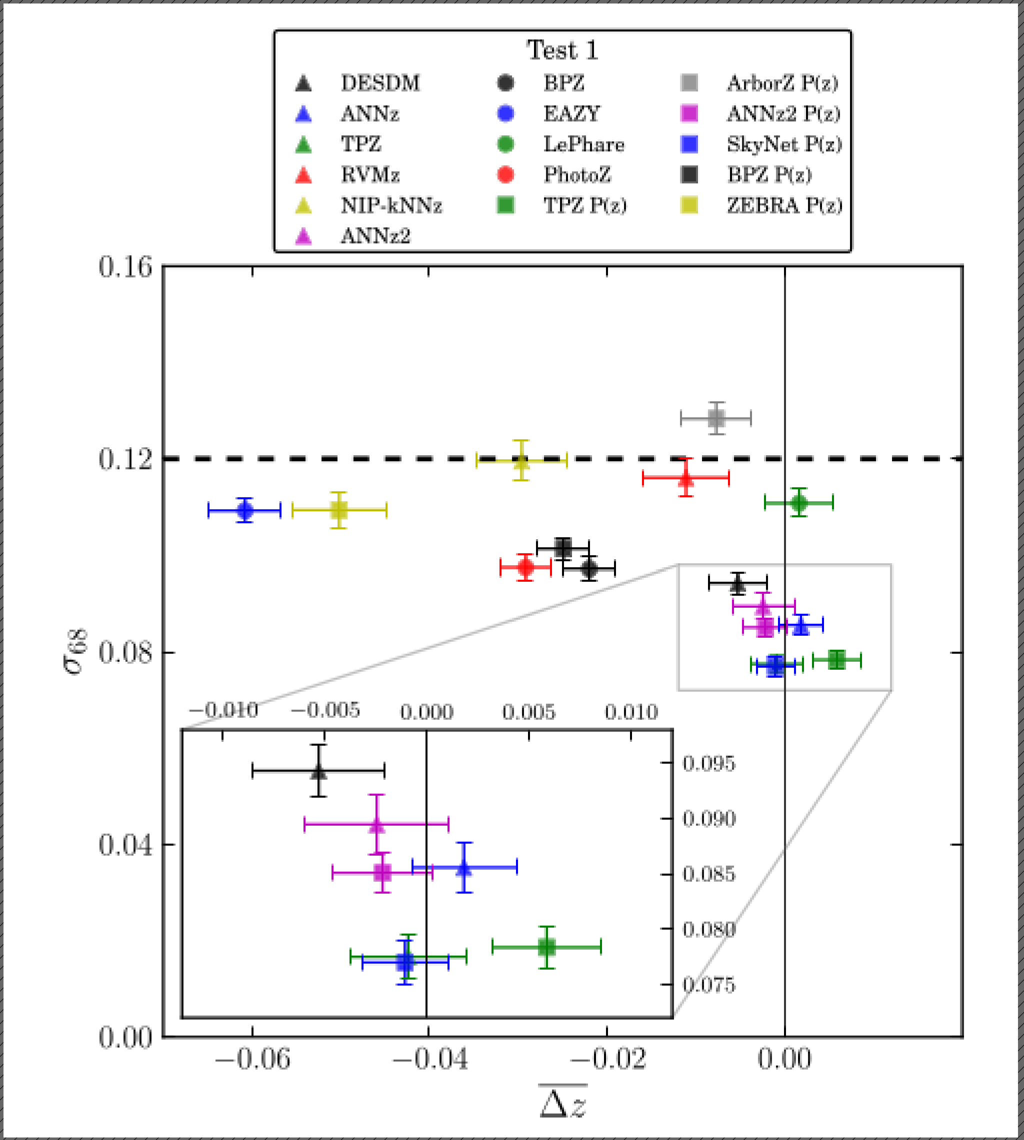
Why are template codes not performing as well?

Same information per galaxy is available.

To understand this, we need to look at where failures are occurring

In the data-rich regime template photo-z do very well.

→ photo-z driven by galaxy evolution, no previous need for vast improvements.



Sanchez et al. (2014)

Caveat for the whole of photo-z (but ML in particular)

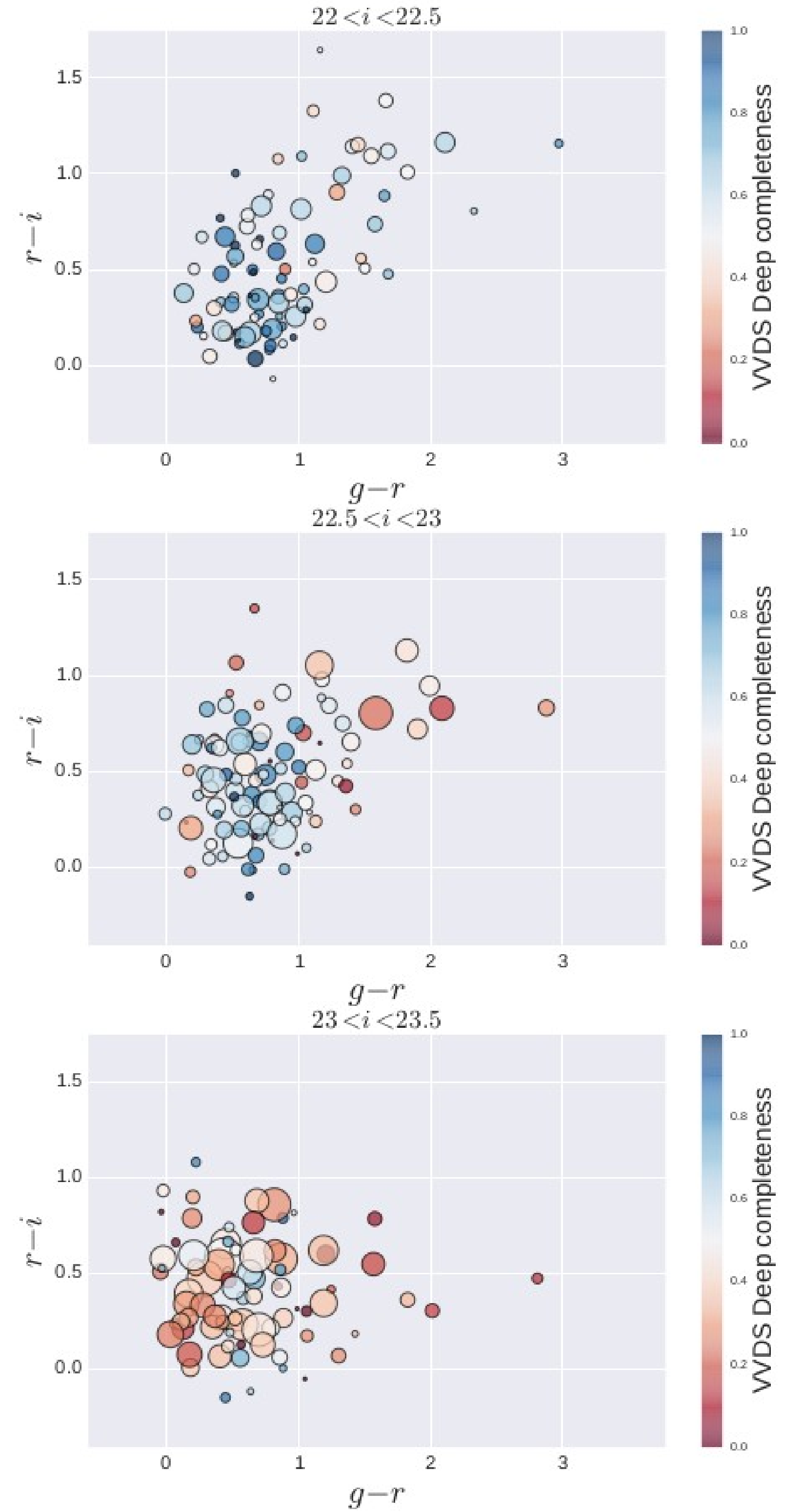
– spectroscopic incompleteness.

There are presently NO complete samples fainter than i=20.

Cannot currently differentiate effects of incompleteness from sample variance on the scale of a 1.5 sq deg field.

True validation of photo-z for DE with spectra, as it stands, is essentially impossible.

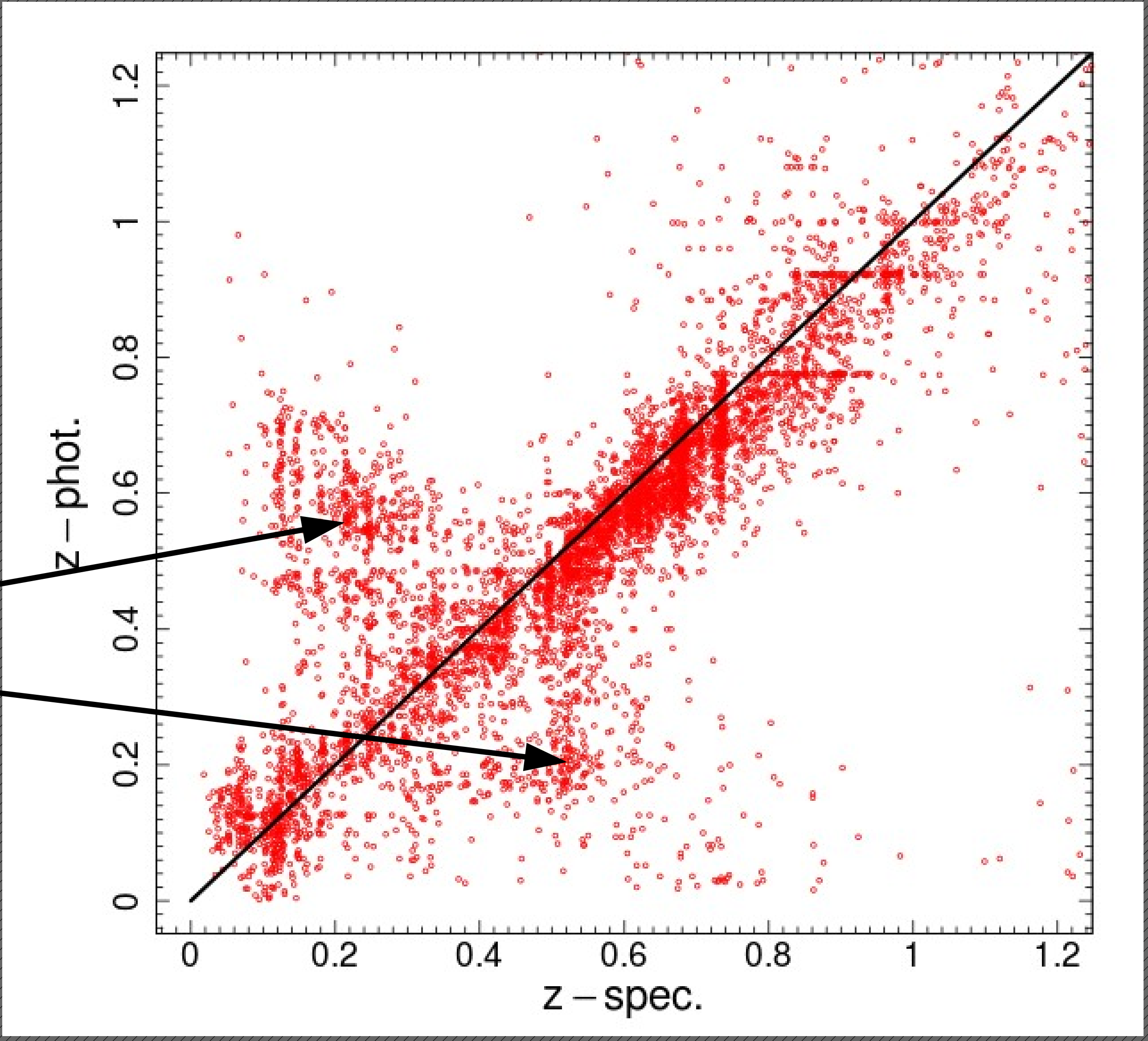
Modelling approaches to photo-z will have to work for future surveys!



Bonnett, Troxel, Hartley et al. (2015)

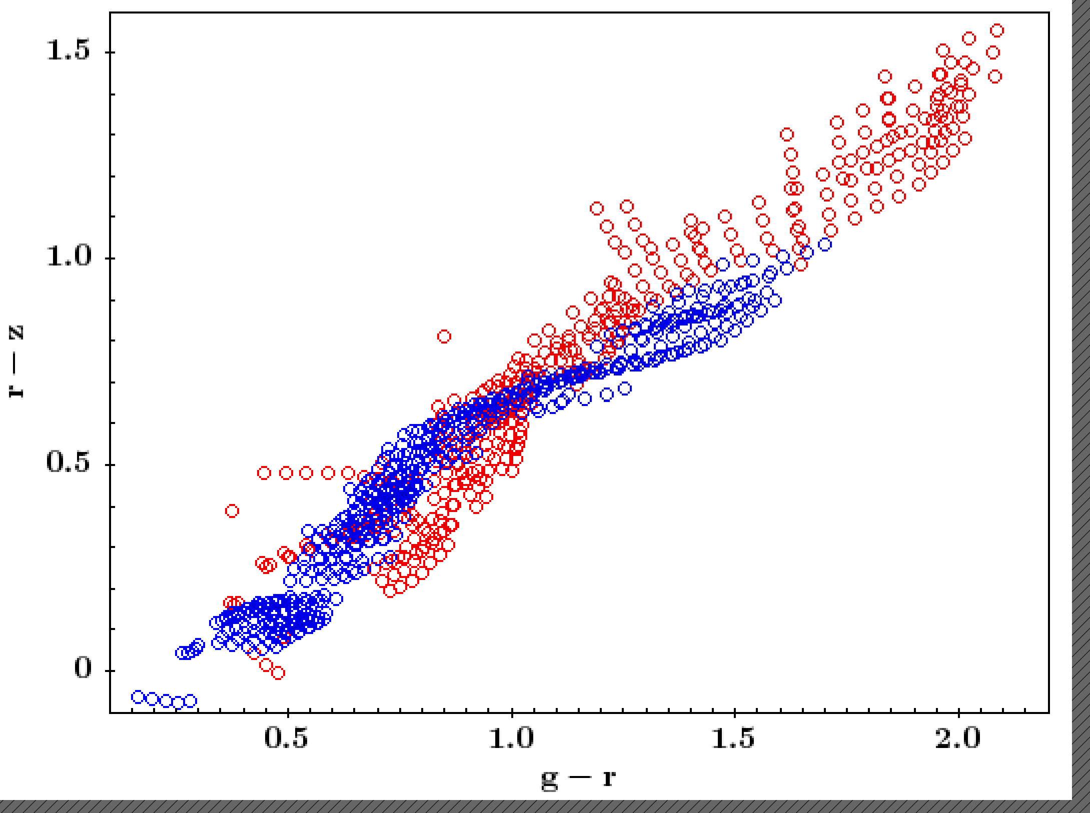
Where are template codes failing?

Raw χ^2 – i.e. uninformative prior – with CWW (empirical) templates



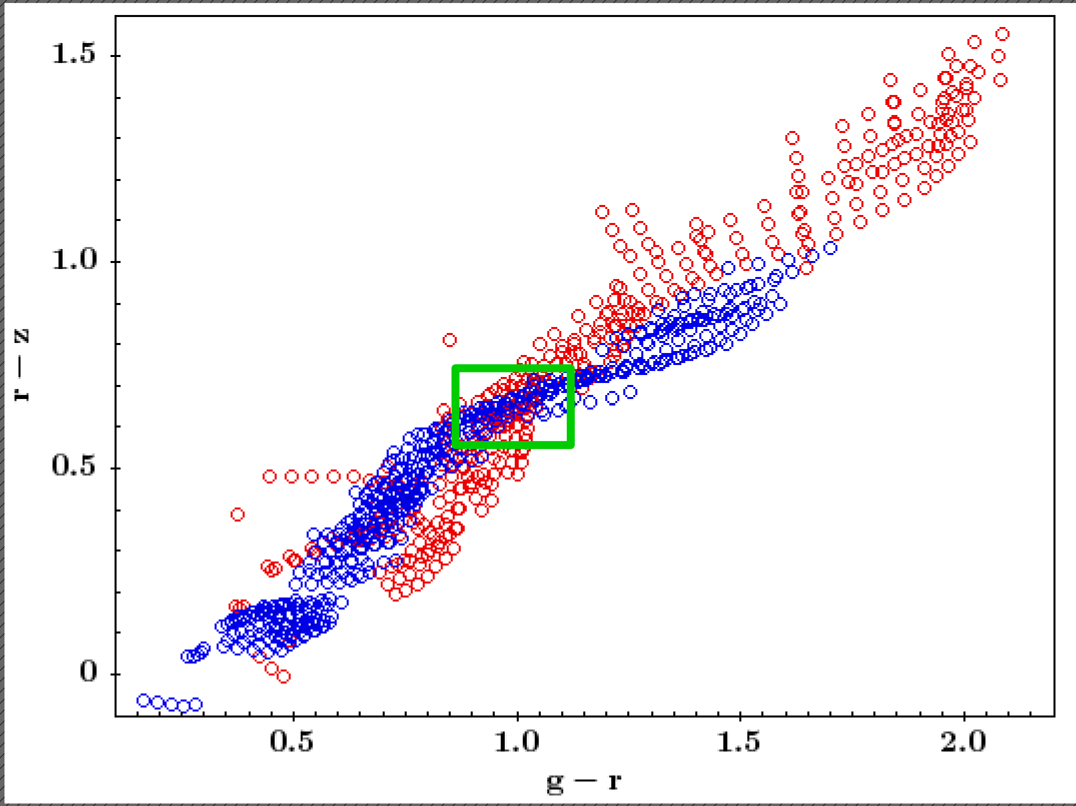
Oh dear!

A clear redshift degeneracy in templates colours...

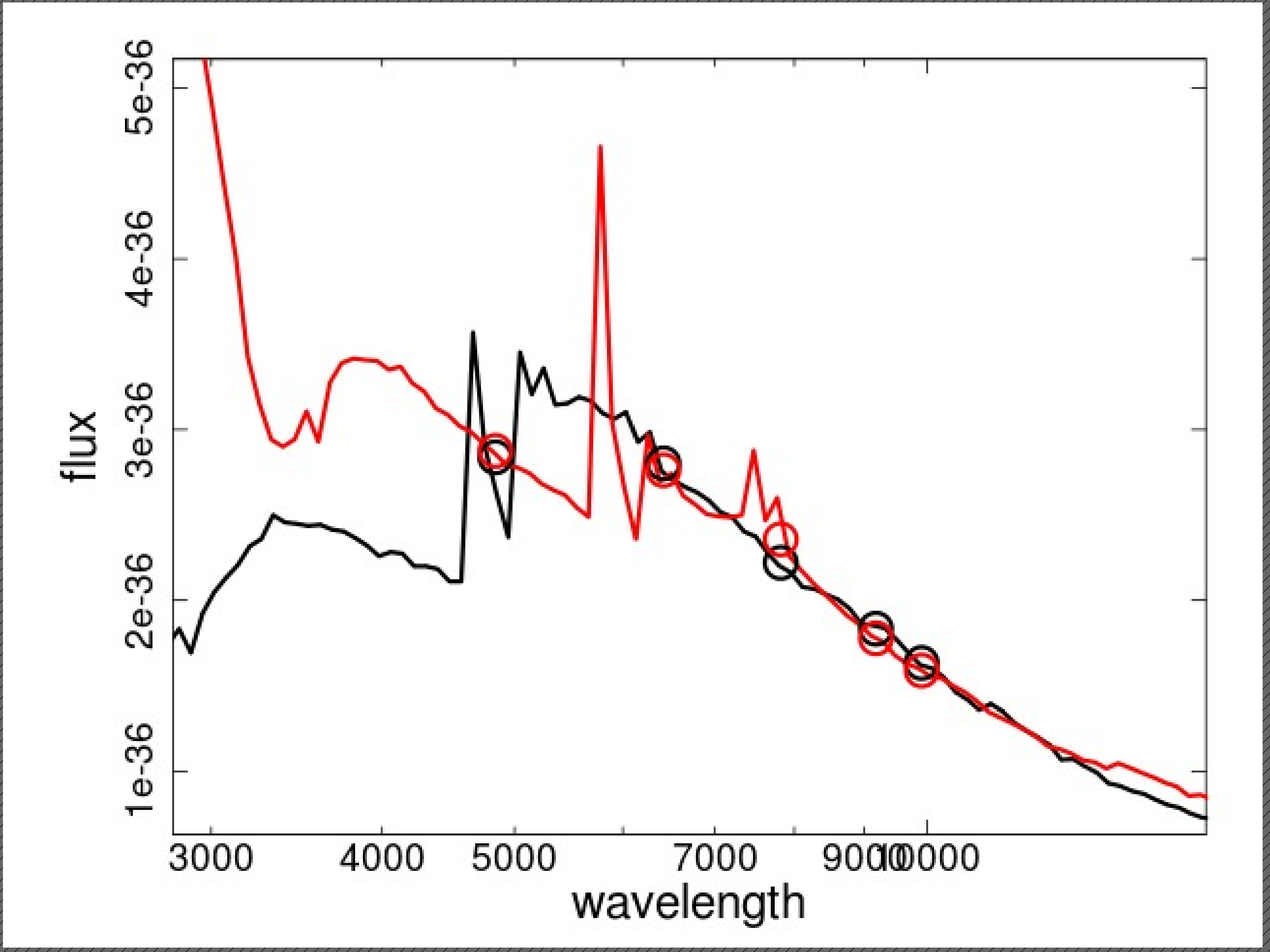


Blue: z ~ 0.25
Red: z ~ 0.55

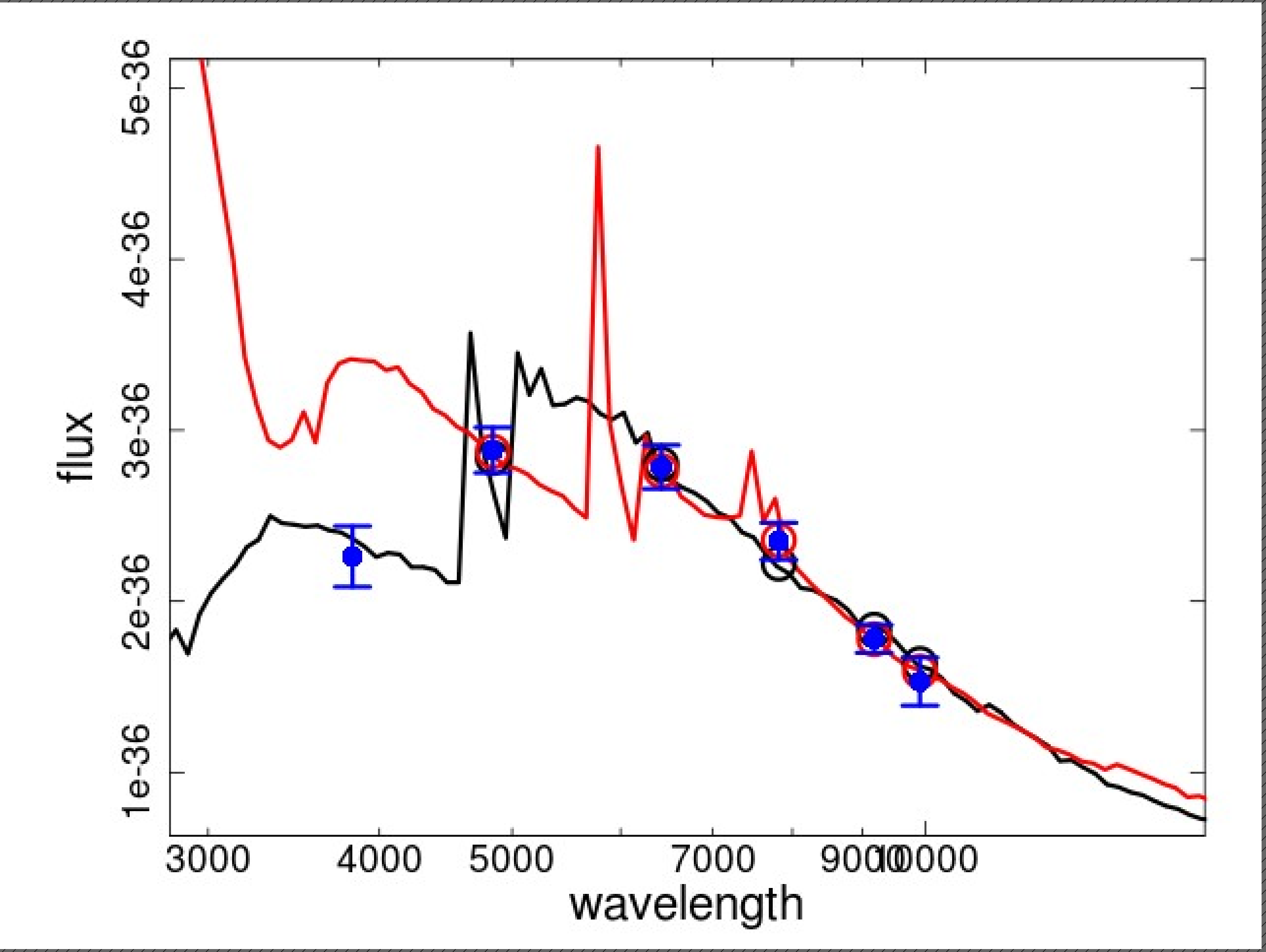
A clear redshift degeneracy in templates colours...



