Impact of SED Uncertainties on Cosmological Constraints

Sam Schmidt
Pittsburgh DESC Meeting
Dec 4, 2013

Galaxy SEDs

- Template/training set (+ prior distribution) is the heart of any photo-z algorithm, need set to span all SEDs that we will actually encounter

- What happens if we have the SEDs slightly wrong?
  - How well do we currently understand the SED set?

- How do we parameterize our lack of knowledge realistically in simulations?

Possible Pitfalls: Non-representative SEDs

- Discretized template set
- Non-representative/missing data
- slightly incorrect model data
- AGN contamination/galaxy evolution
- Effect of such errors on cosmological predictions

Discrete Templates: Template Mismatch

- Galaxies actually drawn from a fairly continuous distribution, we often represent with a finite set of templates (PCA style formulations help)

- Map to incorrect template maps to (hopefully) small error in redshift

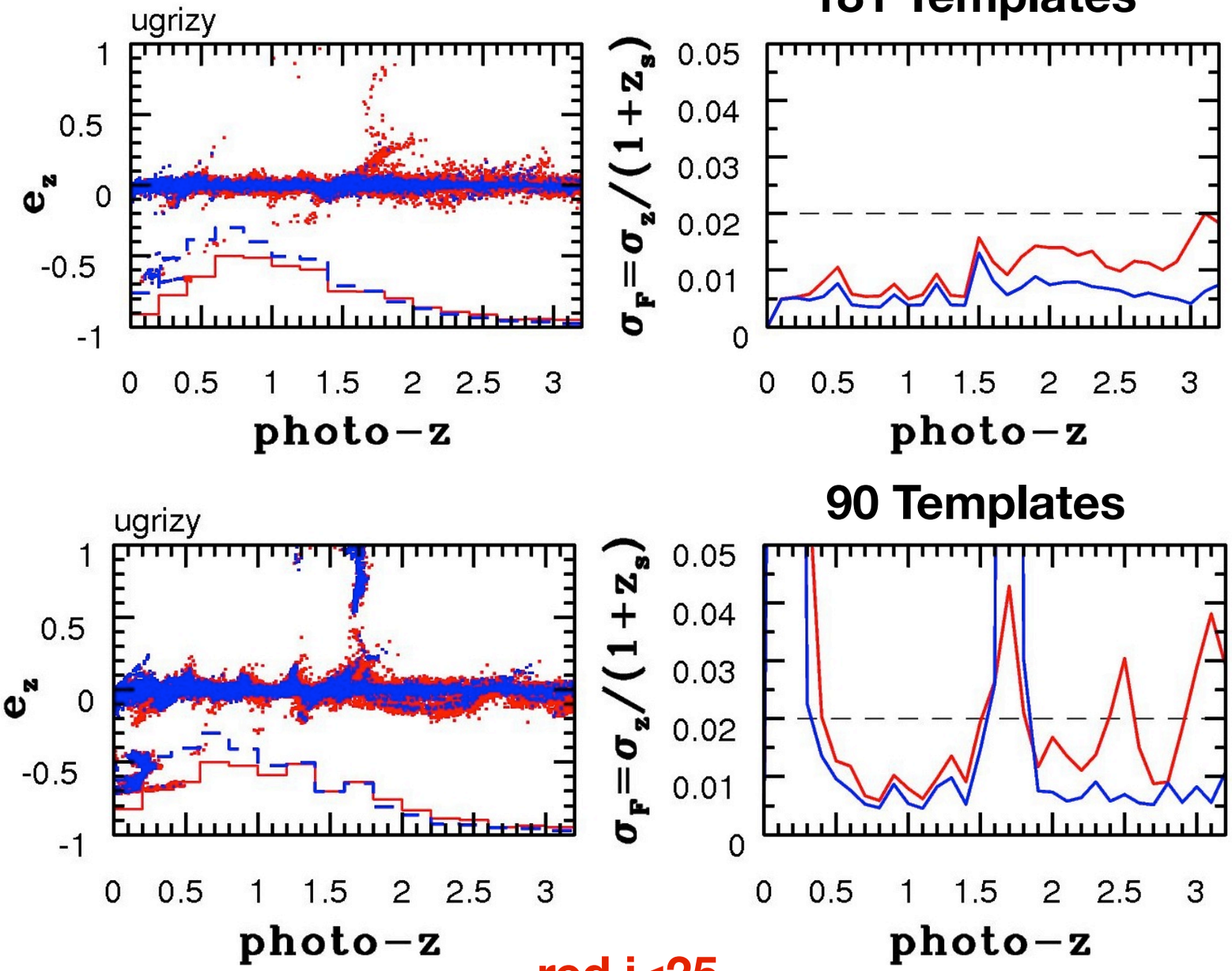
- If templates fairly representative/spanning, this usually manifests as increase in uncertainty, small change in bias, e.g. SciBook sims

