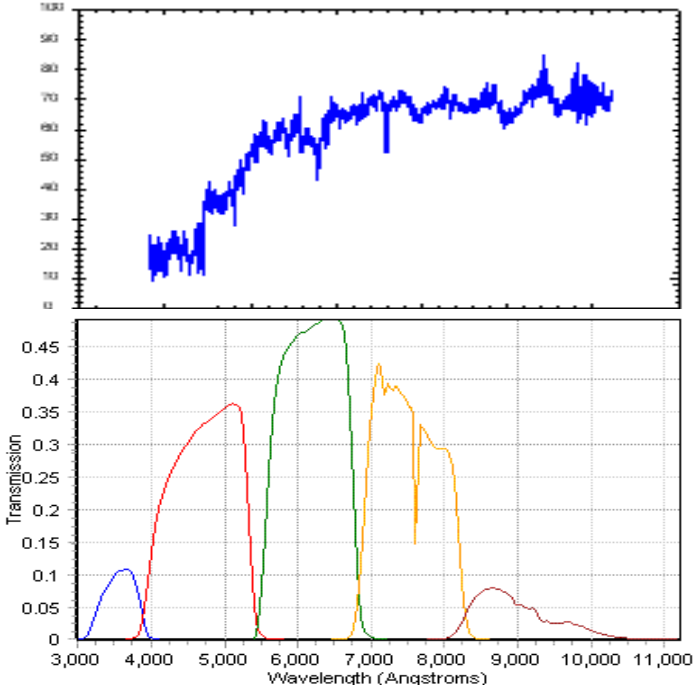
BEYOND PHOTOMETRIC REDSHIFTS USING BAYESIAN INFERENCE

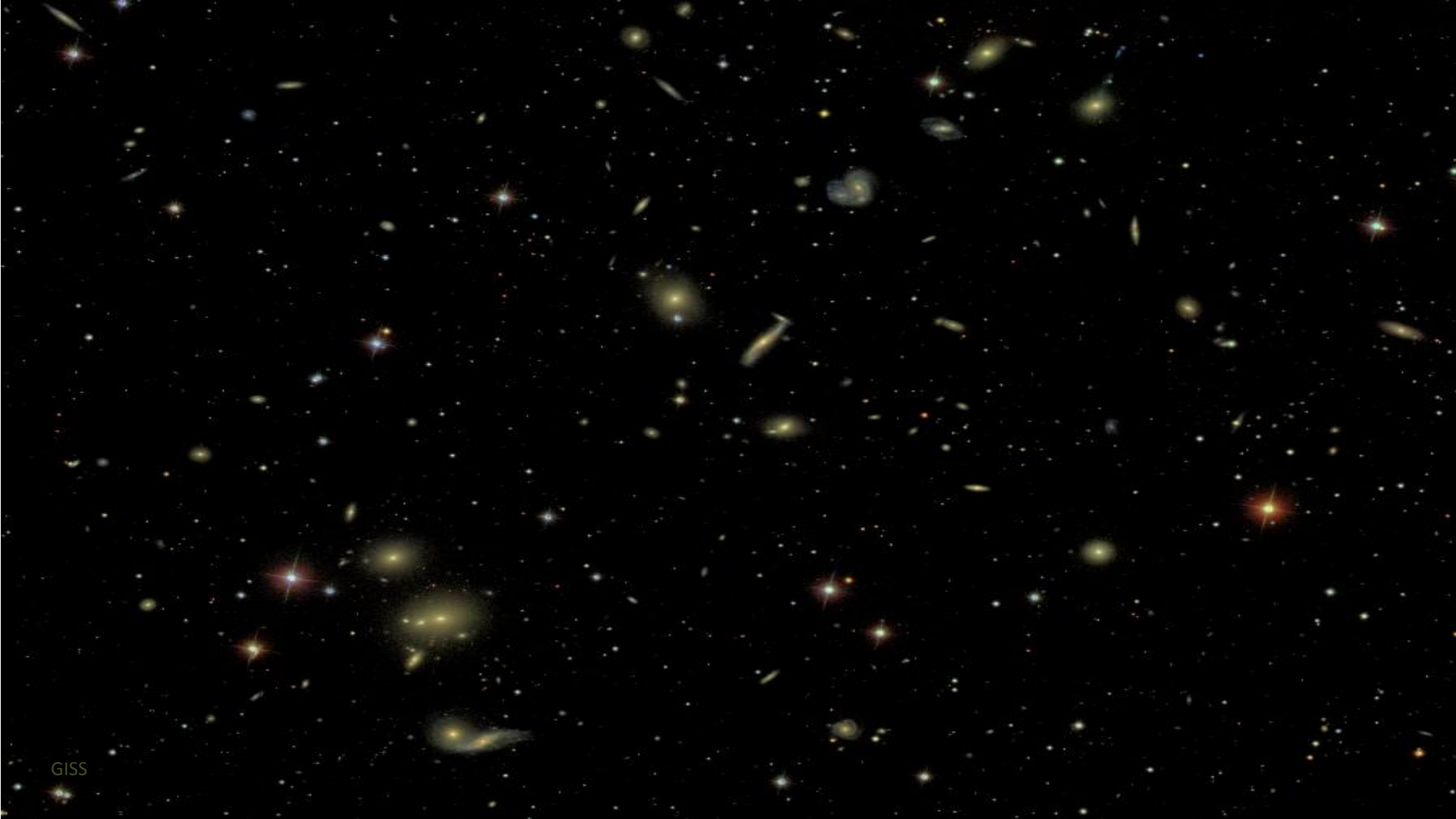
02/25/2011

Tamás Budavári / The Johns Hopkins University

Binning Photons

- Space (and time)
  - Position measurements
- Wavelength
  - Spectroscopy: narrow bins
  - Photometry: broad bands
- For example, SDSS
  - Spectro: $r < 17.7$ 1.6M sources
  - Photo: $r < 22+$ 360M sources




GISS

Information Content

- Incredible amount of information in the images
  - Traditionally very pragmatic approaches

THERE IS MORE IN THE DATA!