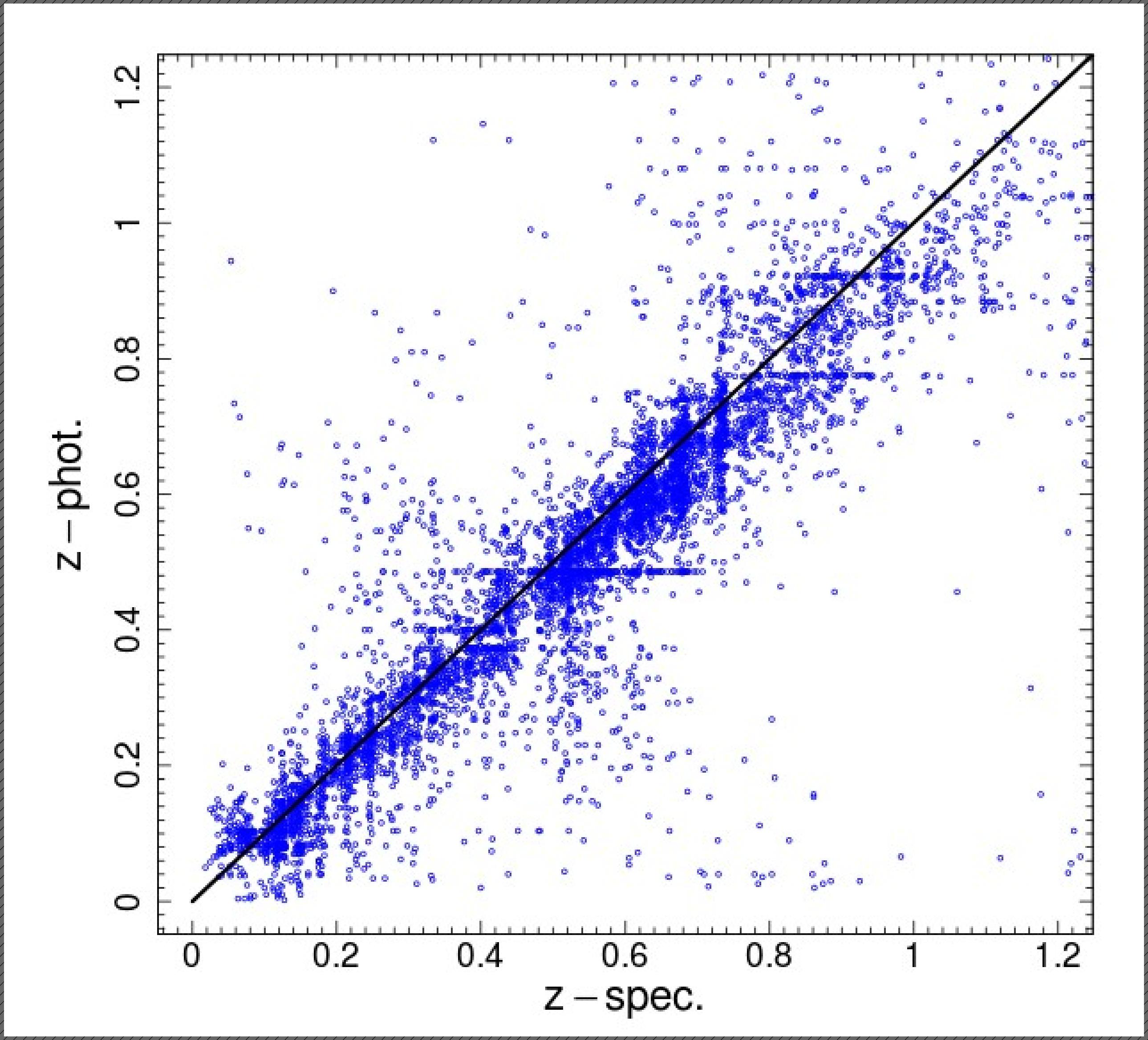
Blue: z ~ 0.25
Red: z ~ 0.55


5e-36
4e-36
3e-36
2e-36
1e-36
flux
3000
4000
5000
7000
9000
10000
wavelength

If only we had u-band data....

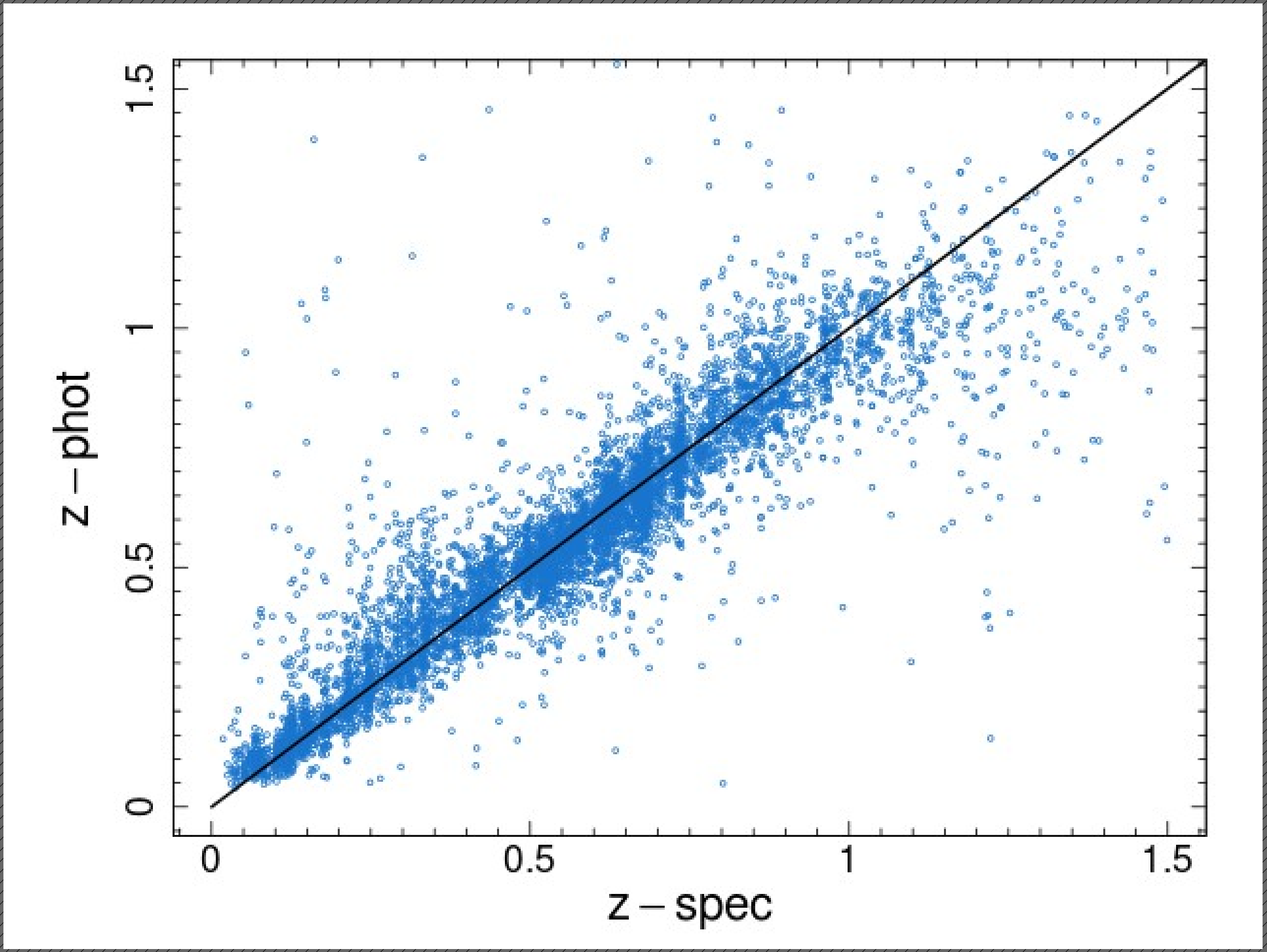


Lost without u? (Raw χ^2 – i.e. uninformative prior, but with u-band)



Lessons from Machine learning.

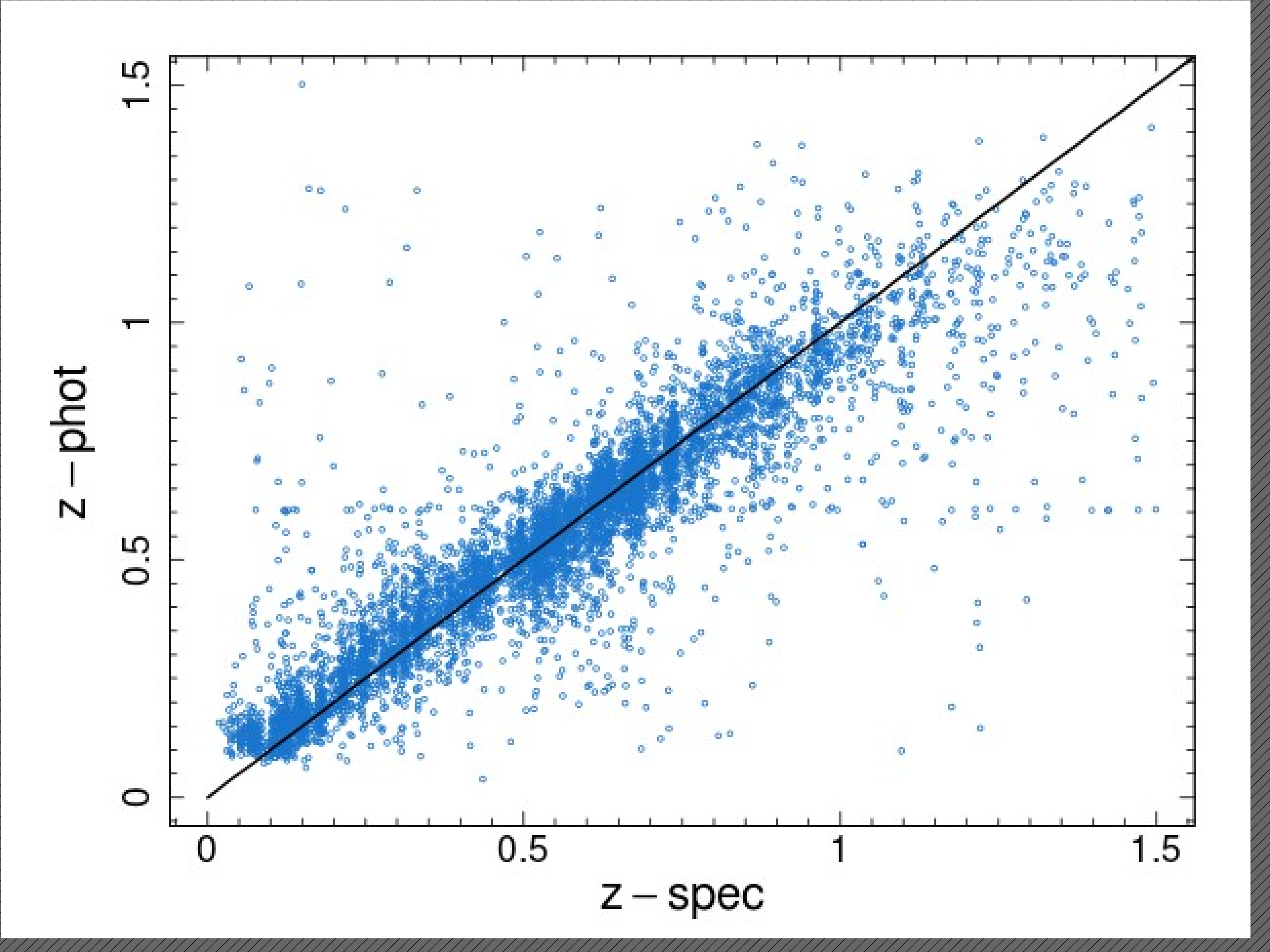
DESDM:
(H Lin)



MAG_AUTO, MAG_DETMODEL magnitudes
g, r, i, z, Y

Lessons from Machine learning.→ Not a simple case of too little information!

DESDM:
(H Lin)



MAG_AUTO colours only
g-r, r-i, i-z – plus other permutations

Most likely causes of template code failures:

1. Lack of a (suitable) prior (*poor model)

2. Photometric zero-point calibration errors

3. Unphysical templates (*poor model)

(4. The degeneracies are genuine, and we simply lack the spectra to show it in ML methods)

The dangers of 'correcting' zero-points

(“A photo-z is only as good as the photometry you have.”)

Example - The UKIDSS UDS DR8

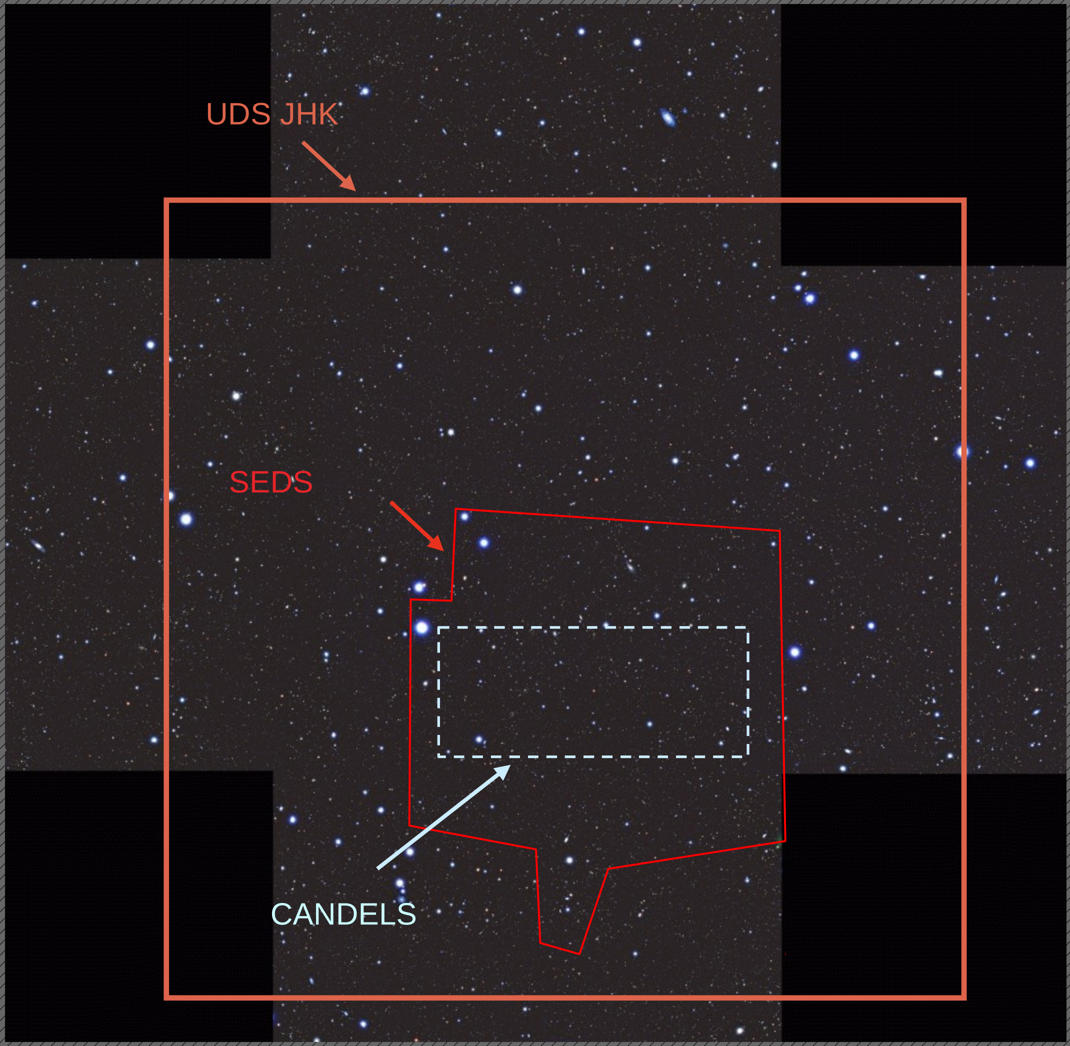
selection
K - 24.6 (5σ, AB)

u – 26.8
B – 27.6
V – 27.2
R – 27.0
i – 27.0
z – 26.0
J - 24.9
H - 24.2
Irac 1 – 24.2
Irac 2 – 24.0

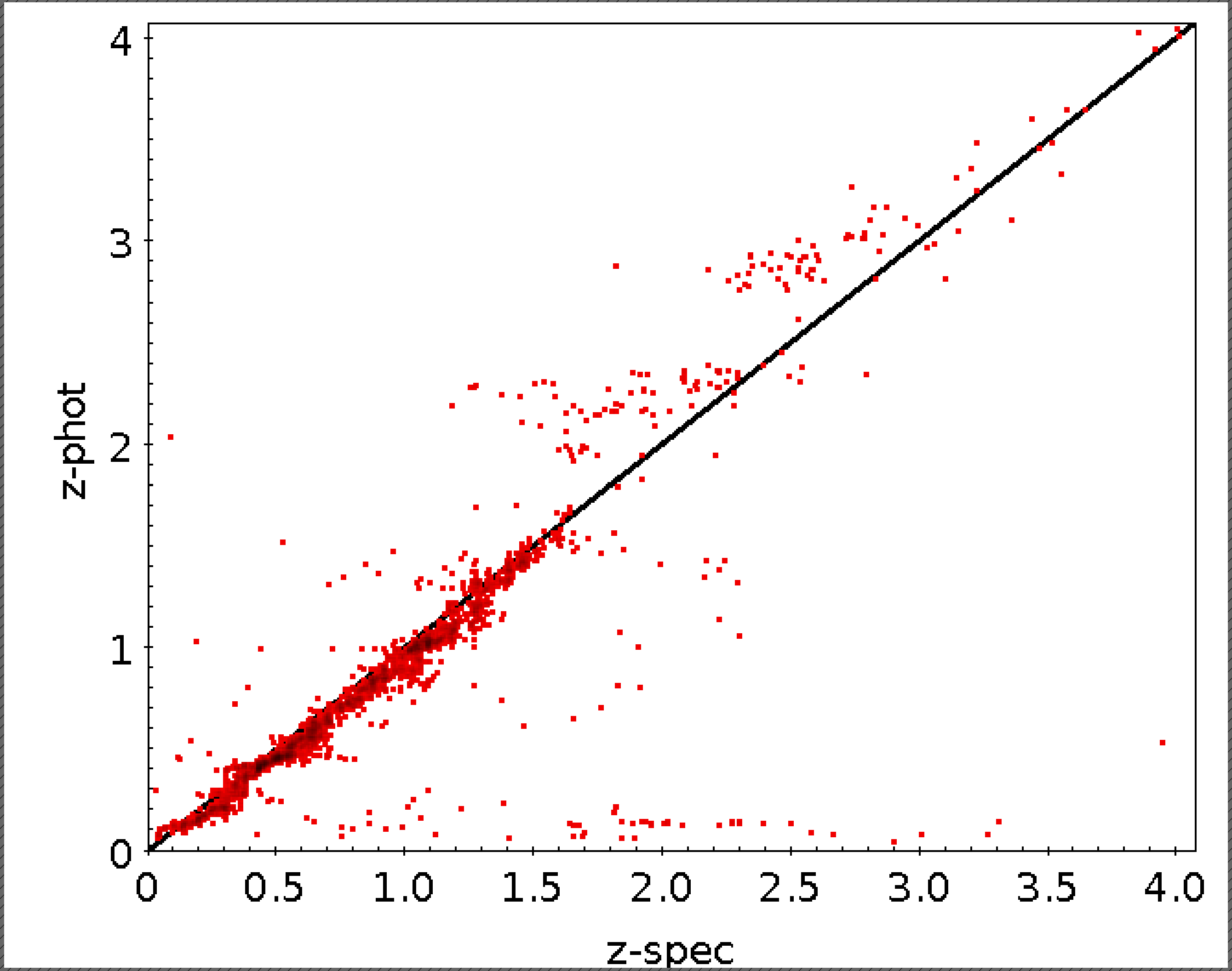
Also:
X-ray
24um
sub-mm (HerMES, SCUBA-2)
Radio

>2000 high confidence spec-z
(largely unbiased, inc. UDSz)

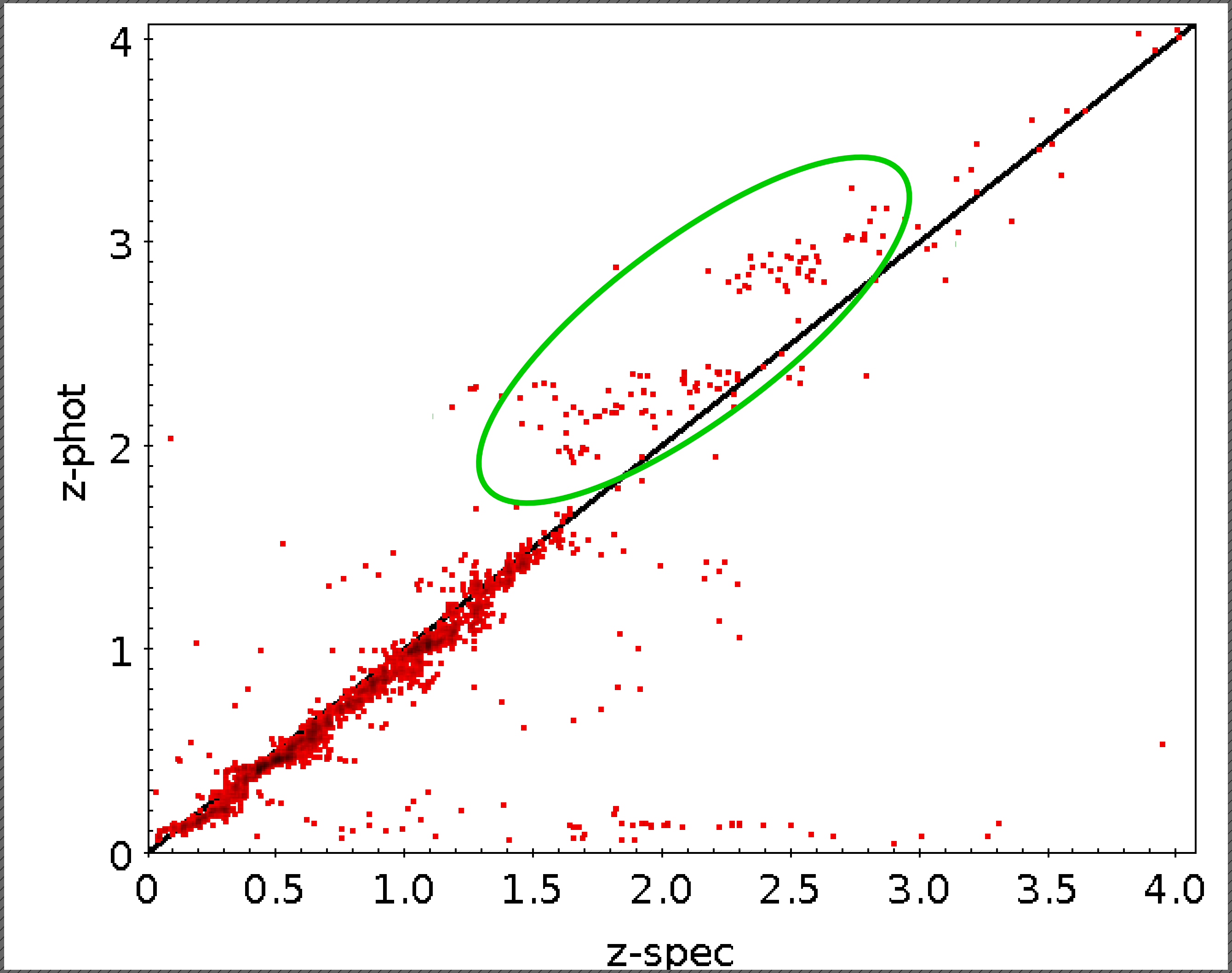
Field size: 0.62 sq. deg.