    - Use 181 templates to generate data, use 90/181 in the photo-z determination

Template Mismatch Example

- 181 templates vs 90 templates
- σ_z increases by factor of ~2, much more at low-z and z~1.4
- Lenient quality cut, could remove gals flagged as bad, but would decrease sample size

$$e_z = (z_p - z_s)/(1+z_s)$$



Non-Representative Data

- What if we do not have redshifts/SEDs for a particular type of galaxy?

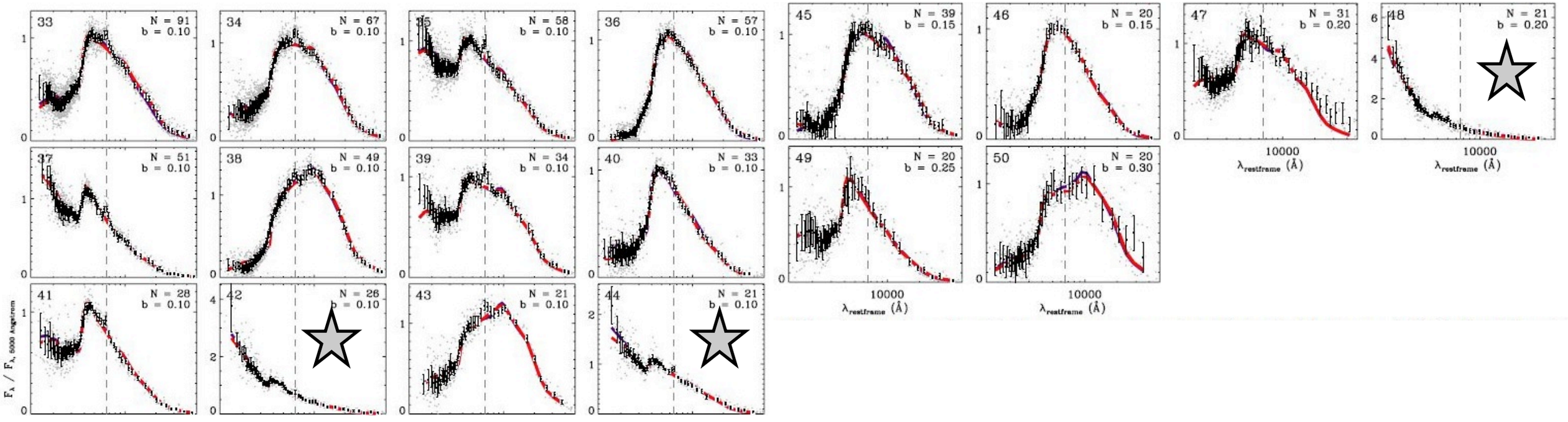
- E.g. an alarming percentage of DEEP2 galaxies do not yield redshifts even with >1hr integration on Keck

- Can try to fill in gaps with synthetic spectra, but SSP/BC models known to be uncertain to at least a few % in continuum, emission lines further complicate