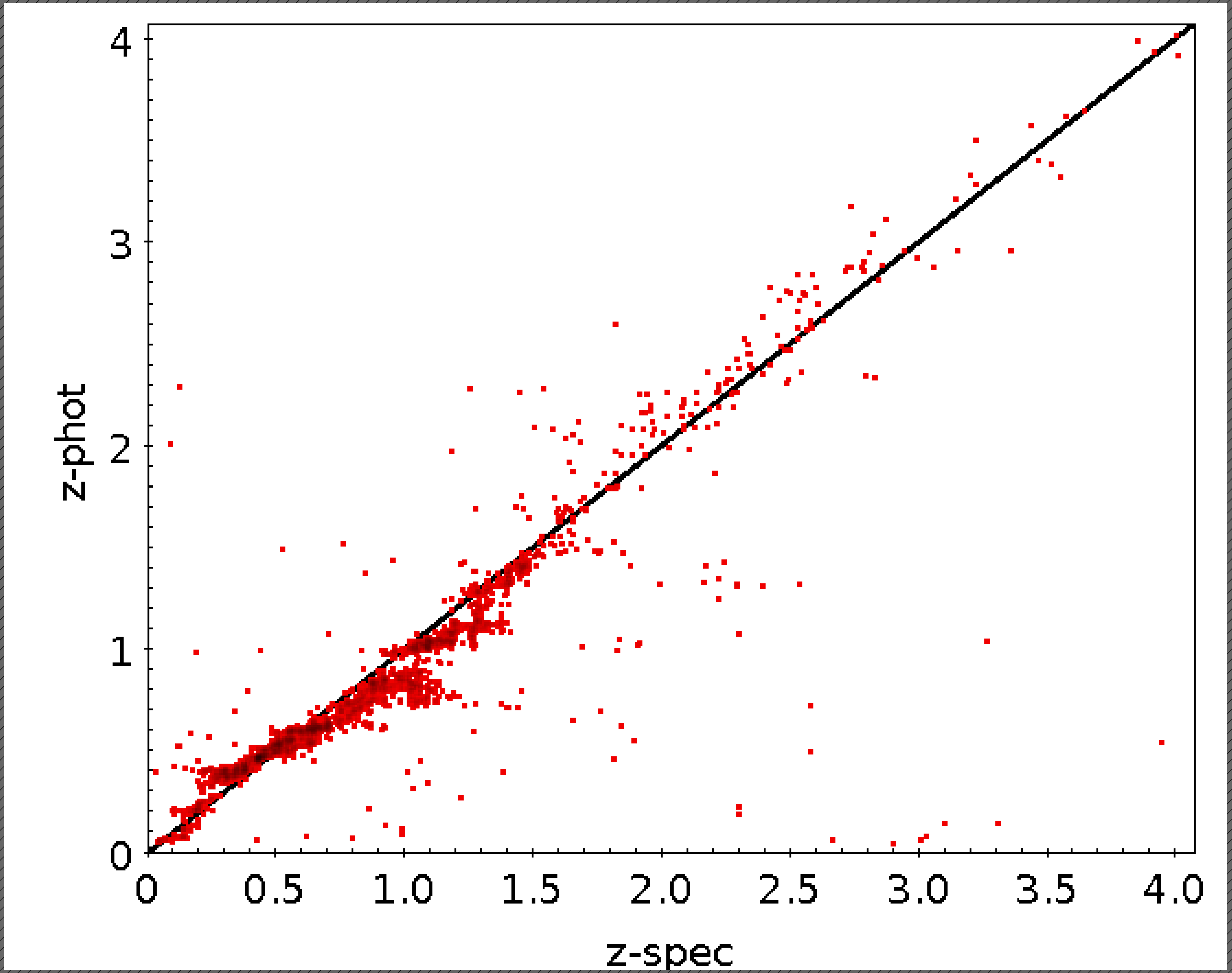
Out of the box EAZY redshifts (linear combination of 6 PCA components)



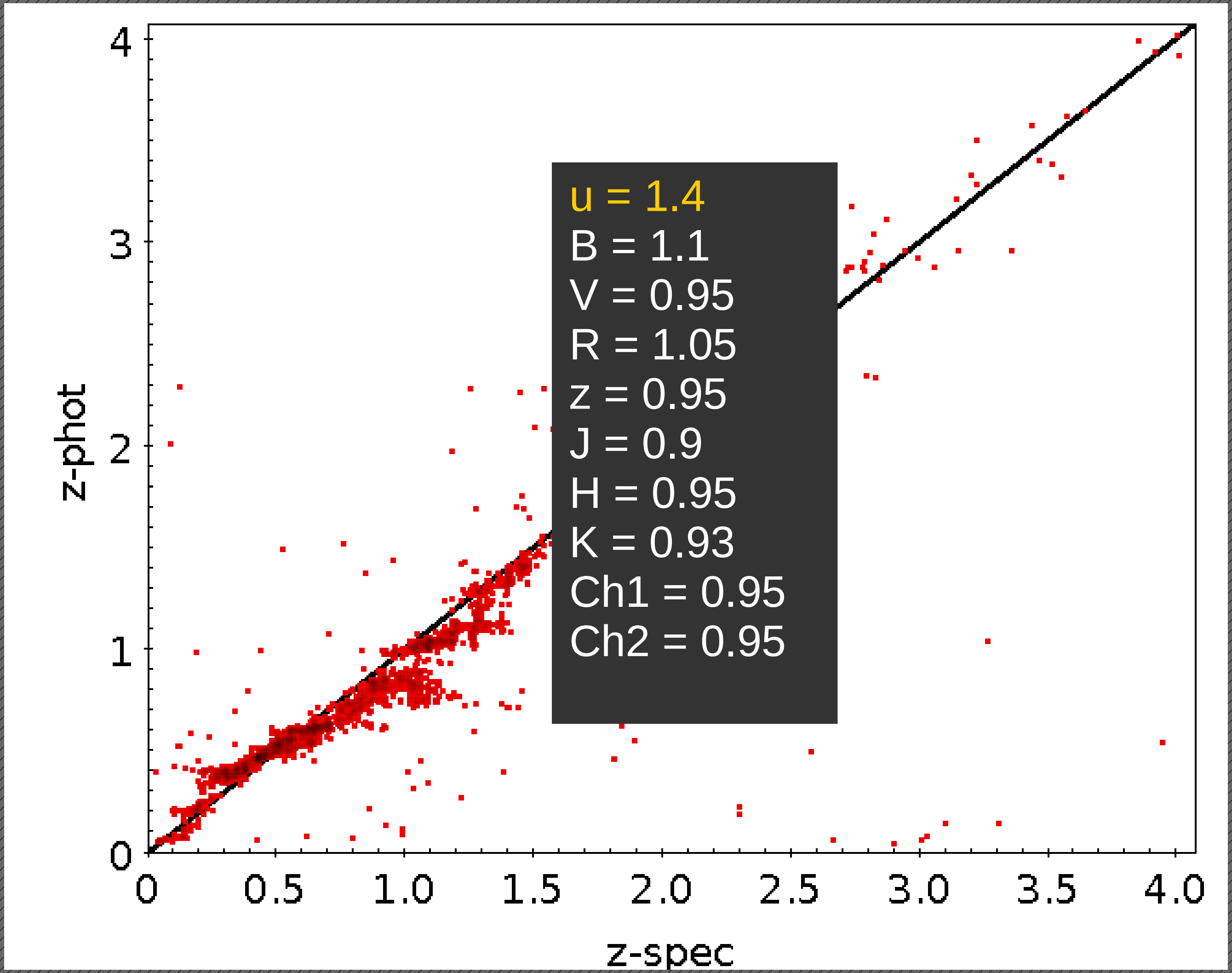
Out of the box EAZY redshifts (linear combination of 6 PCA components)



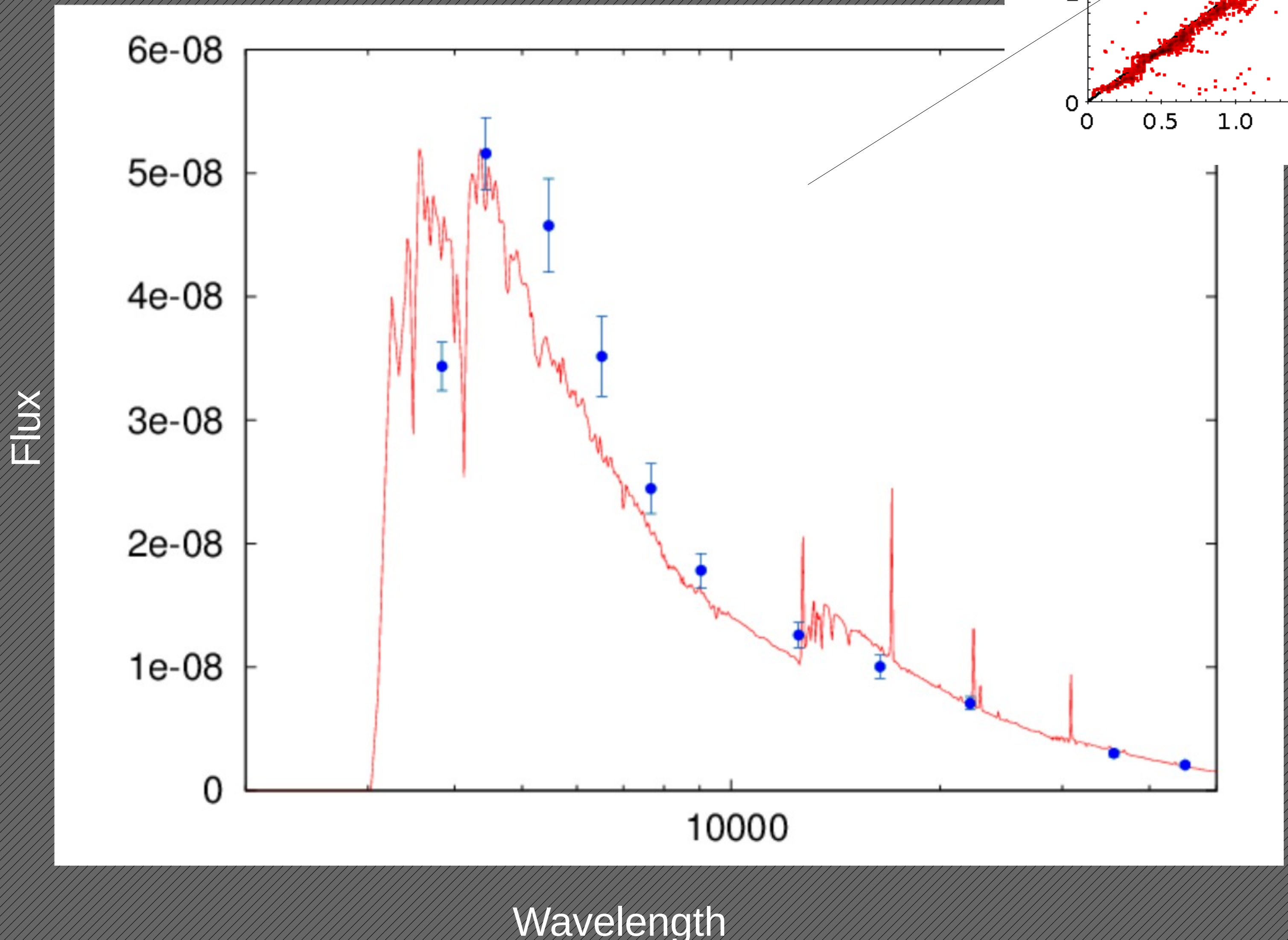
Applying the zero-point fiddle.

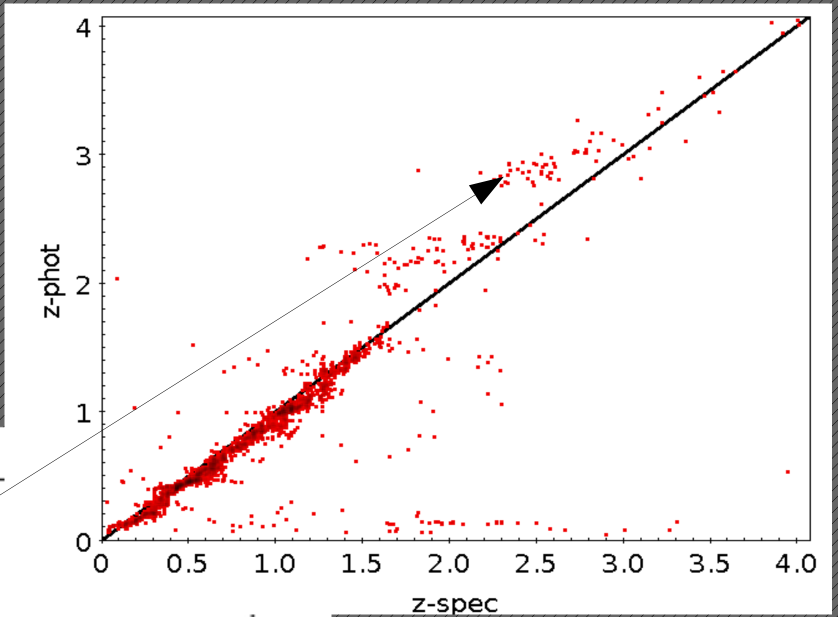


Applying the zero-point fiddle.
(“A photo-z is only as good as the photometry you have.”)

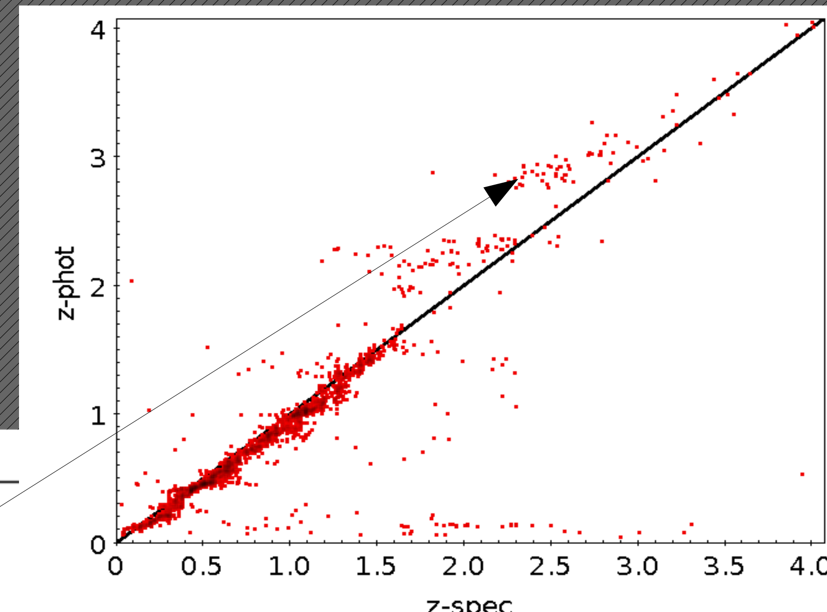


Best fit at fixed redshift, z = z-spec.



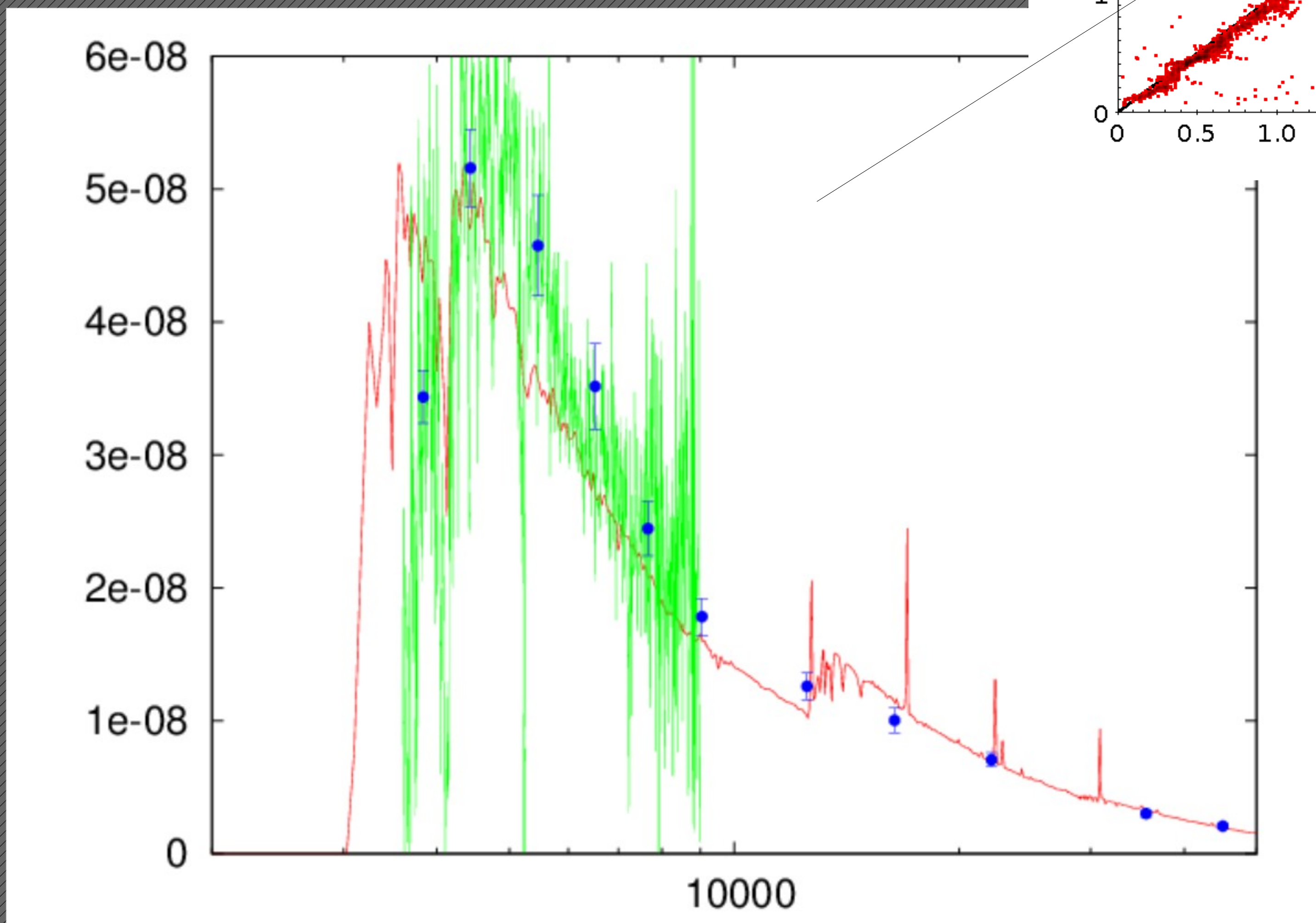


Best fit at fixed redshift, z = z-spec.



Green: UDSz spectrum

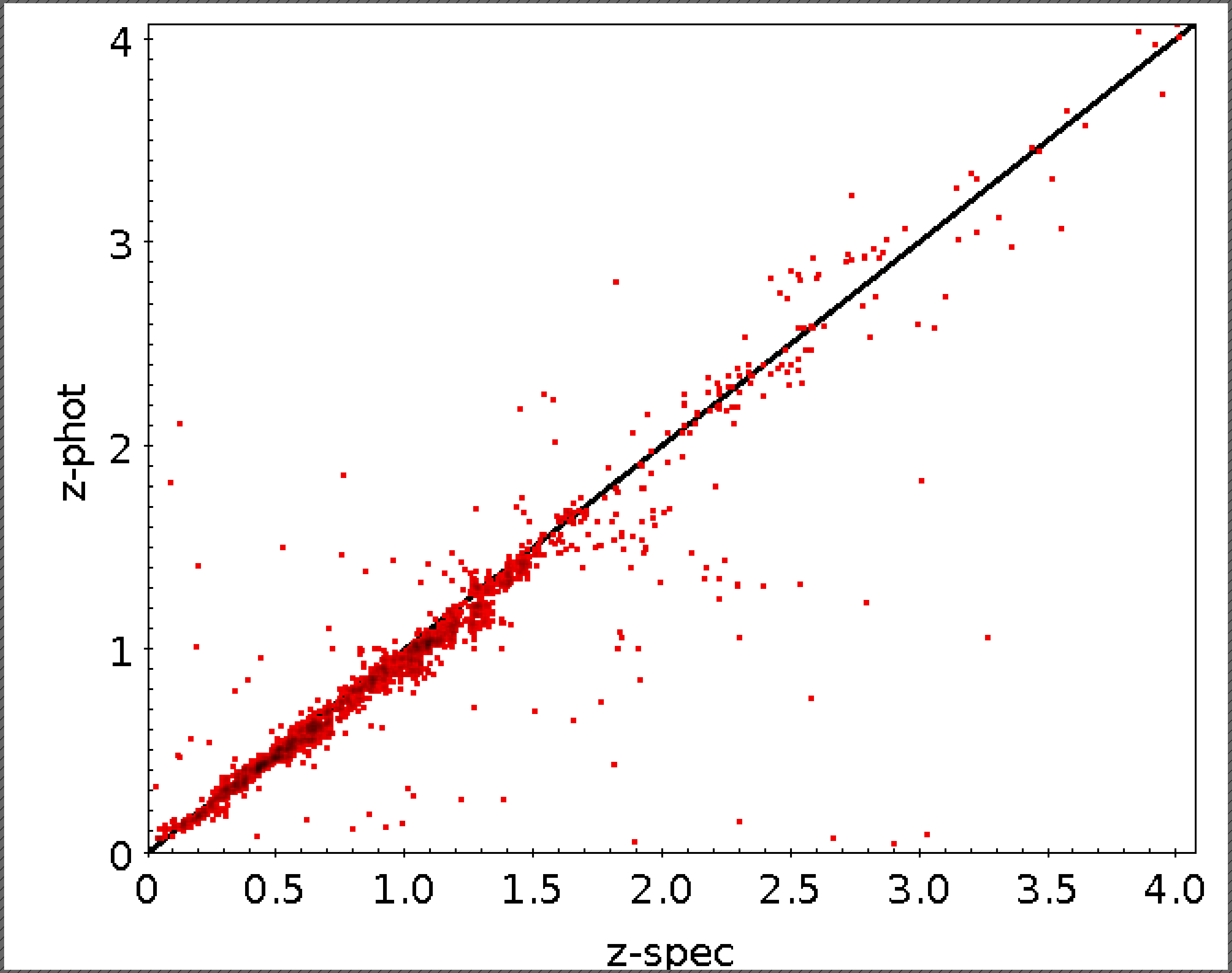
Flux



Wavelength

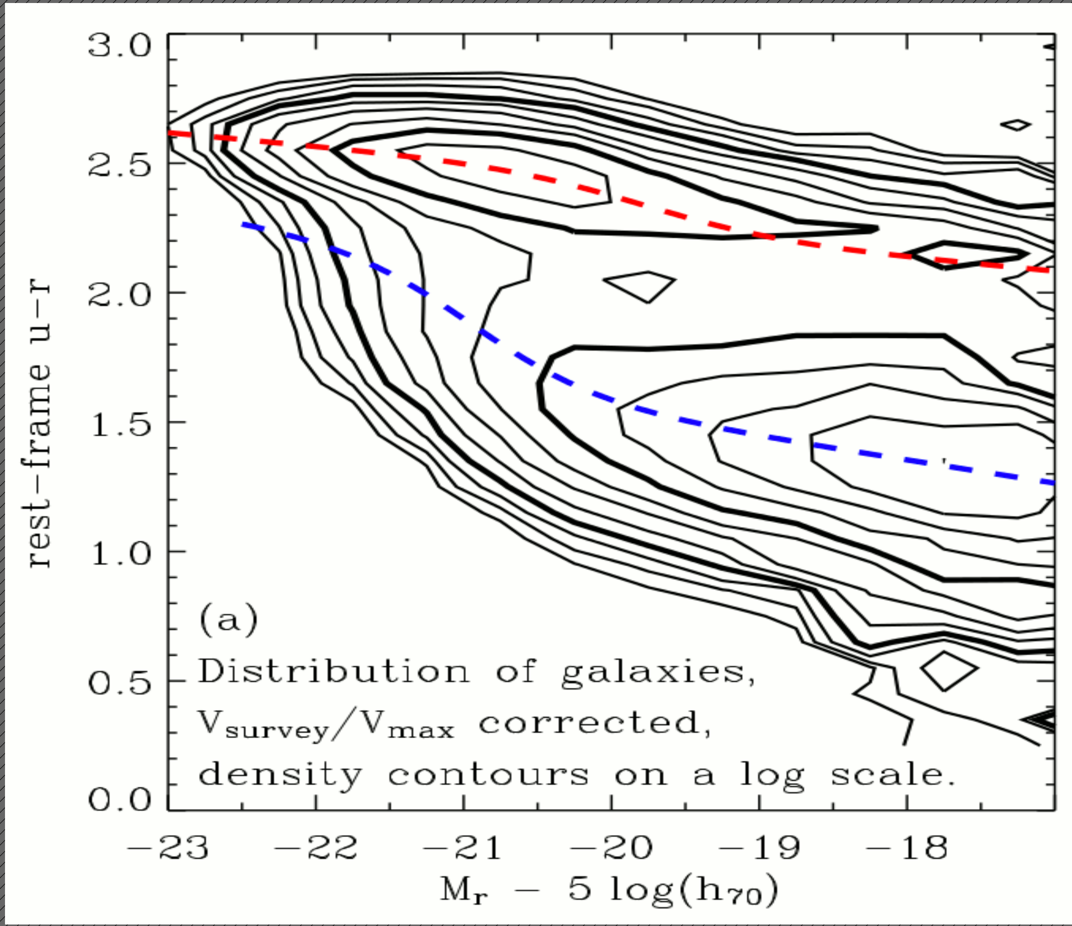
Create a new blue template with a steep dust law (SMC)
→ Much better!

u = 1.15



The current state of model photo-z in the Dark Energy Survey

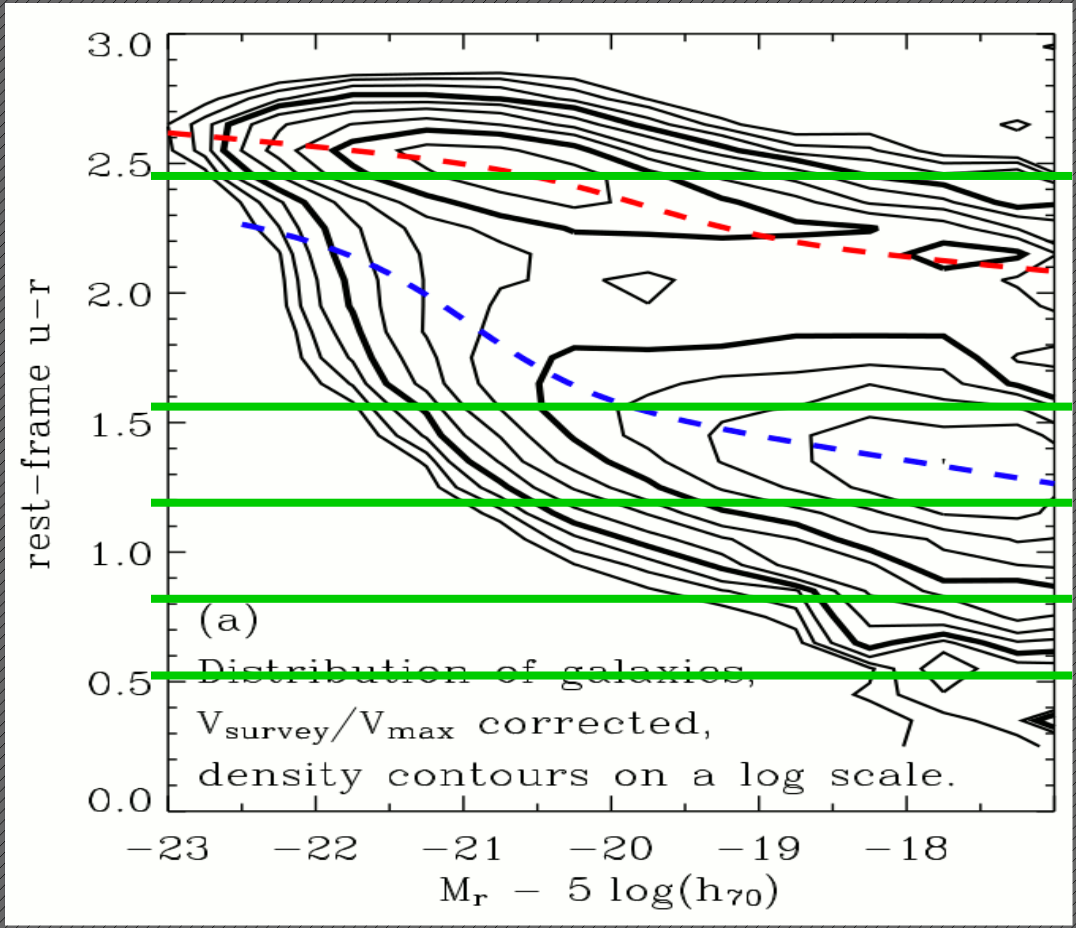
The z=0 galaxy population in the rest-frame...



Baldry et al. (2004)

The z=0 galaxy population in the rest-frame...

Templates are (typically) fixed (4D) colours in the rest frame.

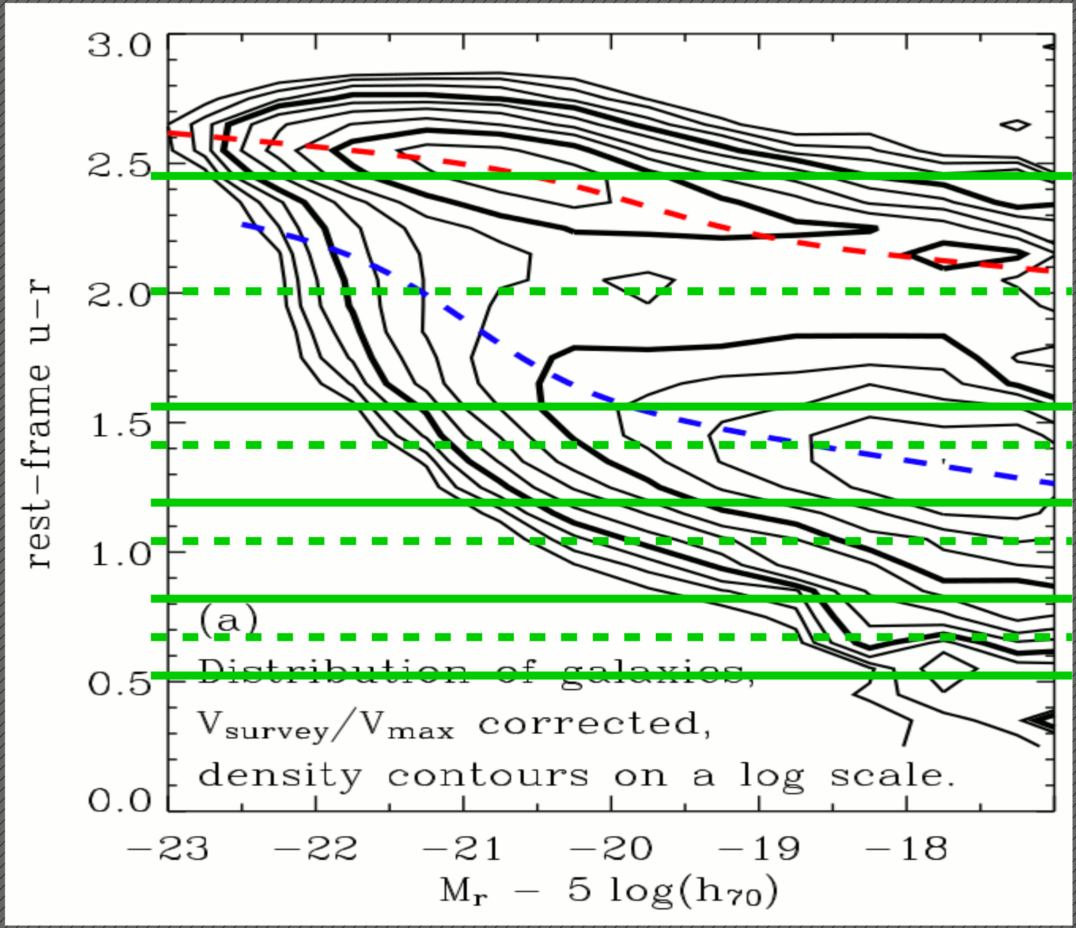


Baldry et al. (2004)

The z=0 galaxy population in the rest-frame...

Templates are (typically) fixed (4D) colours in the rest frame.

Maybe with interpolation.



Baldry et al. (2004)

The z=0 galaxy population in the rest-frame...

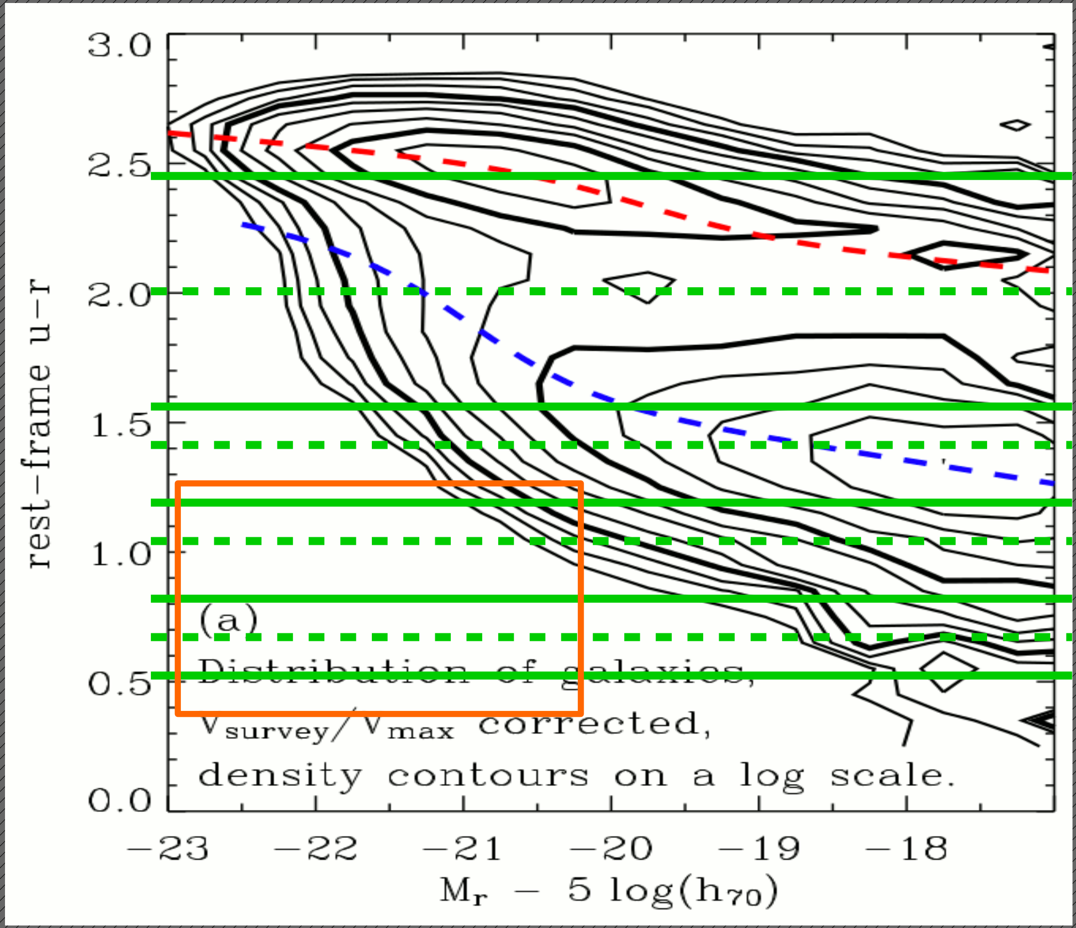
Templates are (typically) fixed (4D) colours in the rest frame.

Maybe with interpolation.

Sampling of the allowed colour-luminosity space is fairly low.

While there are other regions allowed but do not contain galaxies.

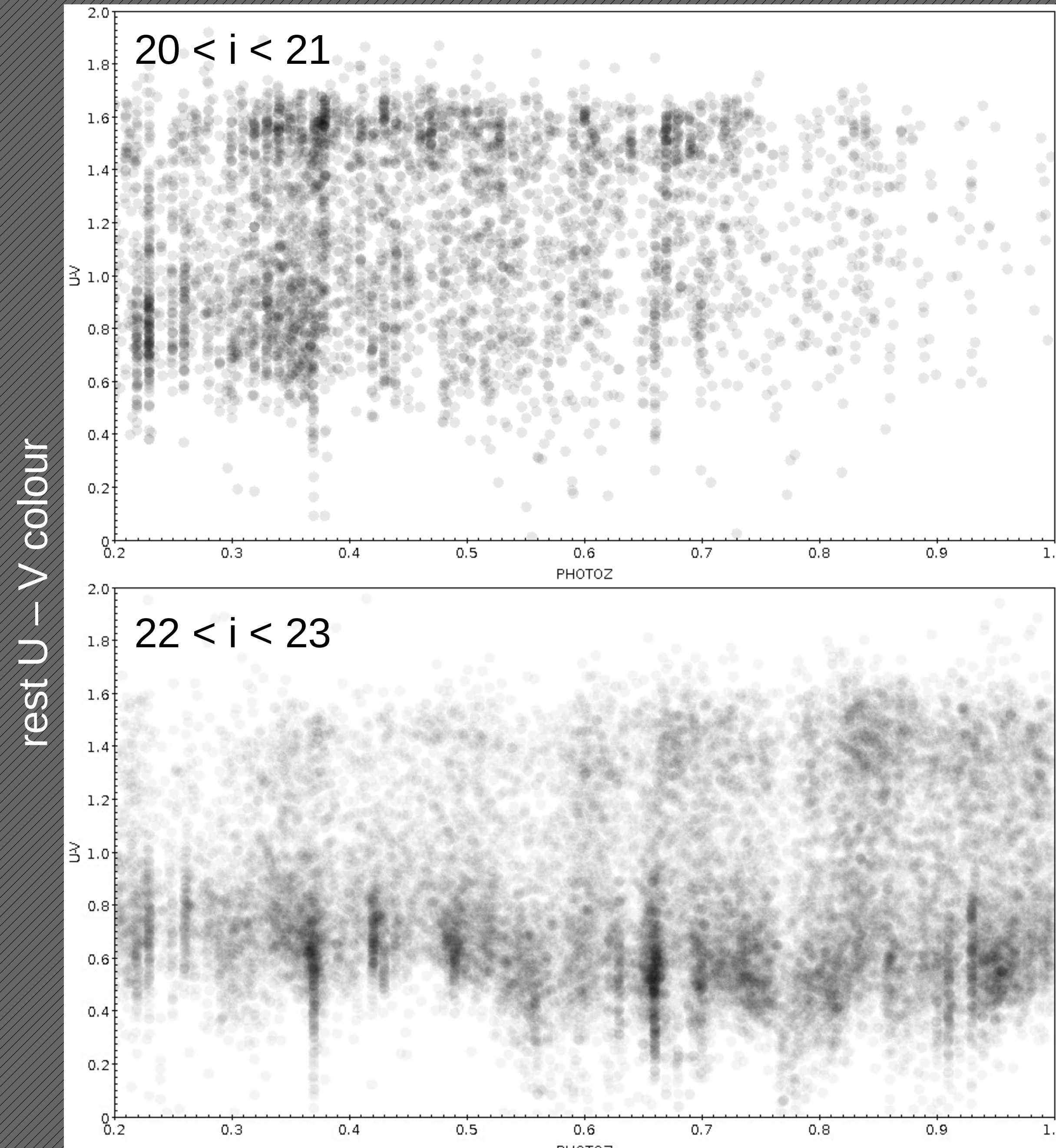
(Not a major problem in data-rich regime)



Baldry et al. (2004)

In the data...

Within the space of just two magnitudes the galaxy distribution in redshift and colour is almost inverted.

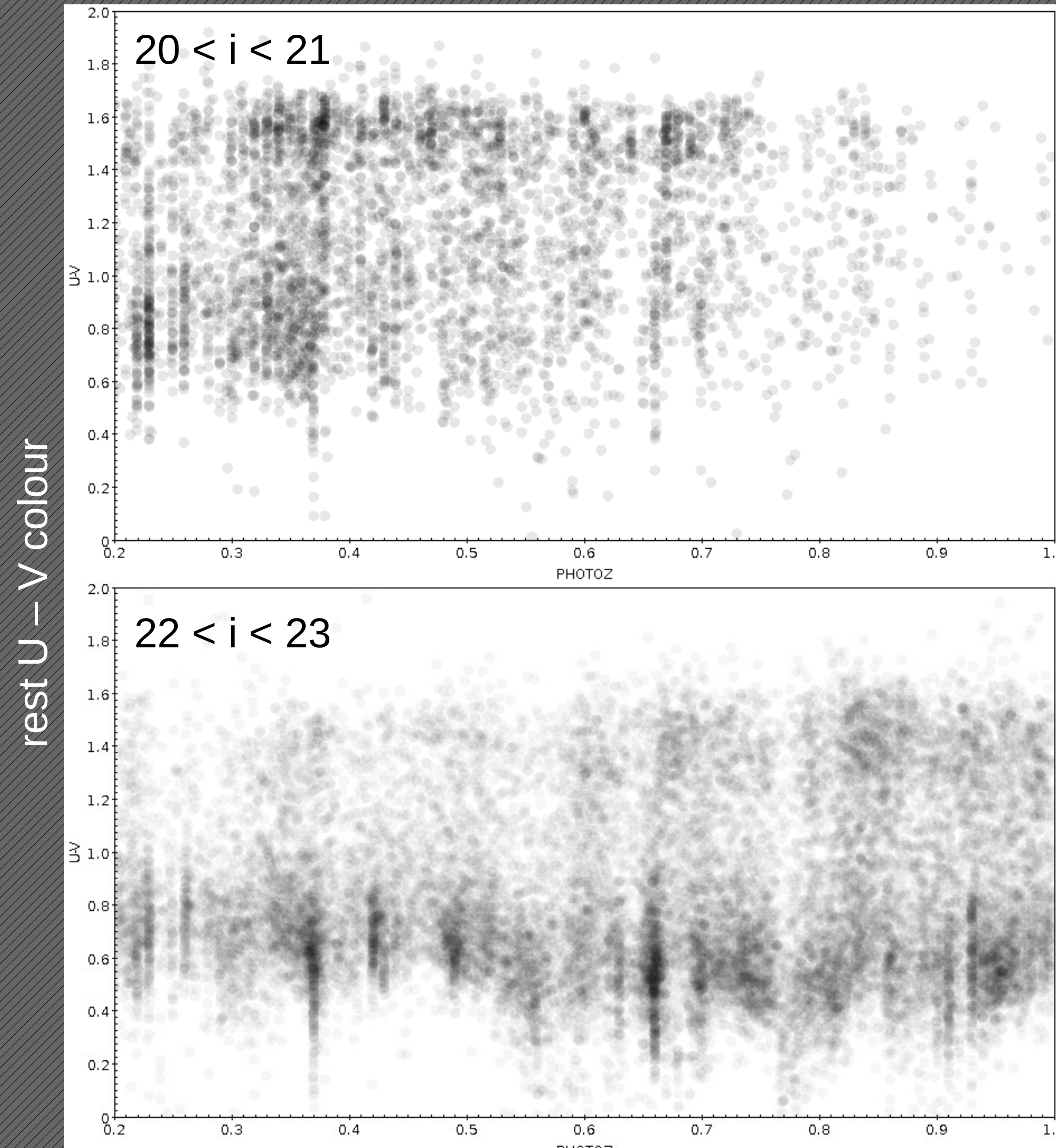


Cosmos phot-z (Laigle et al. 2015)

In the data...

Within the space of just two magnitudes the galaxy distribution in redshift and colour is almost inverted.

Training methods capture 'forbidden' regions automatically.



Cosmos phot-z (Laigle et al. 2015)

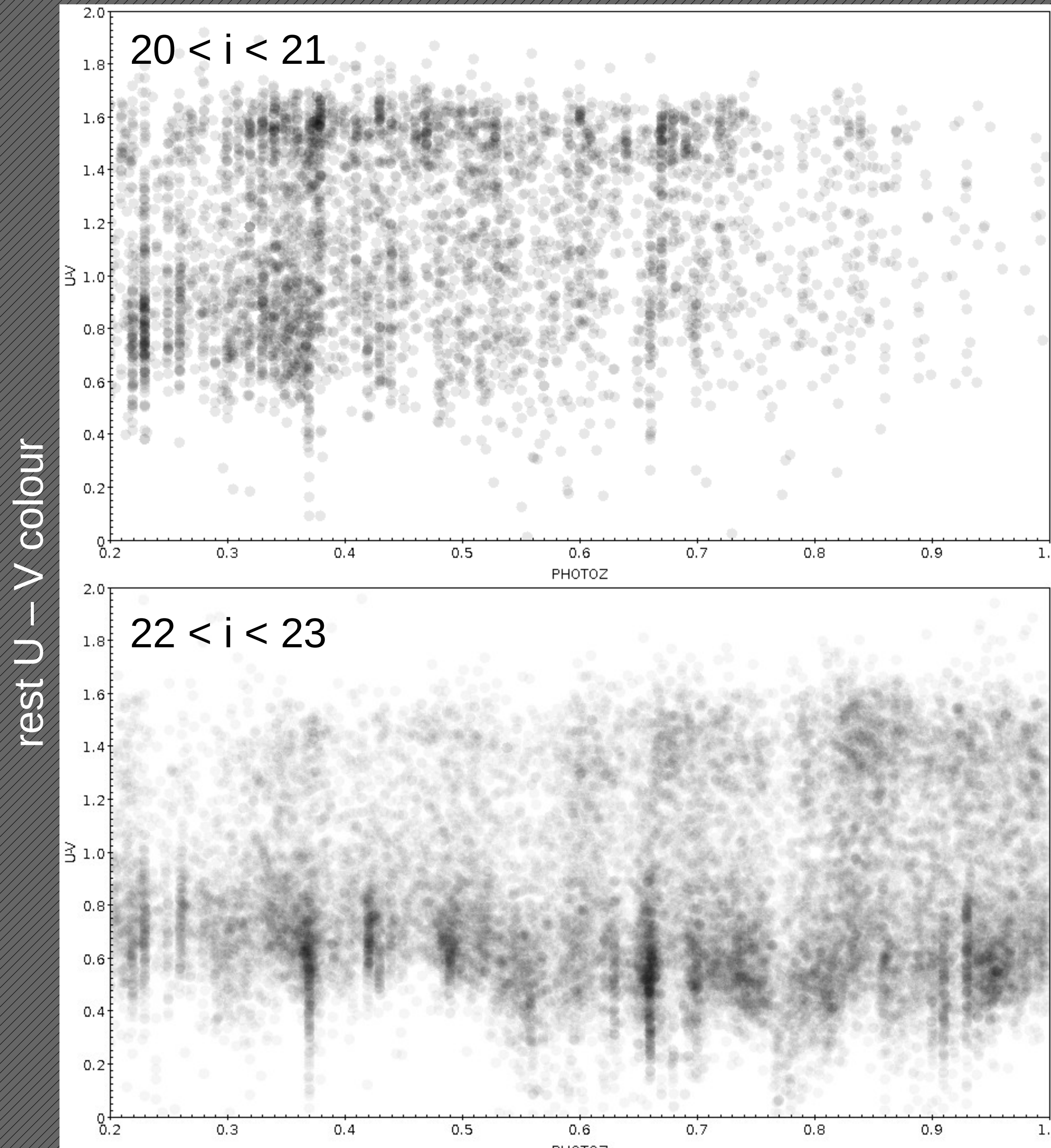
In the data...

Within the space of just two magnitudes the galaxy distribution in redshift and colour is almost inverted.

Training methods capture 'forbidden' regions automatically.

How can we do this in a template code?

→ Use a prior
(i.e. a model of the galaxy population, not just individual objects)



Cosmos phot-z (Laigle et al. 2015)

Constructing a prior...

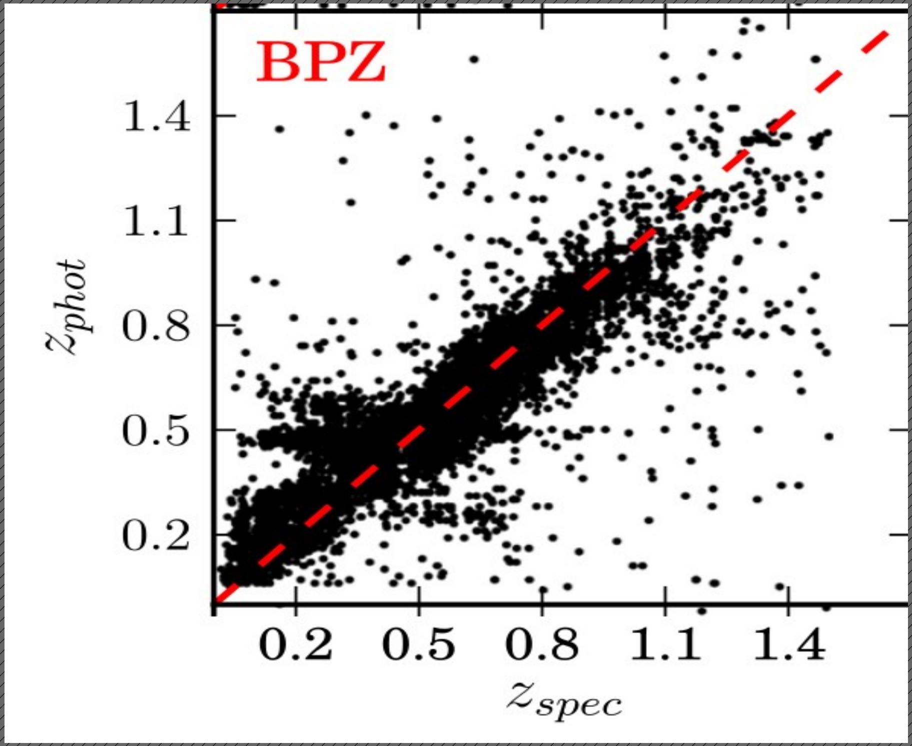
In BPZ run for DES SV:

Prior calibrated directly, using the available spec-z (→ similar weaknesses as ML)

11 parameter model, based on a mix of 3 galaxy types (elliptical, spiral, irregular)

→ mix of types as fn of observed magnitude (exponential in form)
→ each type has mag-dependent redshift probability function (also exponential)

Degenerate solutions are reduced,
but clear features remain.



Sanchez et al (2014)

Constructing a prior...