- Difficult to parameterize our missing templates in simulations

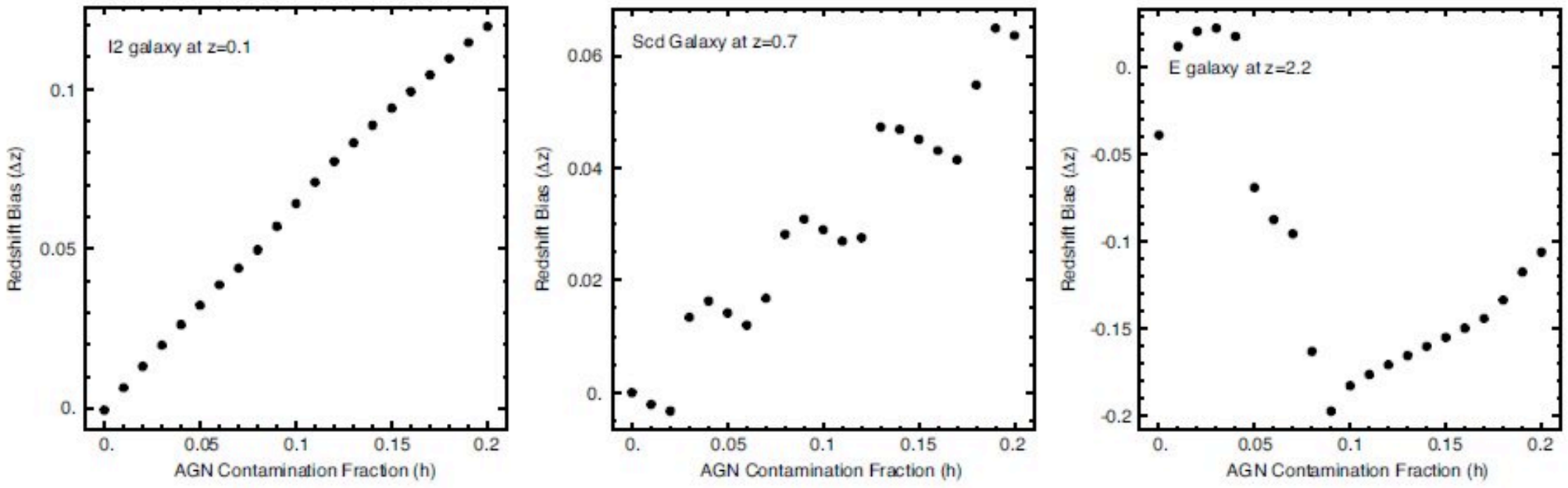
Many-band Photo-z's

- Kriek et al. (2011) performed K-selected survey with NIR medium bands and supplemental data covering UV to NIR. (see also COSMOS 30-band)
    - Group galaxies by how similar they are, find 32 groups that contain 83% of the sample (Good news: look similar to low-z SEDs)
    - Remaining 17% has 3 templates not represented by the main 32 (very unobscured blue SF galaxies), may be smattering of disparate SED types.



SED Contaminants

- Intrinsic: AGN contamination
- MacDonald & Bernstein 2010 look at AGN contamination, add 0-20% light from AGN template
- Find 1% AGN can cause bias of $\Delta z \sim 0.005$



Can see effect of discretized templates as well

Madau Reddenning

- Currently, all photo-z codes and sims (that I know of offhand) treat Madau reddenning with models of the mean observed.

- Actual reddenning will be stochastic, depends on the amount of IGM along the line of sight, should be dealt with both in models (and possibly templates/ clustering)