In BPZ run for DES SV:

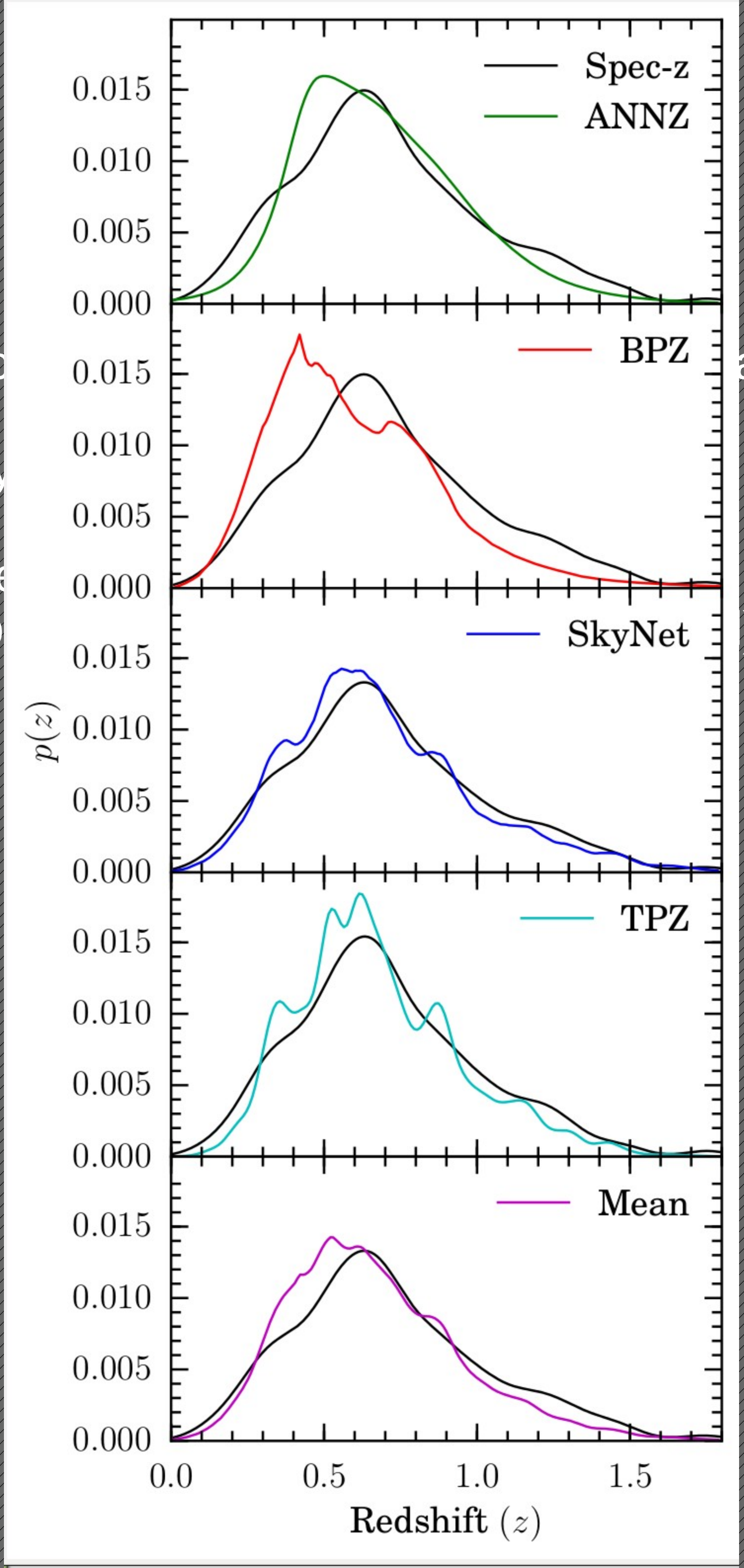
Prior calibrated directly, using the available spec as ML)

11 parameter model, based on a mix of 3 galaxy

→ mix of types as fn of observed magnitude
→ each type has mag-dependent redshift p)

Degenerate solutions are reduced,
but clear features remain.

The quantity that we need – the stacked
PDF – has a major issue!



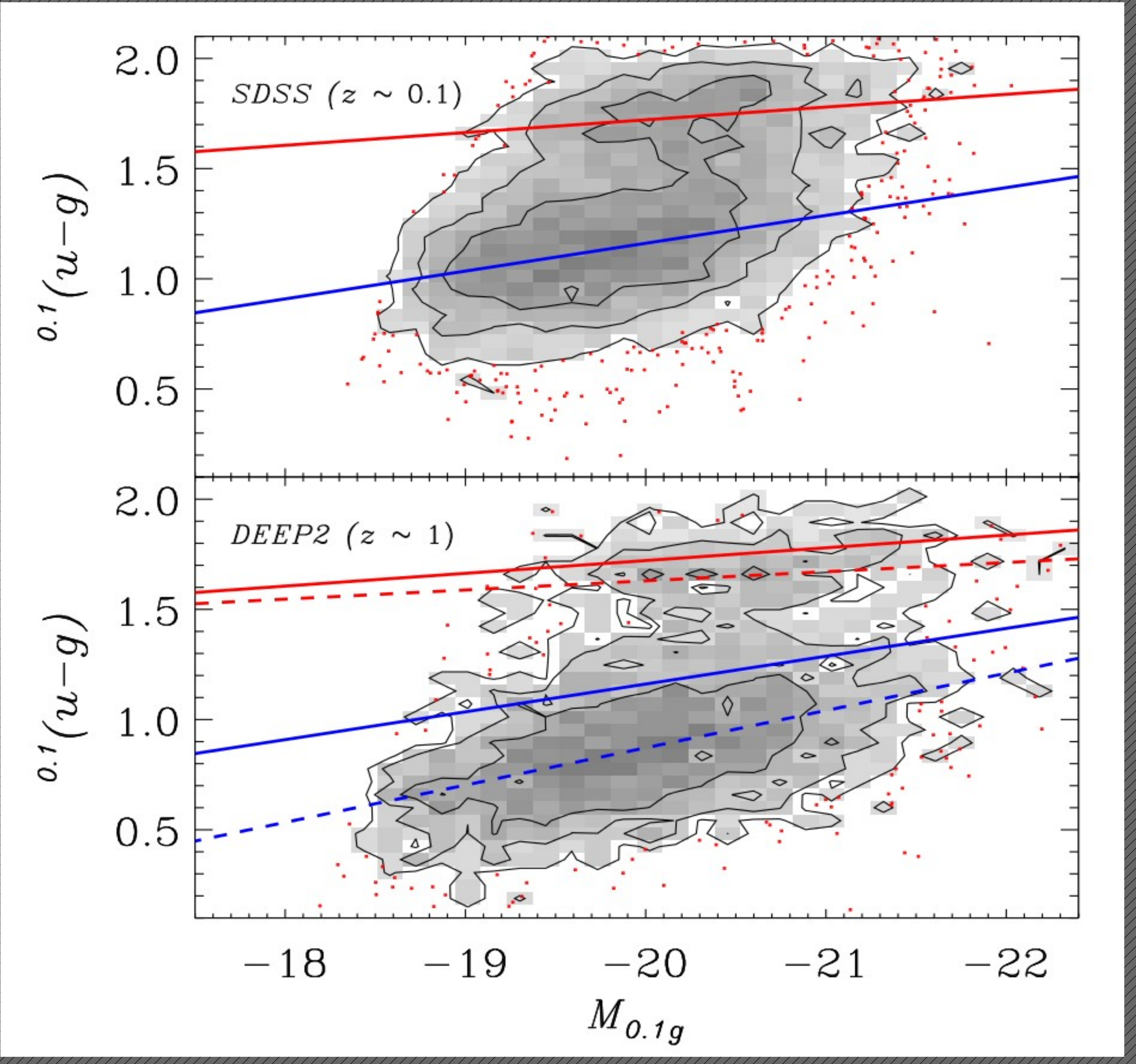
Despite including a data-driven prior, the model is still wrong!

Templates do not account for:

→ colour – luminosity dependence
for given galaxy type

→ redshift evolution of galaxies

→ incomplete coverage of
galaxy types



Blanton et al. (2006)

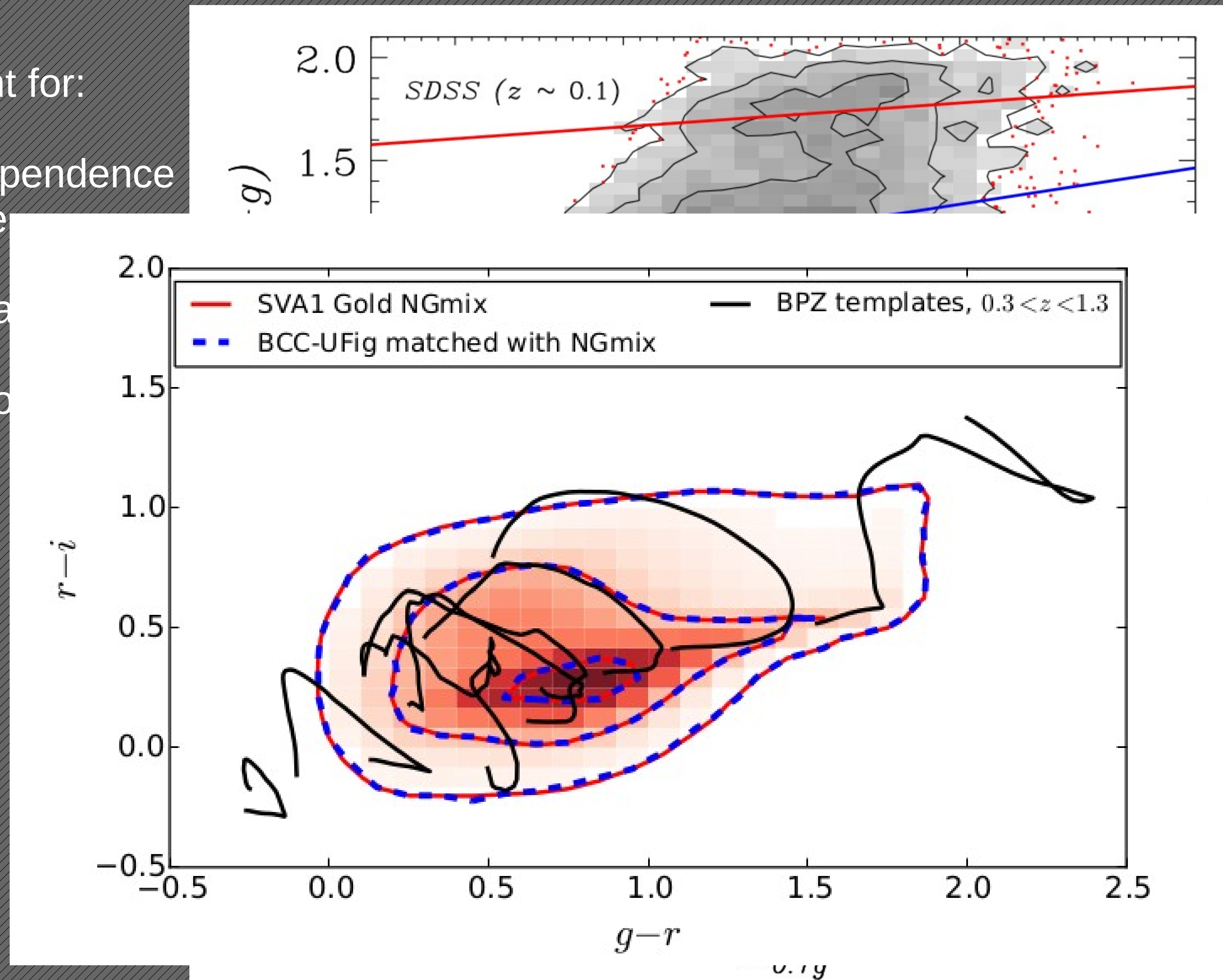
Despite including a data-driven prior, the model is still wrong!

Templates do not account for:

→ colour – luminosity dependence for given galaxy type

→ redshift evolution of ga

→ incomplete coverage o galaxy types



Bonnett, Troxel, Hartley et al (2015)

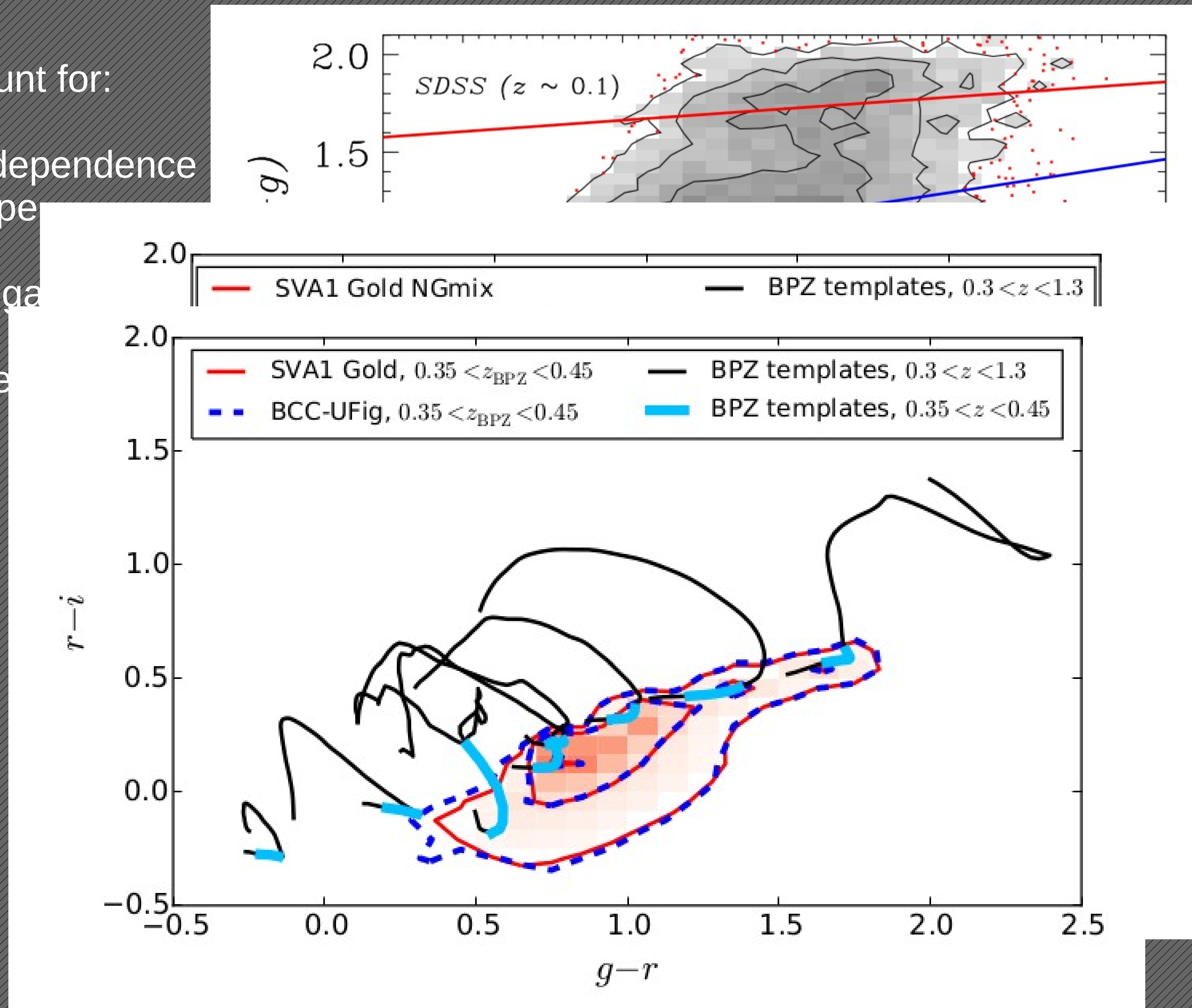
Despite including a data-driven prior, the model is still wrong!

Templates do not account for:

→ colour – luminosity dependence
  for given galaxy type

→ redshift evolution of ga

→ incomplete coverage
  galaxy types



Bonnett, Troxel, Hartley et al (2015)

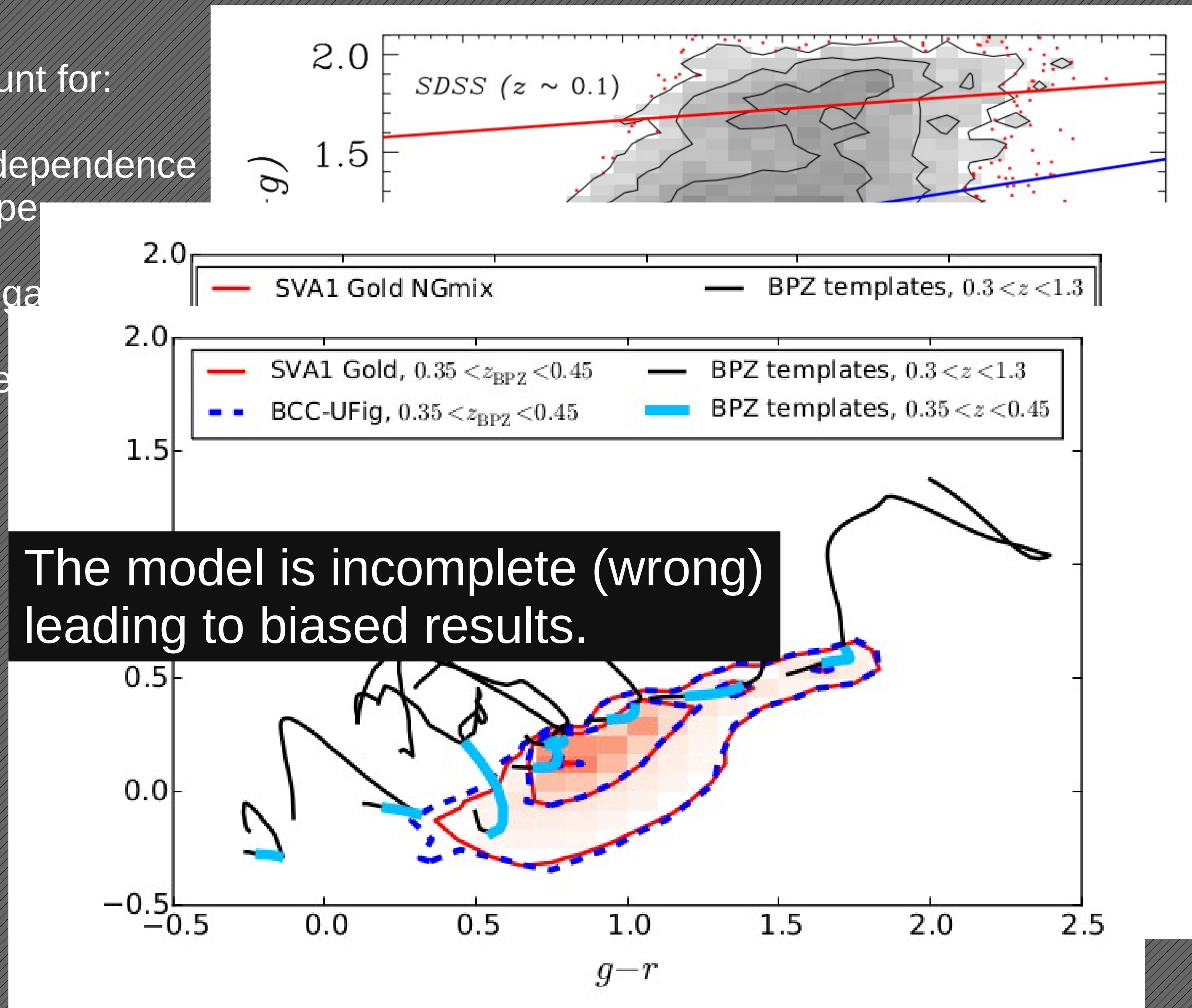
Despite including a data-driven prior, the model is still wrong!

Templates do not account for:

→ colour – luminosity dependence
for given galaxy type

→ redshift evolution of ga

→ incomplete coverage
galaxy types



The model is incomplete (wrong) leading to biased results.

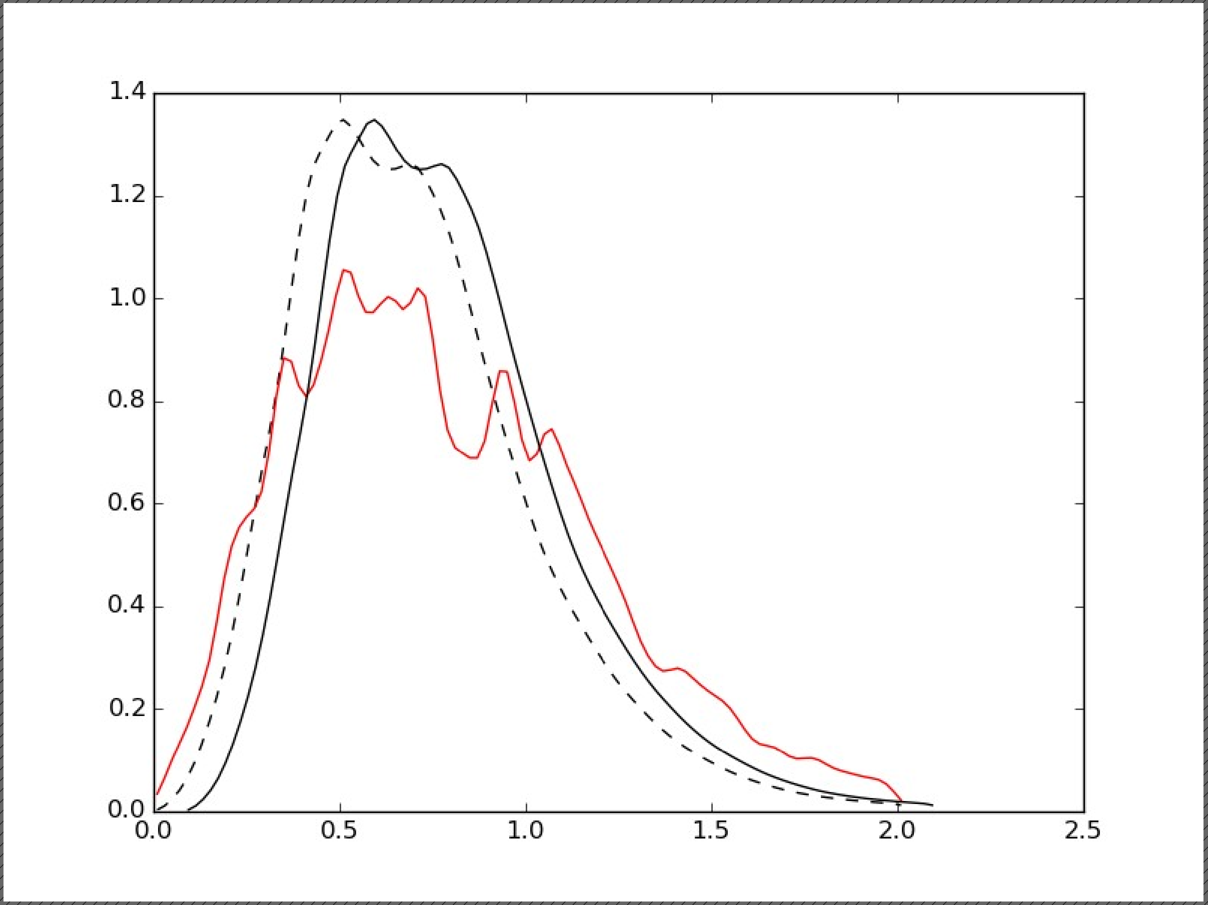
Bonnett, Troxel, Hartley et al (2015)

Solution: Model the bias introduced via simulation

BPZ was run on N-body + image simulation data (Wechsler+, Refregier+, Leistedt).
A weak-lensing-like sample was culled* and we explored methods to map the BPZ PDF to the truth.