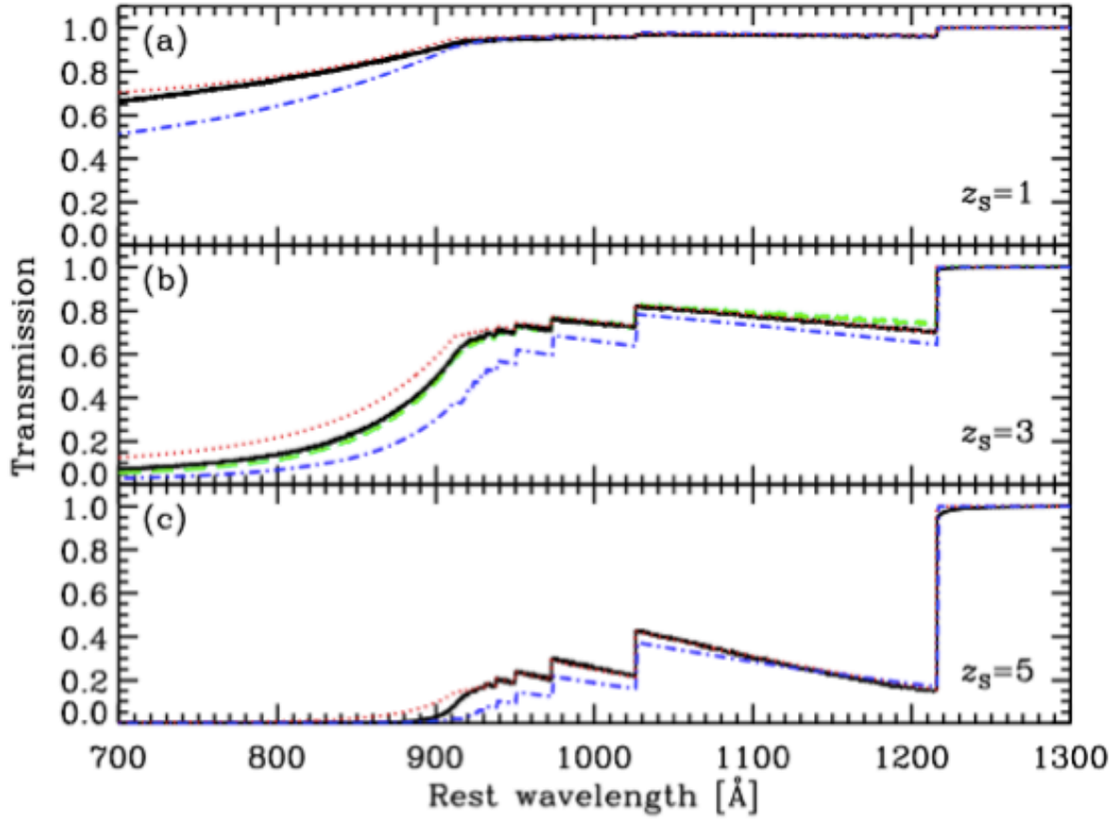


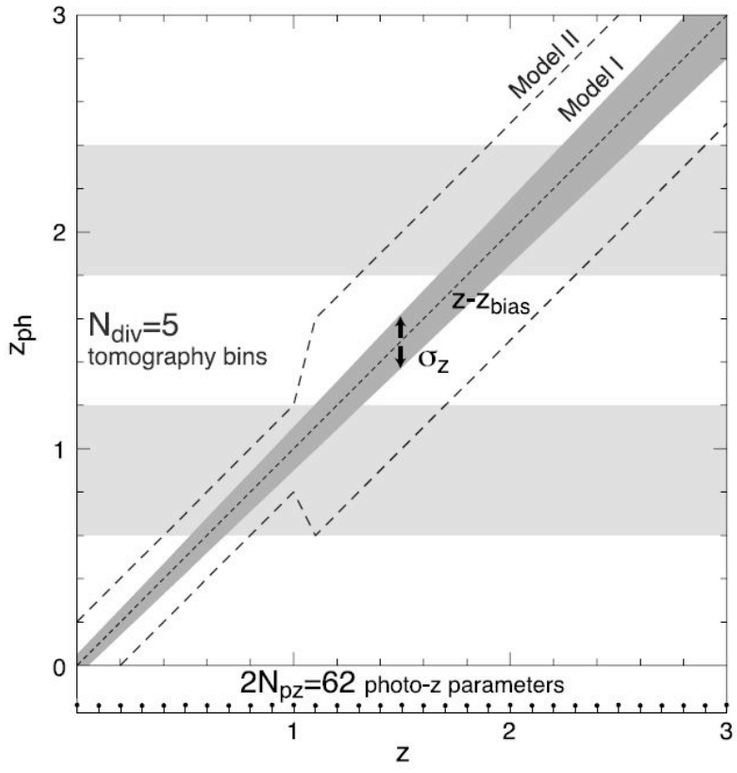
Figure 6. Average IGM transmissions. The horizontal axis is the wavelength in the source rest-frame. The source redshifts are noted in the panels. The solid lines are our Monte Carlo results; 10,000 lines of sight are averaged in each panel. The dot-dashed lines are the mean transmission models of Madau (1995). The dotted lines are the mean transmission models of Meiksin (2006) but the updated version. The dashed line in the panel (c) is the average transmission of the Monte Carlo simulation by Bershady et al. (1999) (MC-Kim model).