Calibration implies a simple offset: $\Delta z = 0.08$

n(z)

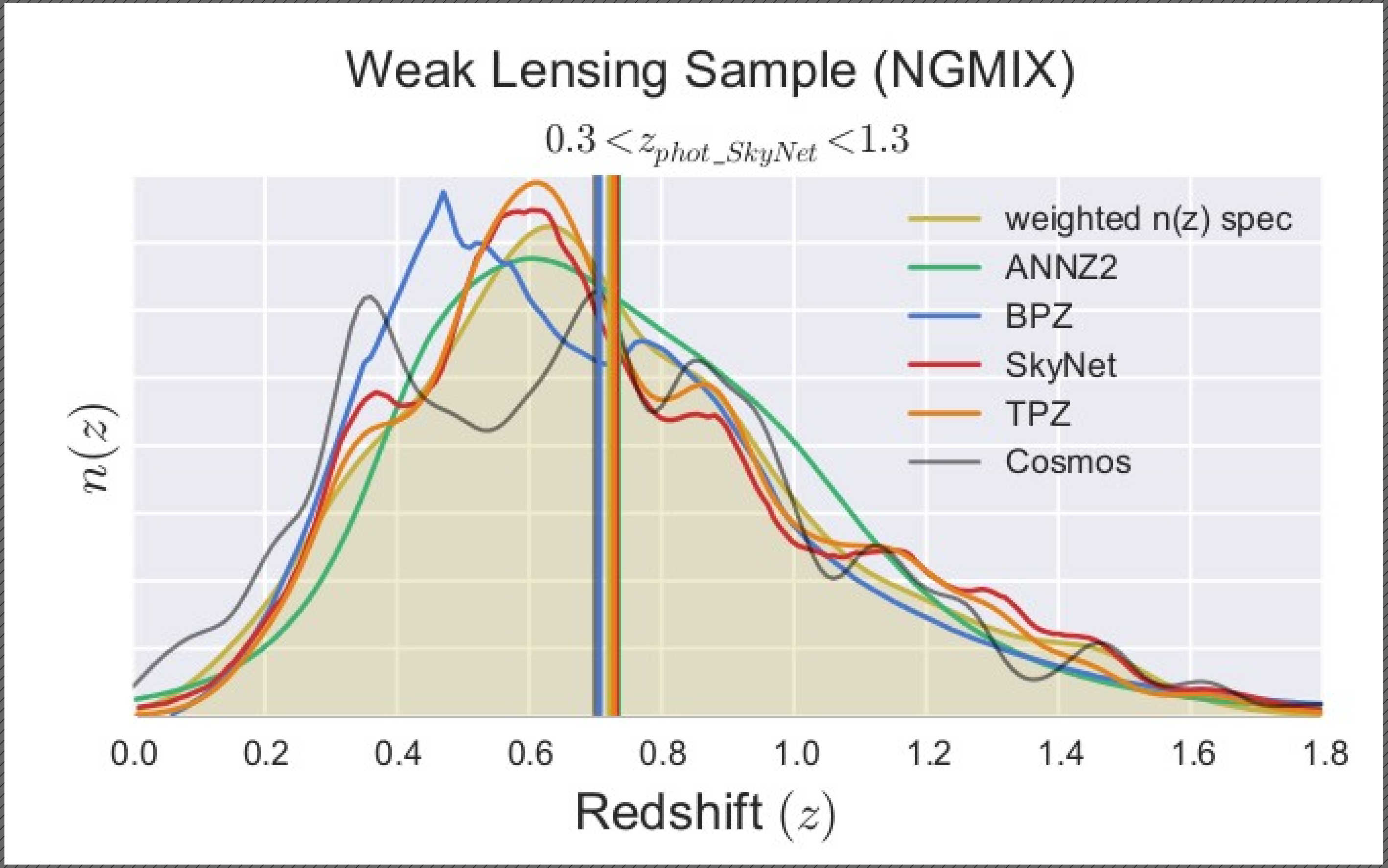


redshift

Red: simulation truth
Black dashed: BPZ n(z)
Black solid: bias corrected BPZ

*non-trivial to do, but we were able to get close enough.

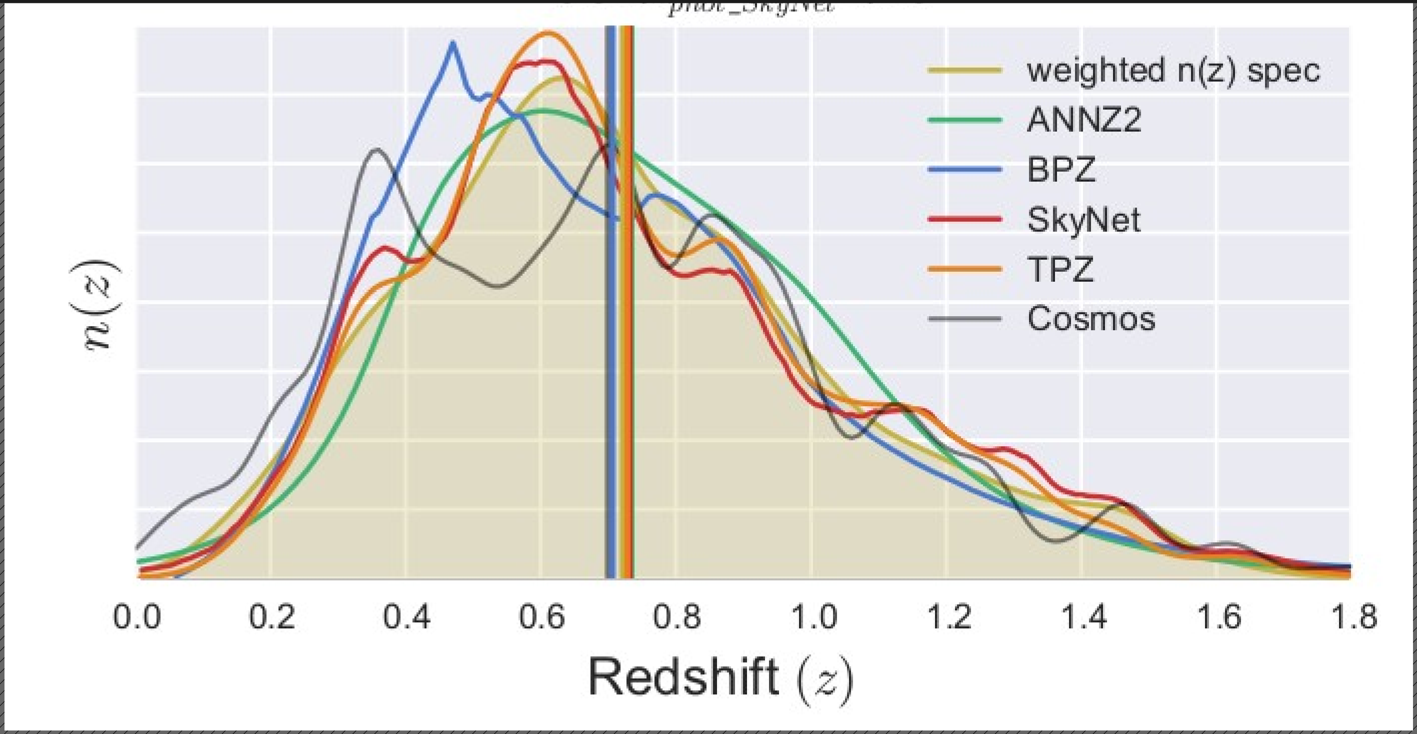
	DES SV -WL sample	Validation sample
Spectra	0.72 (weighted)	0.64
ANNZ2	0.73	0.65
SKYNET	0.73	0.65
TPZ	0.73	0.64
BPZ	0.71	0.64
Matched COSMOS	0.70	-



Bonnett, Troxel, Hartley et al (2015)

	DES SV -WL sample	Validation sample
Spectra	0.72 (weighted)	0.64
ANNZ2	0.73	0.65
SKYNET	0.73	0.65
TPZ	0.73	0.64
BPZ	0.71	0.64
Matched COSMOS	0.70	-

After correcting for model bias, the template code, ML methods, spectroscopic and COSMOS photo-z mean redshifts agree to within σ=0.02.
(0.05 in tomographic bins)



Bonnett, Troxel, Hartley et al (2015)

Next steps:

Improve the templates and prior to capture galaxy evolution

→ empirically?
→ synthetically?

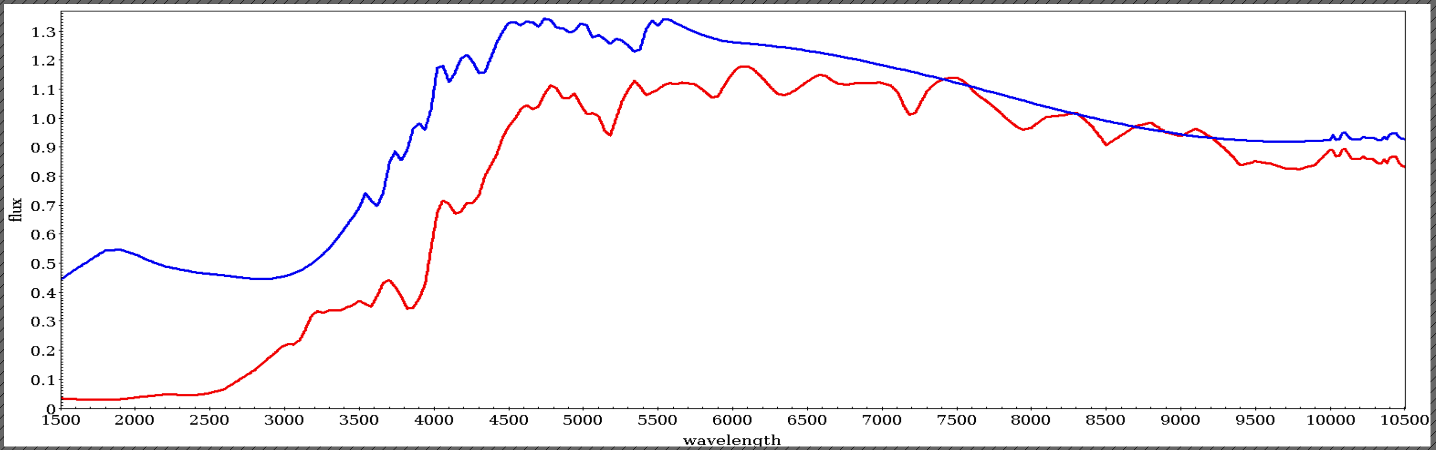
(work in progress)

Empirical modification using PRIMUS galaxies

- fit PRIMUS galaxies at their z_spec to the CWW templates
- look at the photometric offsets in different redshift ranges for different templates as a function of rest-frame wavelength.

Blue: Sbc
Red: Elliptical

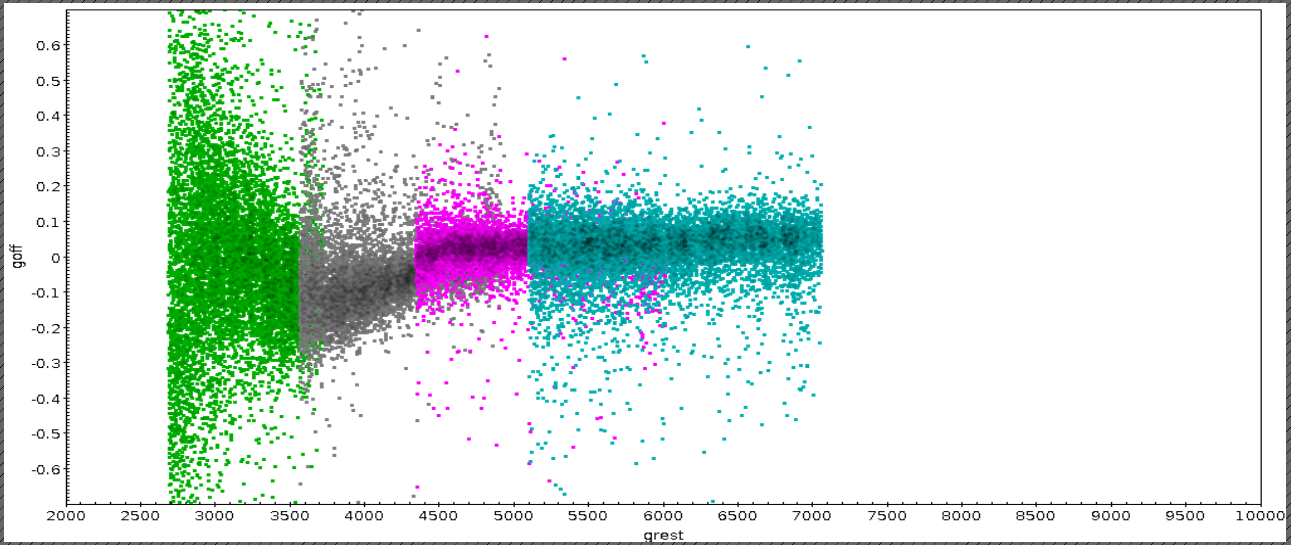
flux



rest wavelength

Elliptical galaxy at
$0.3 < z < 0.8$

mag difference



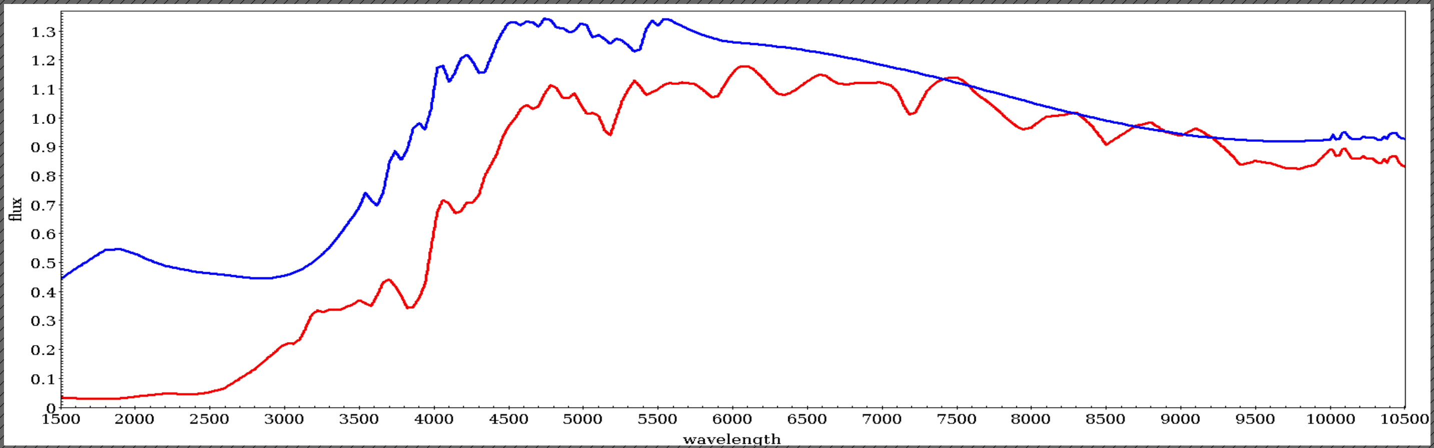
rest wavelength

Empirical modification using PRIMUS galaxies

- fit PRIMUS galaxies at their z_spec to the CWW templates
- look at the photometric offsets in different redshift ranges for different templates as a function of rest-frame wavelength.

Blue: Sbc
Red: Elliptical

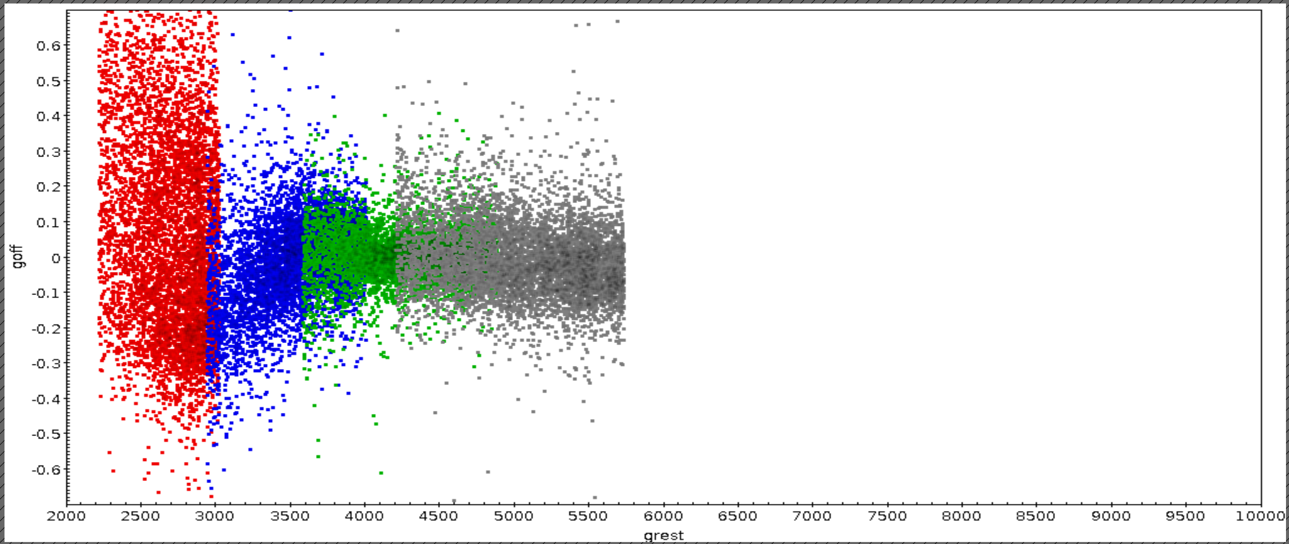
flux



rest wavelength

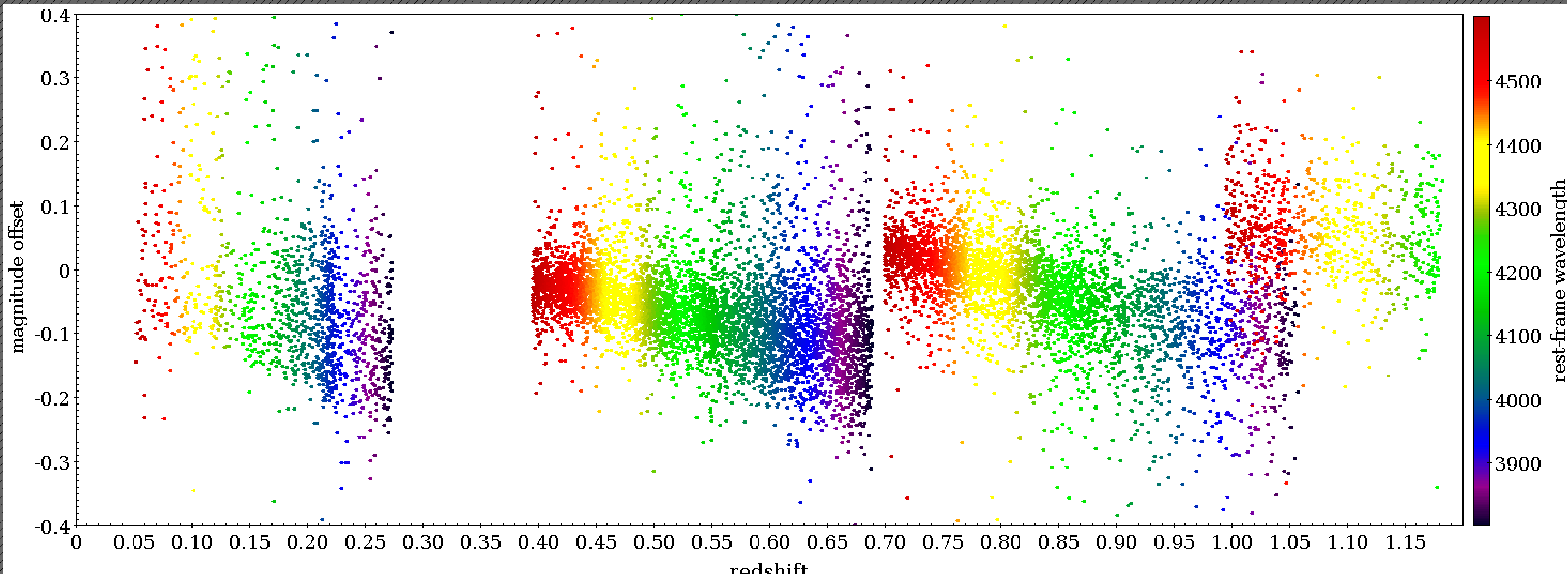
Sbc at
0.6 < z < 1.18

mag difference



rest wavelength

In practice, such wide redshift intervals are not ideal



→ Evolution within the redshift intervals needed to ensure overlap of bands

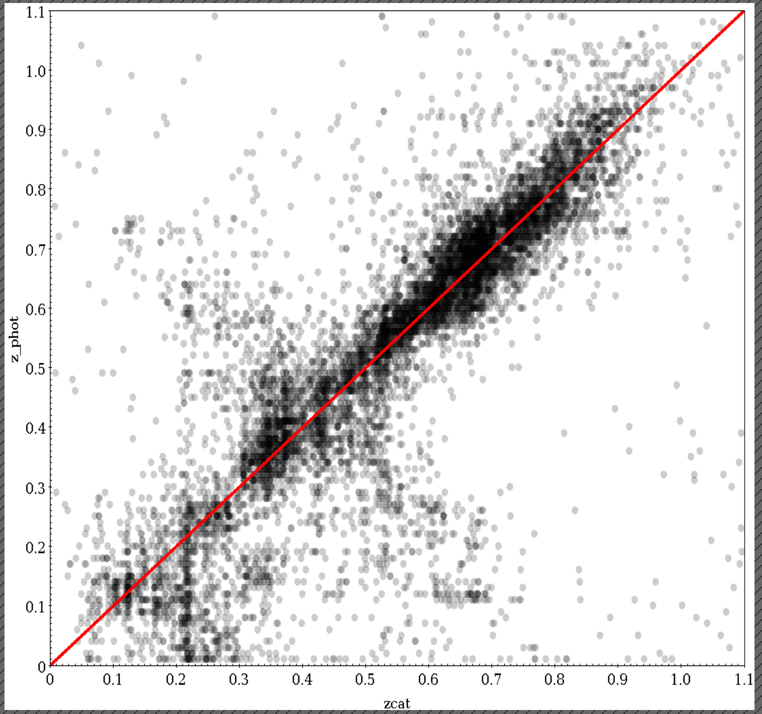
→ Instead, fit a 2D function to describe, offset(z,λ), for each basic template

→ Scatter tell us the level of systematic uncert. we need to apply to the photometry.

Performance:

20 < i < 21

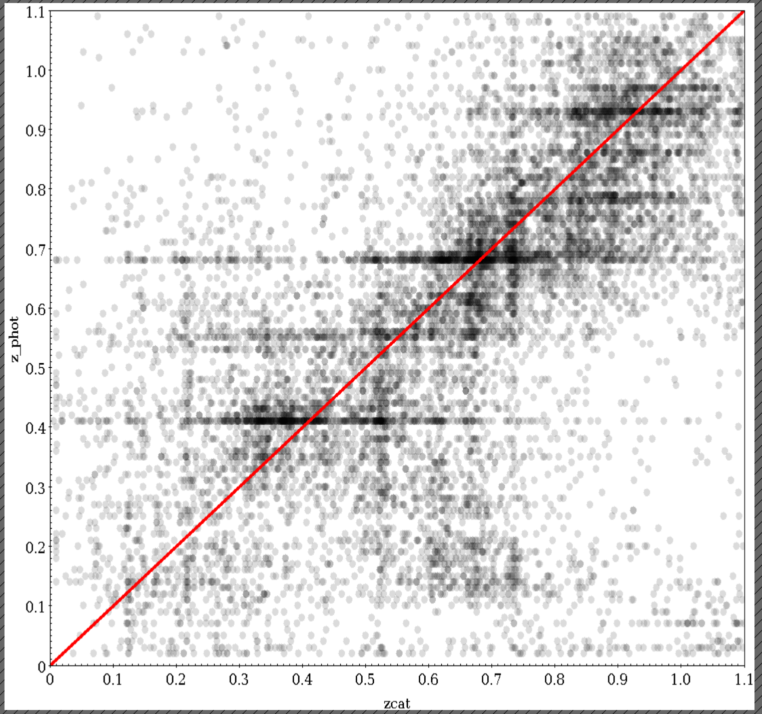
z-phot (ZEBRA)



z-spec

i > 22

z-phot (ZEBRA)



z-spec

→ DES Y1 data (i_{lim} ~ 23.)

→ internally calibrated prior

→ overall bias is improved

→ z~0.4 feature absent

→ degenerate solutions somewhat under control

→ external prior should provide further Performance boost

→ overall, somewhat encouraging

Towards realistic synthetic templates

(see also, Tanaka 2015 and Chevallard & Charlot 2016)

The 'classic' implementation of synthetic templates in photo-z:

→ exponentially declining SF histories

→ single metallicity per template

→ range of allowed ages, limited by age of the Universe

→ otherwise, all templates allowed at all redshifts

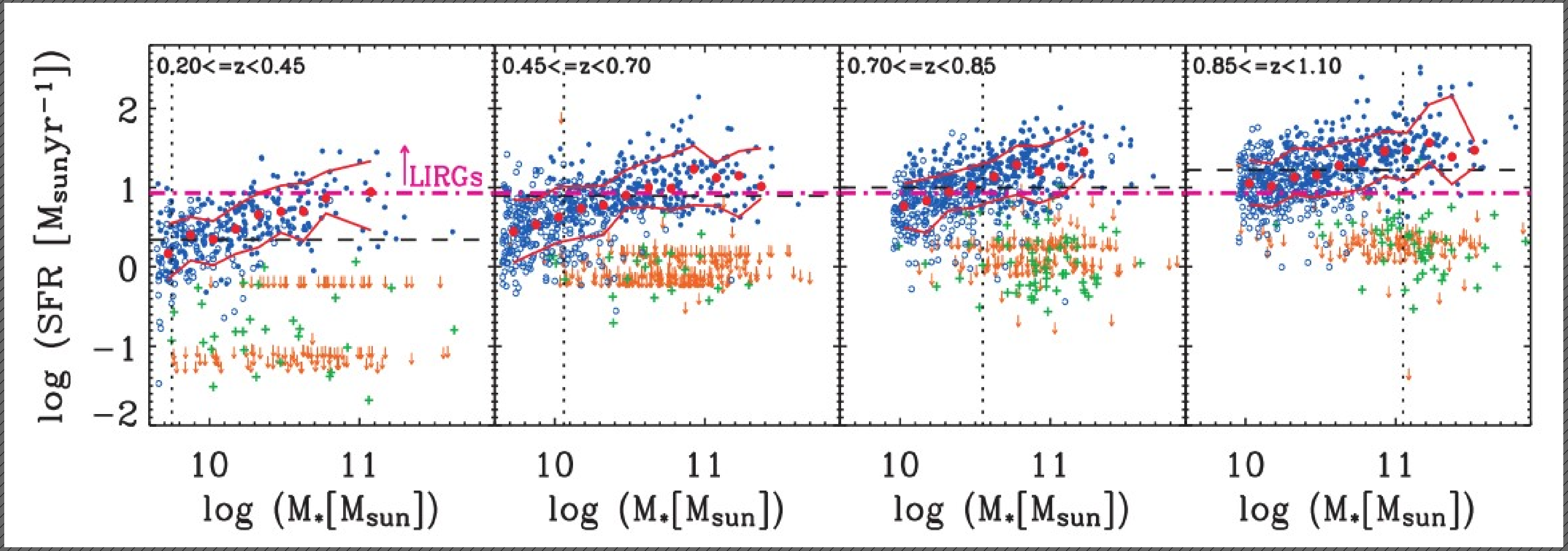
→ inter-galactic extinction applied as usual

Similar to empirical model methods, typically good enough for acceptable performance in data-rich regime.