Other effects

- Deblending! “photo”-z very dependent on excellent photometry
- Star/galaxy separation
- Dust law uncertainties, very dusty galaxies

Effects of Uncertainties on Cosmology

- Ma, Hu, & Huterer (2006) model photo-z uncertainties for WL tomography
- 31 ‘micro-bins’ with σ_z and δ_z, 5 tomographic bins (Gaussian pz, no cat outliers!)
- Use Fisher matrices to study information loss from photo-z errors
- Find strong degeneracy between w and photo-z bias
- control σ_z and δ_z to 0.003-0.01 to not degrade dark energy params by factor of 1.5 (per microbin)



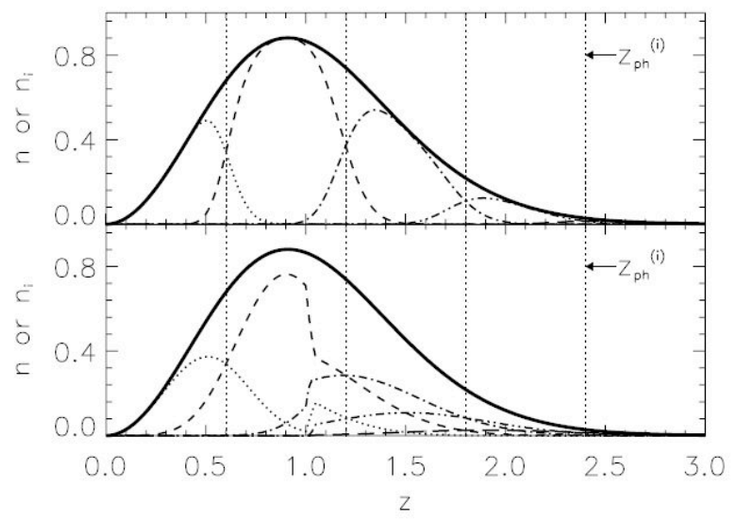


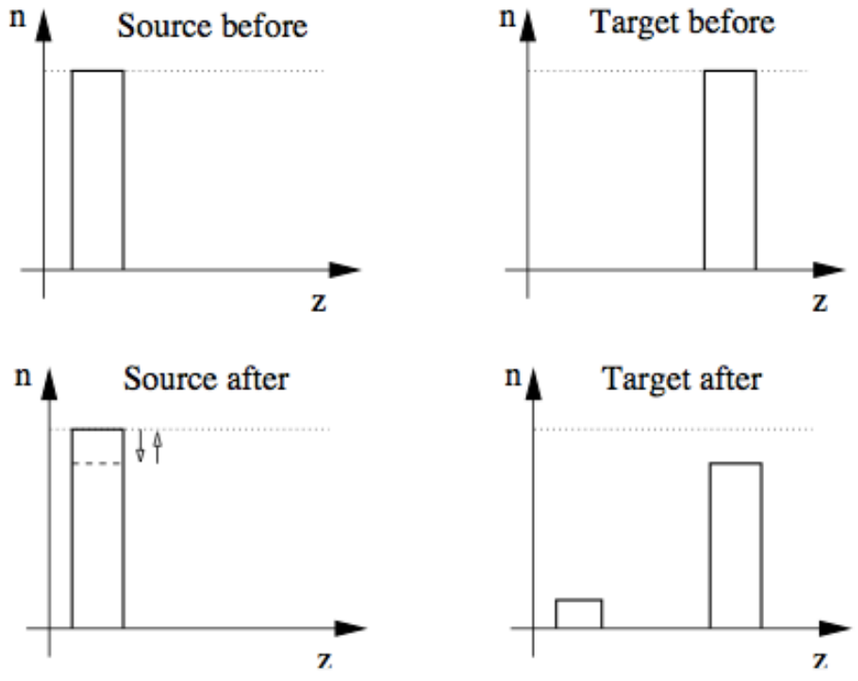
FIG. 2.—Source galaxy redshift distribution $n(z)$. *Top*: Photo-z model I. *Bottom*: Photo-z model II. The solid line is the overall galaxy distribution defined in eq. (2). The other lines are the true (spectroscopic) distributions that correspond to the sharp divisions in photo-z space (denoted by dotted vertical lines). [*See the electronic edition of the Journal for a color version of this figure.*]