But templates do not reflect current models of galaxy evolution.

21st Century galaxy evolution

By using the same templates from z=0 → z=n we are assuming that galaxies do not evolve.
But of course they do:



Noeske et al. 2007

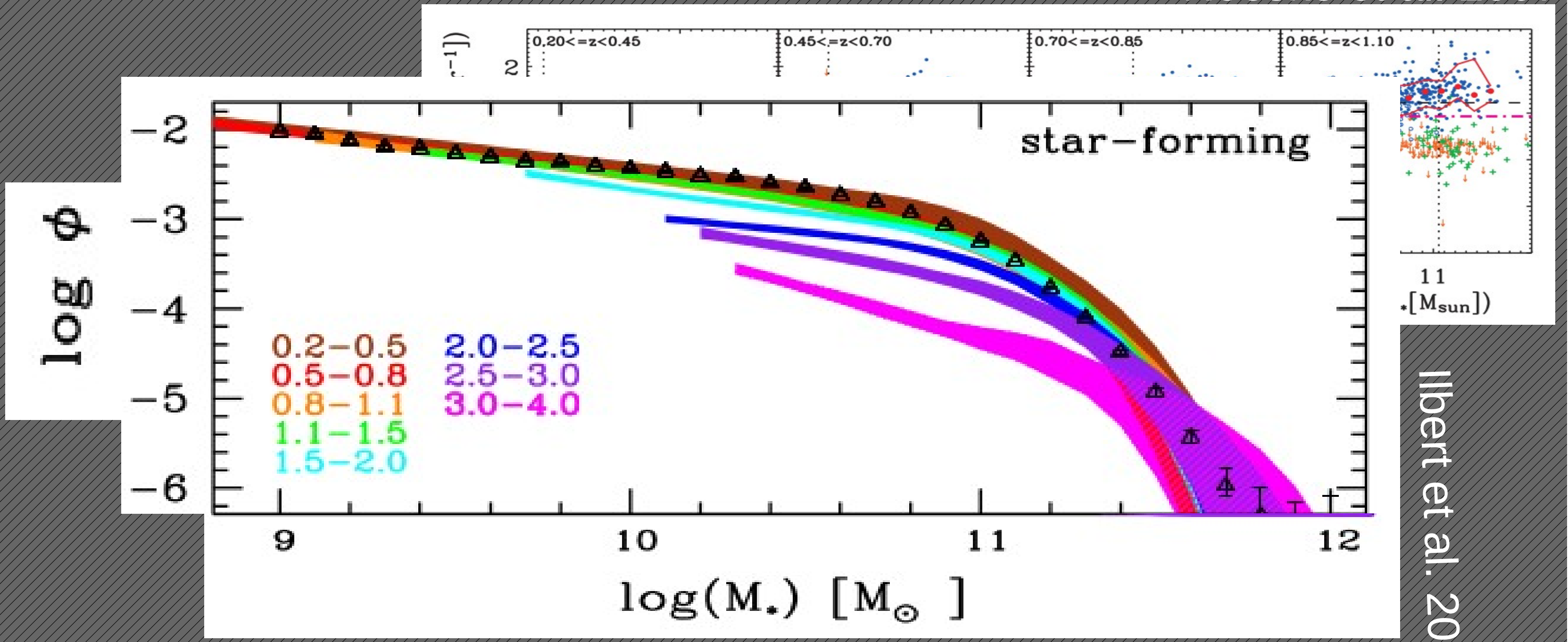
Higher typical specific star-formation rates at high-z.
Scatter in SFR – mass relation is only ~0.3 dex.

→ 'Main sequence' of SFing galaxies

→ Now characterised to z~4 (with claims for a relation at even higher-z).

21st Century galaxy evolution

Noeske et al. 2007



Ilbert et al. 2013

Stellar mass function of star-forming and quenched galaxies now traced to high redshift (z~4)*.

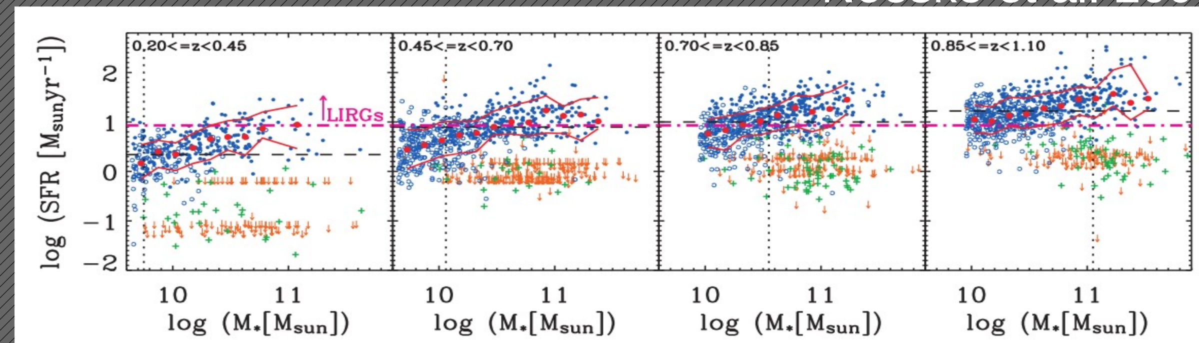
Shape of SFing mass function does not change with time – only in normalisation.

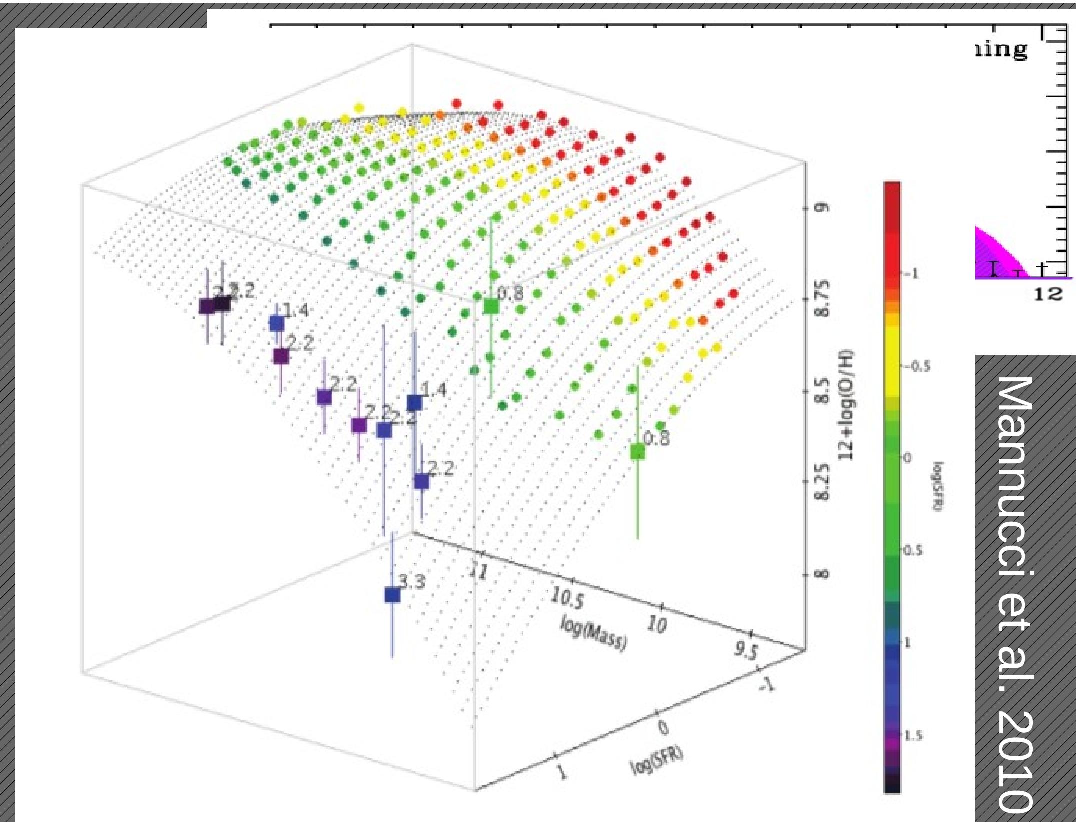
→ Together with M.S., lead Peng et al. (2010) to define the form of 'mass quenching' (switch off of star-formation due to processes correlated with stellar mass).

* subject to sample variance

21st Century galaxy evolution

Noeske et al. 2007





Ilbert et al. 2013

Mannucci et al. 2010

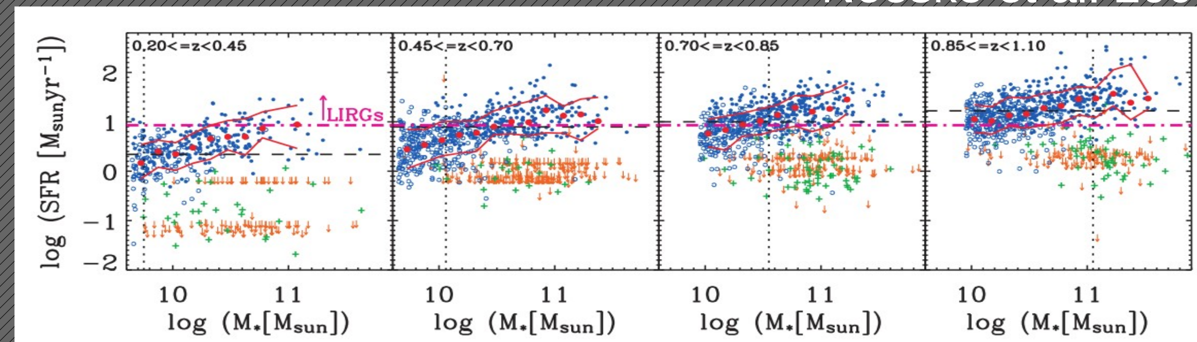
Metallicity of star-forming gas is a function of galaxy stellar mass and SFR which appears to be independent of redshift.

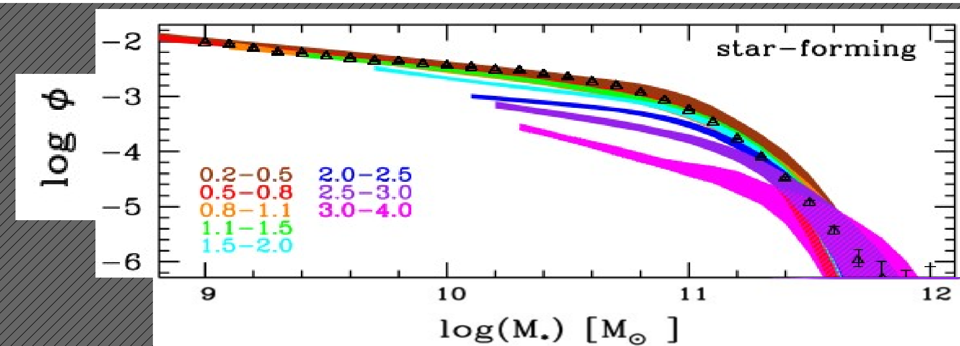
→ 'Fundamental metallicity relation'

→ Avoid unphysical templates (e.g. high metallicity in high-sSFR galaxies)

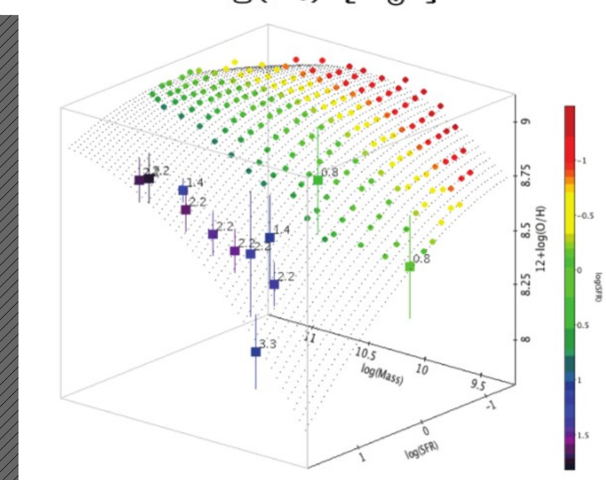
21st Century galaxy evolution

Noeske et al. 2007





Ilbert et al. 2013



Mannucci et al. 2010

→ Include these aspects in a self-consistent, redshift-sensitive set of templates.

→ Follow galaxies as they evolve on the Main Sequence.

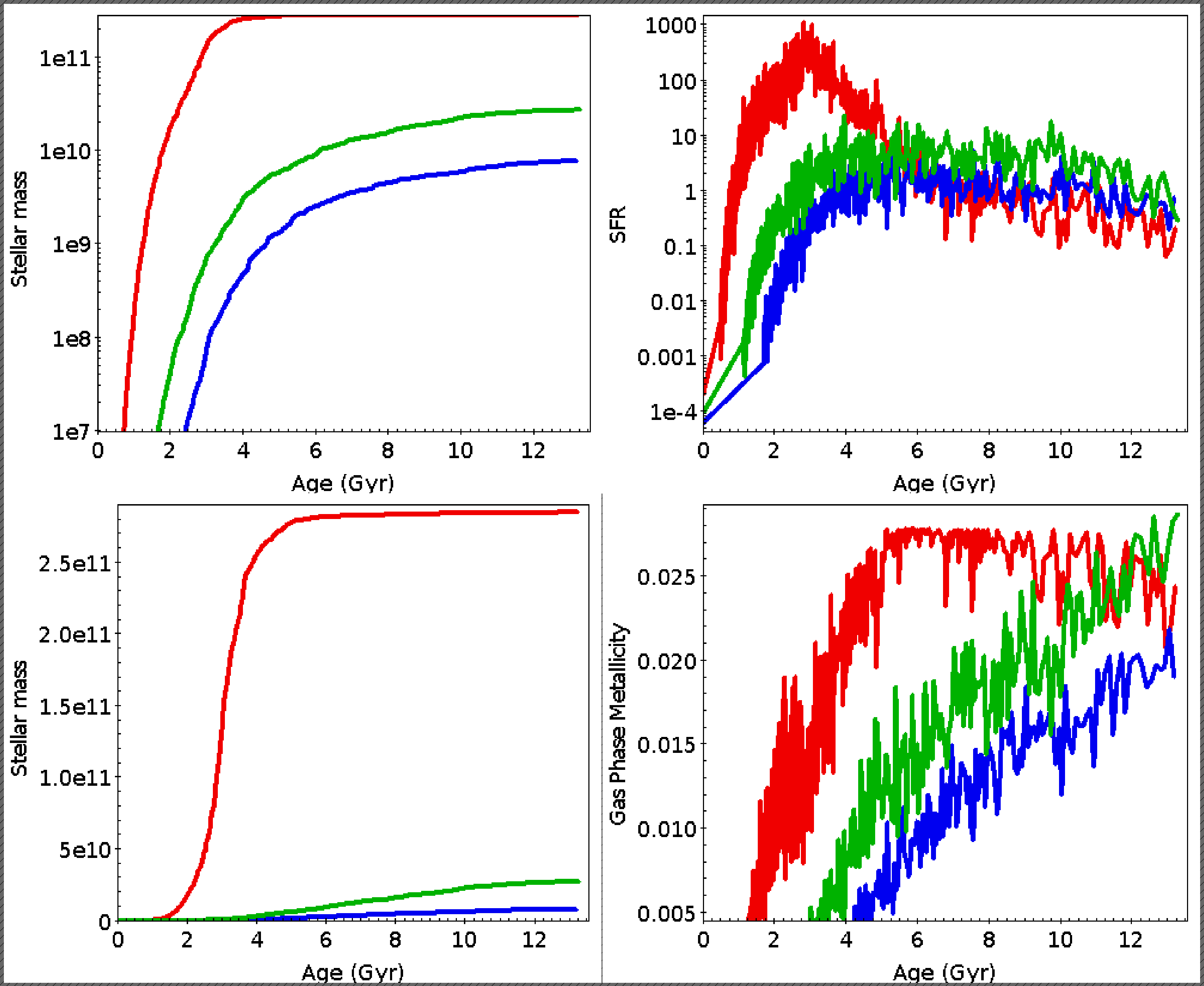
→ Quench according to Peng et al. (2010).

→ Form stars with metallicty according to FMR.

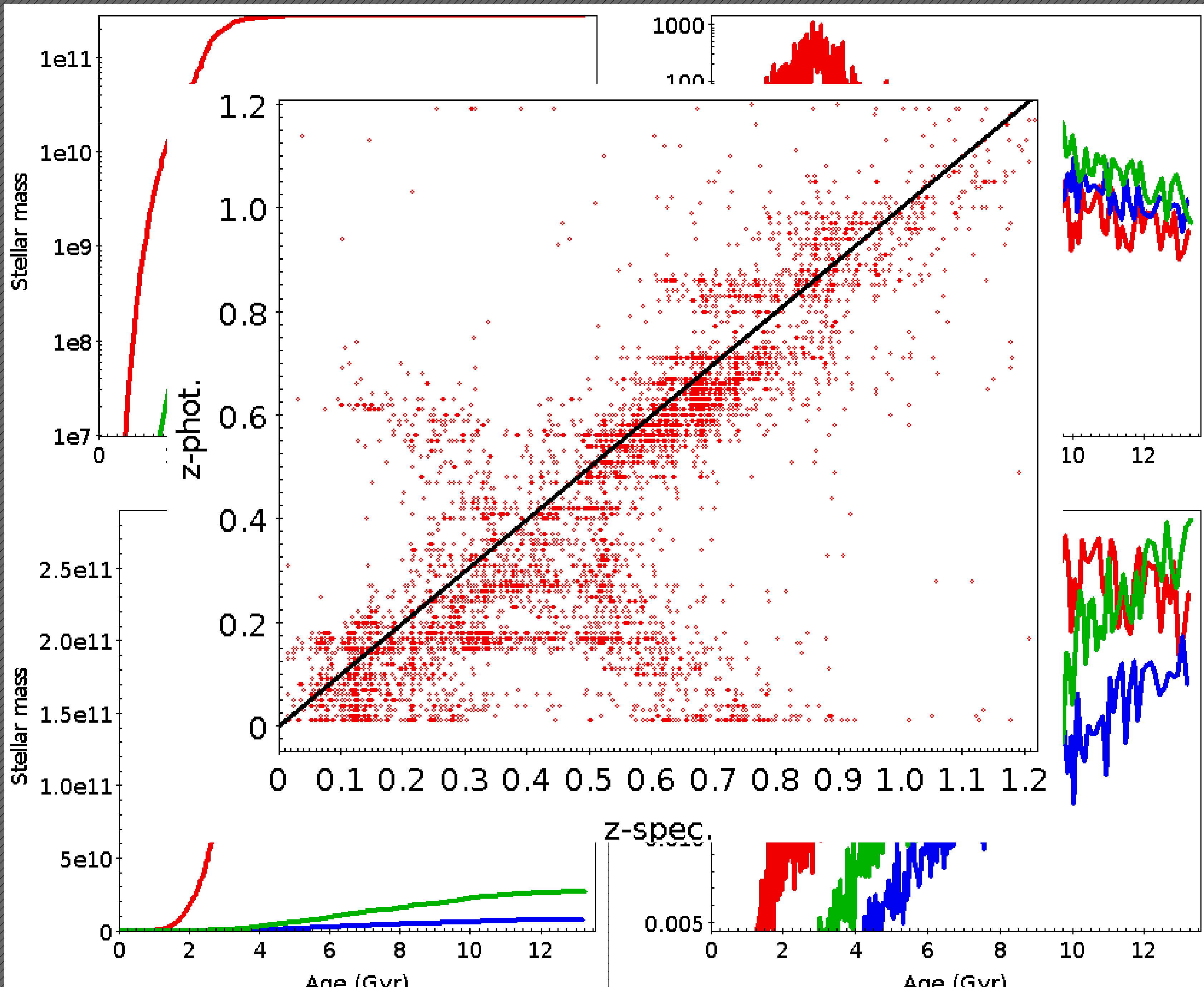
→ Only allow templates to be fit at the appropriate redshift.

(many other aspects not yet included)

Synthetic galaxies (FSPS) tracked along the star-forming sequence and quenched according to current ideas in galaxy formation.



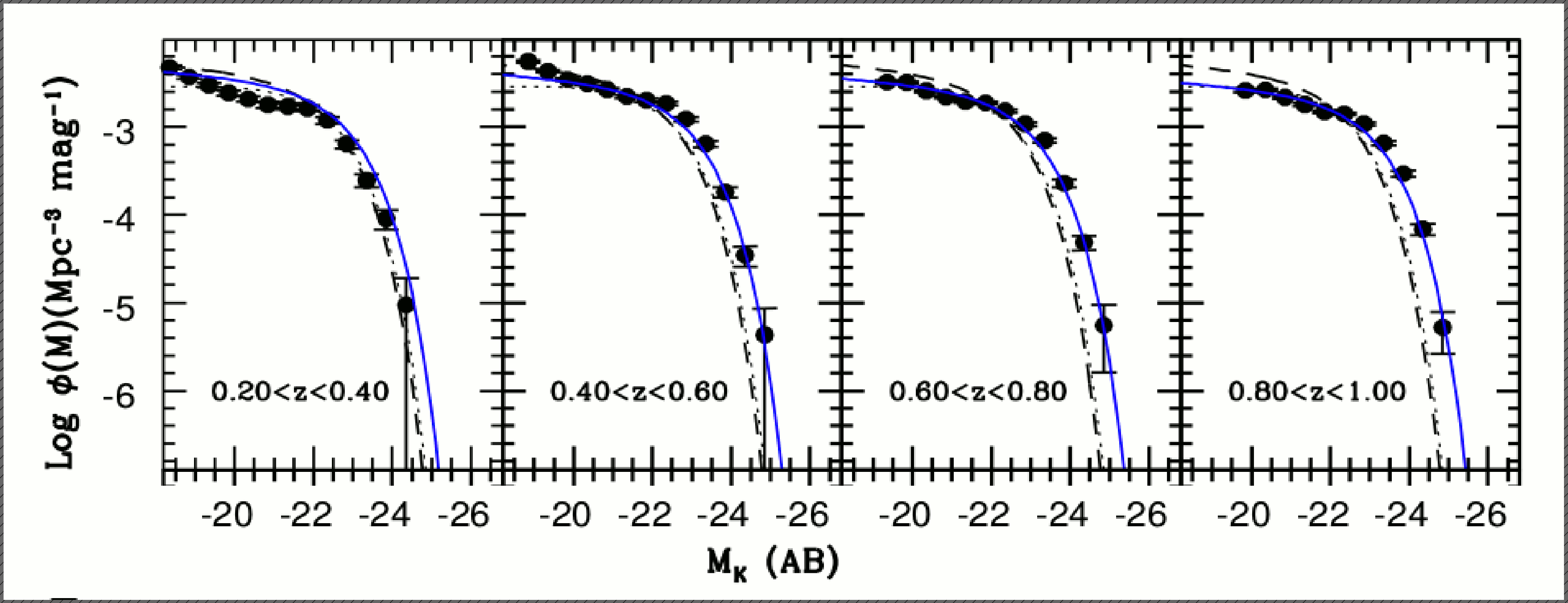
The results are not inspiring, but no prior has been used yet...



Galaxy-evolution-motivated priors

Constructing a prior...