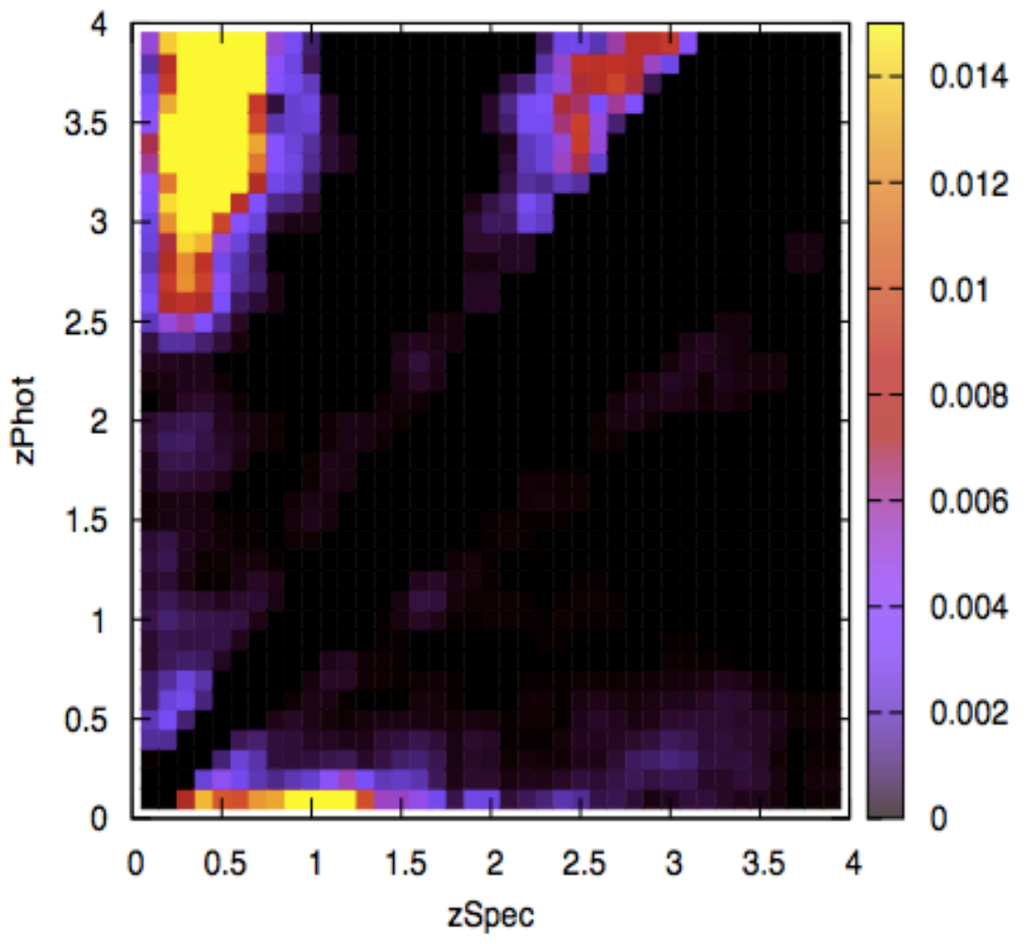
Catastrophic Outliers

- Bernstein & Huterer (2010) estimate the spec-z samples needed to constrain outlier rates for a fiducial SNAP like survey (assume complete spec-z coverage)

- Need ~10^6 if you use all photo-z's, dramatic drop if you restrict to z<2.5 (contaminant scattering to hi-z bins is a significant portion)*(no mag prior assumed!)

- For clustering recovery, estimate you need ~10% a priori knowledge of bias for outlier population, as degenerate with magnification bias



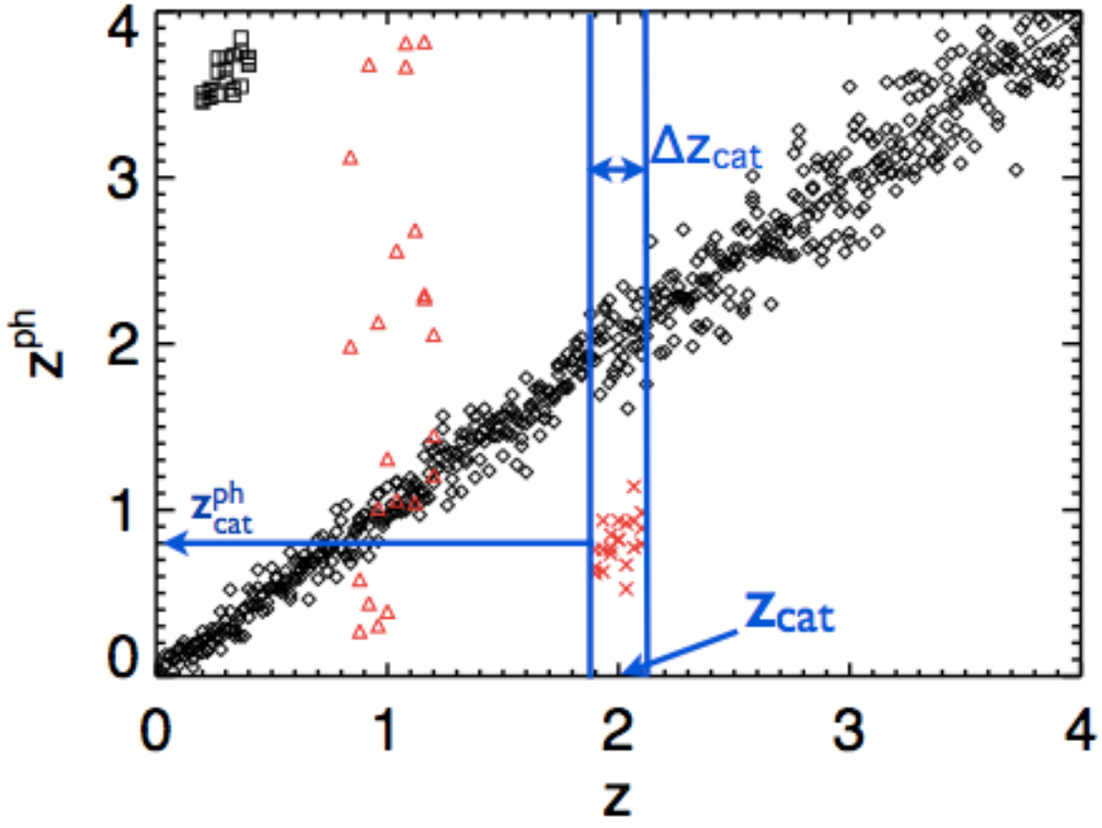


Catastrophic Outliers

- Hearin et al (2010) also look at cat outliers, both localized and widely scattered in photo-z

- Fraction of outliers must be <fewx10^{-4} to not degrade constraints on w_0 and w_a (worse near median of n(z))

- Cutting out z_p<0.3 and z_p>2.4 only degrades constraints ~20%, this is where most of the outliers live



Missing Template Simulation

- Abrahamse et al. (2011) ran MC sims of “LSST-like” photo-z’s, looked at PCA reconstructions of the resulting p(z) distributions

- Toy model: 20 input templates, leave one out of test set for each run

- Parameterization captures “realistic” catastrophic outlier behavior, similar amplitude to Ma, Hu, Huterer distributions

- Not propagated to cosmological predictions (yet)