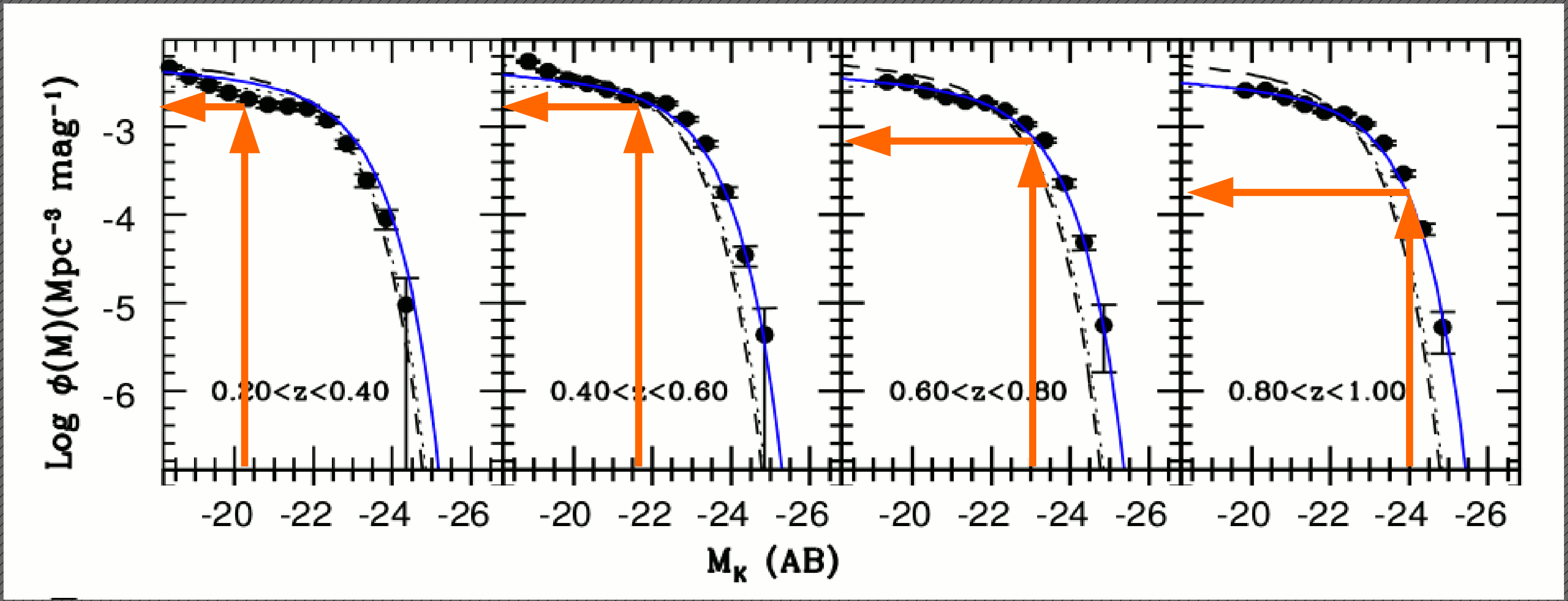
Fundamentally based on the galaxy luminosity (or stellar mass) function:



Cirasuolo et al. (2010)

Constructing a prior...

Fundamentally based on the galaxy luminosity (or stellar mass) function:



Cirasuolo et al. (2010)

- A given apparent magnitude implies a different luminosity at different redshift, and therefore a different space density.

- Including the volume → n(z), with ~exponential fall-off at high-z.

- Can construct luminosity functions for different galaxy types.

- Valid for both empirical and synthetic templates.

- Further quantities (e.g. morphology) can be included easily.

In practice....

→ Spectroscopic samples
  - often cover specific redshift intervals
  - are incomplete
  - often are over small areas of sky (→ sample variance)

→ High-quality photo-z samples
  - cover small areas of sky
  - may contain biases

→ Numerical models (e.g. semi-analytics)
  - do not yet capture all galaxy types sufficiently well

→ Phenomenological models
  - may introduce biases due to over-simplification

These problems are not insurmountable, however.

The magic in RedMaGiC

(Eduardo Rozo & Eli Rykoff)

→ Focus on the easiest galaxies

→ Use a single empirical 'template' for the lumin. and redshift-dependent red sequence.

→ Template from RedMapper cluster finder.

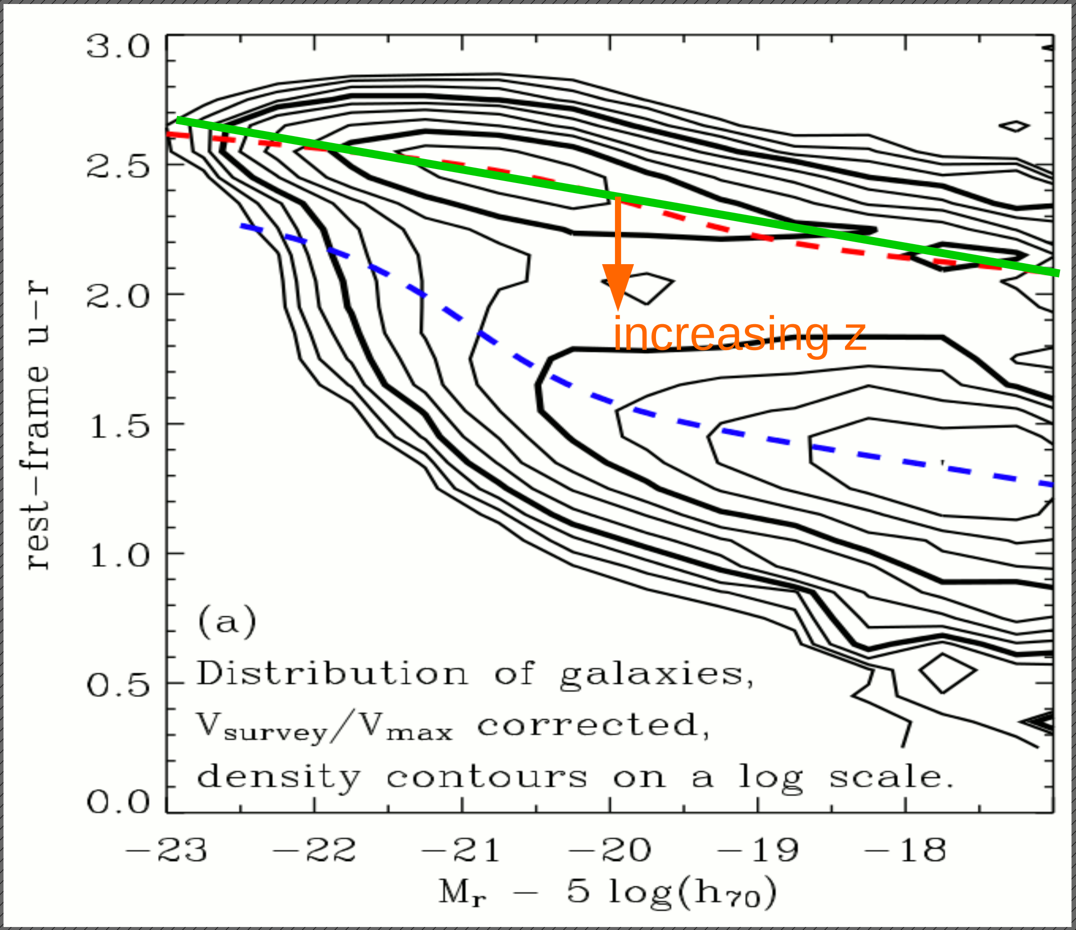
→ Prior on the form of the lumin. fn.
- a Schechter function.

→ Redshift range limited by cluster finder (currently z<0.8).

→ Relies on cluster R.S. galaxies being representative of field R.S.

→ Spectra used to refine solution.

→ Complete solution for one particular galaxy type.

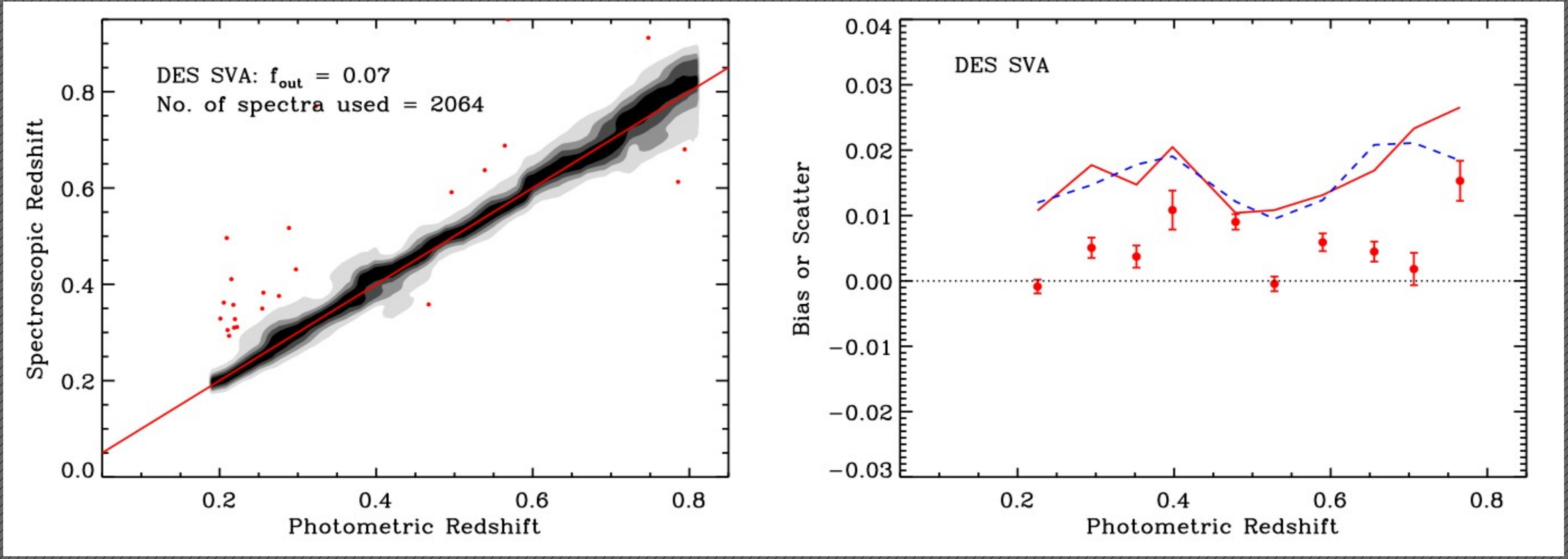


Baldry et al. (2004)

RedMaGiC performance:

- Photo-z bias at the level required, even in narrow redshift intervals.

- Caveat:
    - R.S. galaxies are the easiest to get precise redshifts for.
    - need to extend this success to general population for WL.



Rozo et al. (2015)

points: bias
red: actual scatter
blue: predicted scatter

Take home points

The requirements on photo-z accuracy and precision in cosmology experiments present a considerable challenge.

Modelling type approaches to photo-z will *have* to work due to the difficulty in compiling complete spectroscopic samples.

Traditional template sets and prior calibrations, together with a model bias correction were just enough for DES SV.

Improved templates, either empirical or synthetic, together with appropriate priors are required.

Some distance to go, but there are some encouraging signs.

One complete solution (RedMaGiC) for the easiest galaxy subset demonstrates that it can work.

The problem of an unknown error model

Uncertainties in the redshift solution from:

Catalogue flux errors

→ often poorly defined, especially for model magnitudes

(what is the error on the error estimation?)

'Systematic' flux errors

→ ~2% flux error arbitrarily inserted to cover calibrations

Template errors

→ rarely used, and need tuning

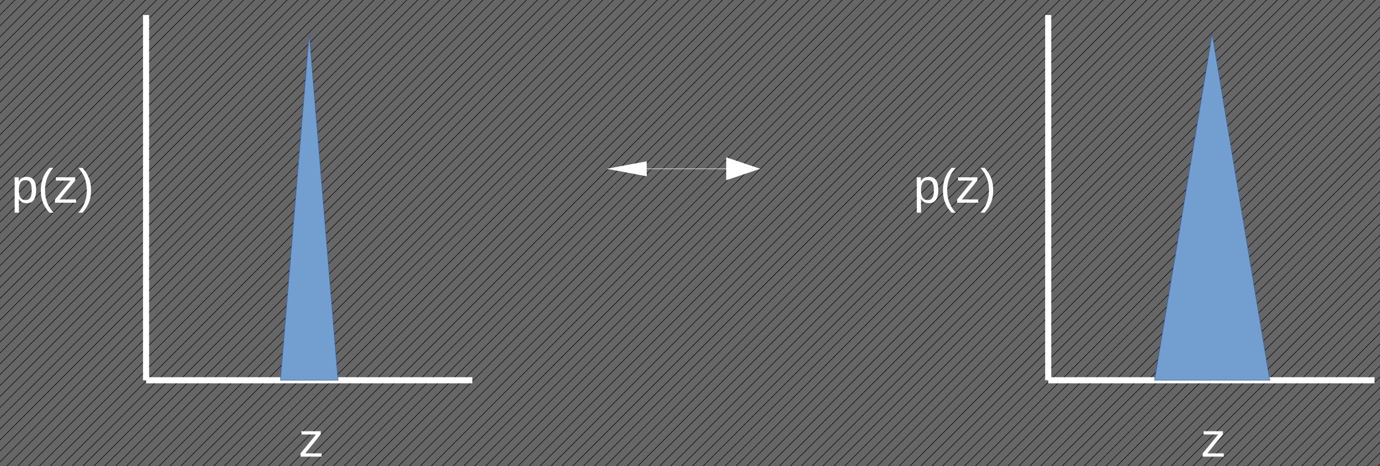
→ redshift dependent error at λ_{eff}

Uncertainties in the redshift solution from:

Catalogue flux errors
→ often poorly defined, especially for model magnitudes

'Systematic' flux errors
→ ~2% flux error arbitrarily inserted to cover calibrations

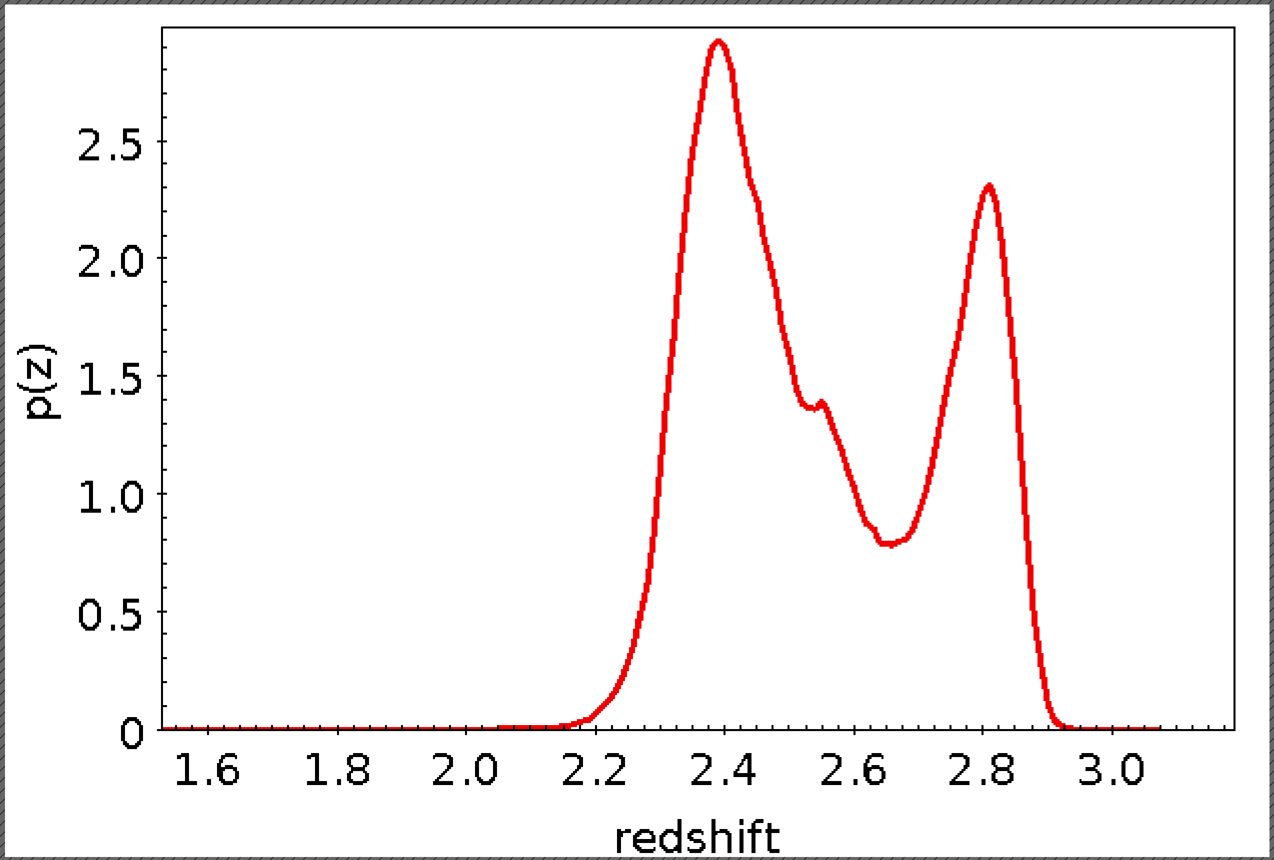
Template errors
→ rarely used, and need tuning
→ redshift dependent error at λ_{eff}



Width of the PDF matters for cosmology!
(much more than for galaxy evolution)

In principle, this error bias can also be calibrated by sims in same way as the model bias.

Alternative is to use information in spectroscopic samples to re-map PDFs. (Bordoloi et al. 2010).

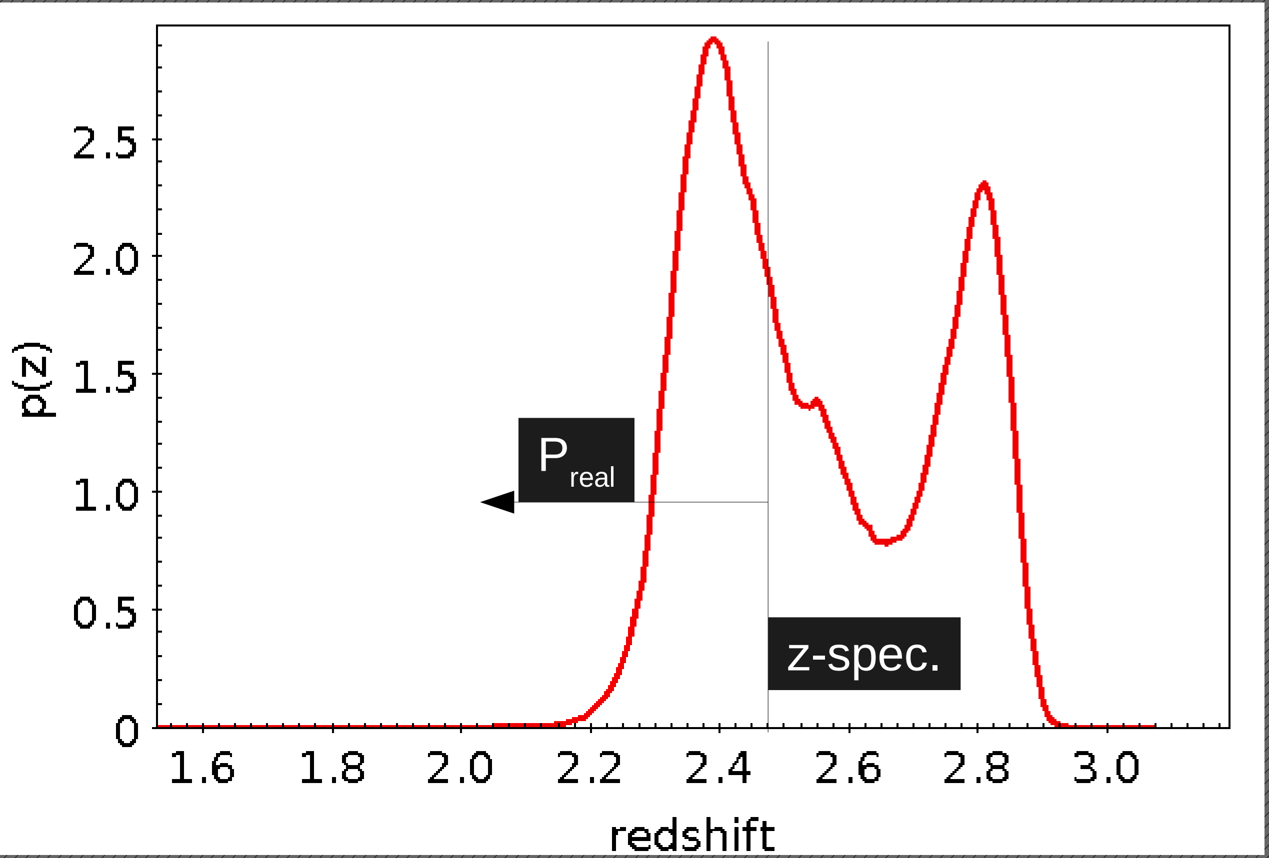


In principle, this error bias can also be calibrated by sims in same way as the model bias.

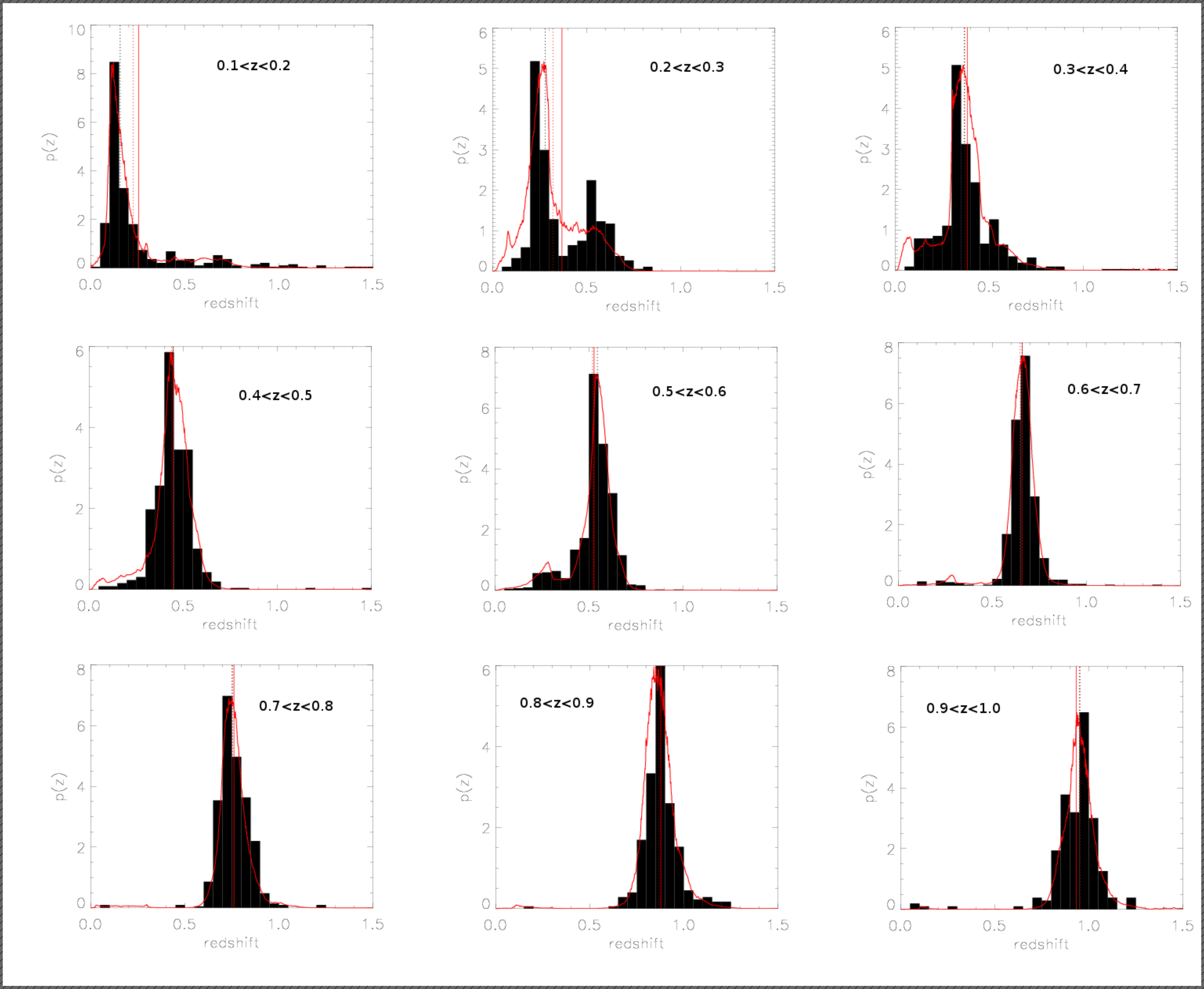
Alternative is to use information in spectroscopic samples to re-map PDFs. (Bordoloi et al. 2010).

Spec-z should be a random draw from the PDF.
i.e. for a population, the distribution of cumulative probabilities up to the spectroscopic redshift should be flat between 0 and 1.

Where they are not, each PDF can be remapped so that the population becomes flat.



Remapped PDFs (includes removal of 10% worst PDFs before applying algorithm).



Solution?

Remapping PDFs can boost performance considerably, but alone is not enough.

However, it is perhaps better to use the spectra (and Bordoloi method) for *validating* redshifts, and put more effort into understanding the error model.

Anatomy of a template-fitting code