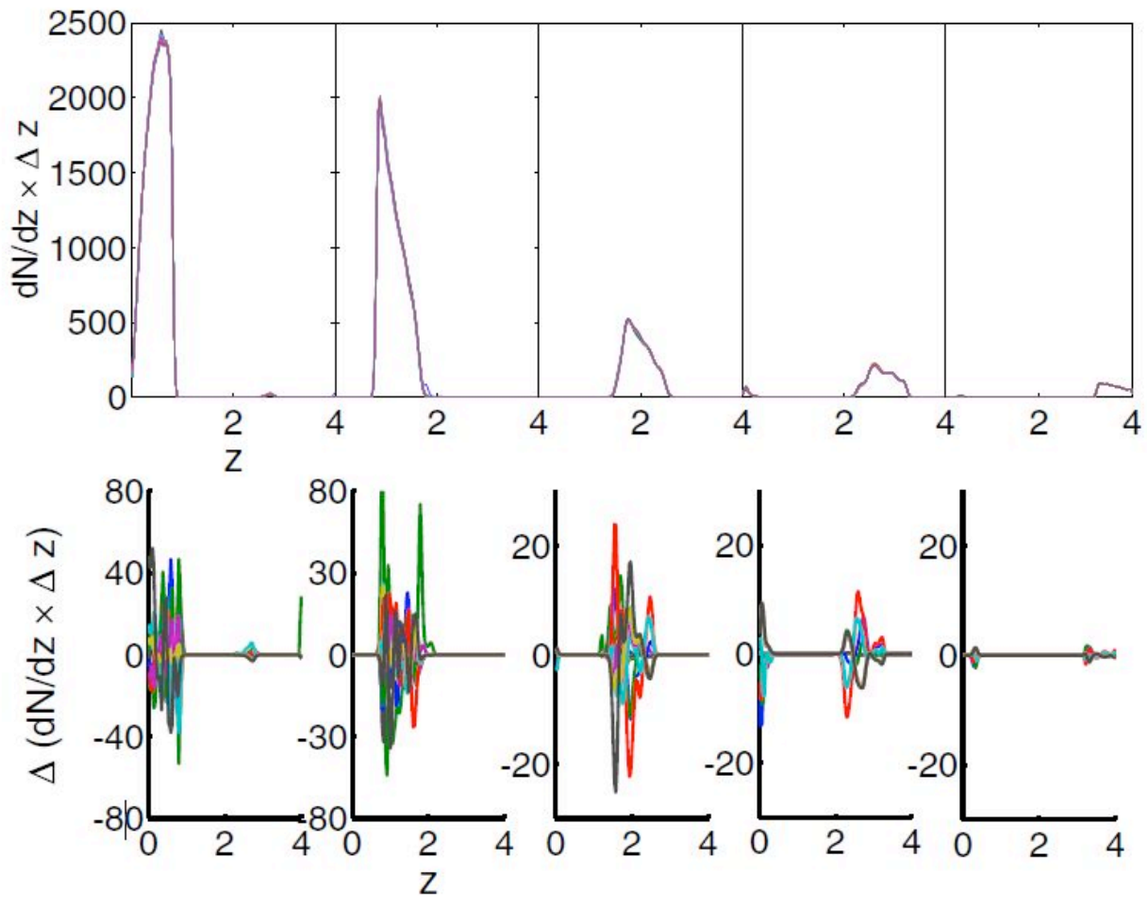


Figure 4. LRT template sets: top panels show the dN/dz’s from summing the $P(z)$ of galaxies in each collection defined for a given redshift range. The bottom panels show the residuals for each realization—that is, the difference between the fiducial (all 20) template set and the reduced (19 only of the 20) template sets, $\Delta dN/dz_i = dN/dz_{\rm fid} - dN/dz_i$, for each template set realization.

Conclusions

- Many potential problems with SEDs (or measurements) that can introduce bias and scatter to the photo-z's

- What can we do to accurately model these in simulations realistically to study/mitigate the effects? (we have some solutions, should discuss how to put into simulations)

- Propagate photo-z uncertainties to cosmological constraints to determine best solution to using/removing catastrophic outliers.