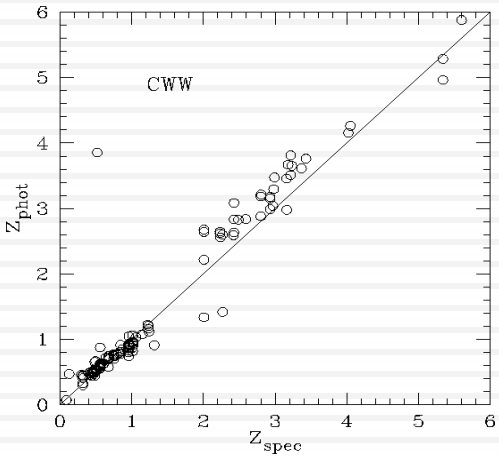
- Models are useful
  - Source catalogs are model-based extractions
  - Statistical and theoretical models of SEDs

Multicolor Challenges

- Cross-identification of sources
  - To assemble multicolor catalogs
- Drop-outs from sky coverage
  - To constrain fluxes not detected
- Constraining physical properties
  - To interpret the data

Photometric Redshifts

- Empirical methods
- Template fitting



Historical Overview

- Baum (1962)
  - Compared average SEDs (9 bands) of ellipticals in clusters
- Koo (1985)
  - Color-shape diagrams of and Bruzual-Charlot iso-*z* tracks (-U+J+F-N)
- Loh & Spillar (1986)
  - Template fitting with 6 non-standard filters and 34 known galaxies
- Connolly et al. (1995)
  - Redshift assumed to be a linear or quadratic function of magnitudes
- Steidel et al. (1996)
  - U dropout due to Ly-alpha blanketing (< 912Å) to select $z > 2.25$

Empirical Methods

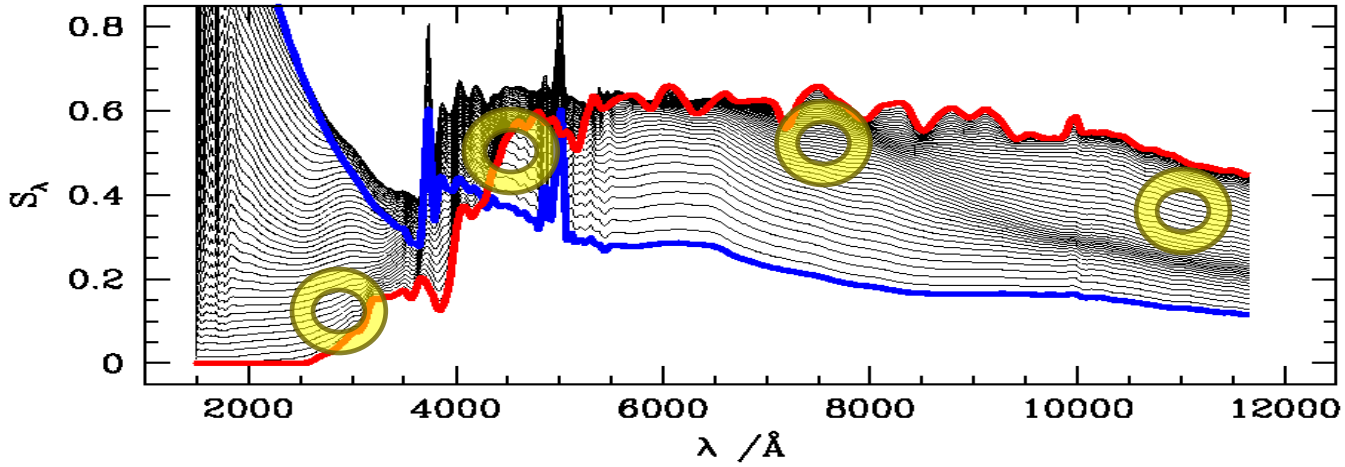
- Regression $z = F(m)$
  - Polynomial fitting
  - Piecewise linear fits
  - Nearest neighbors
  - Neural Nets
  - Support Vector Machines

- Errors?

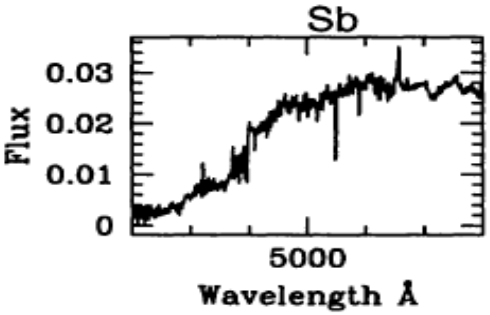
Template Fitting

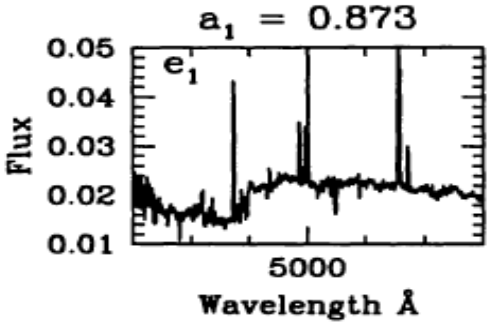
Tamás Budavári

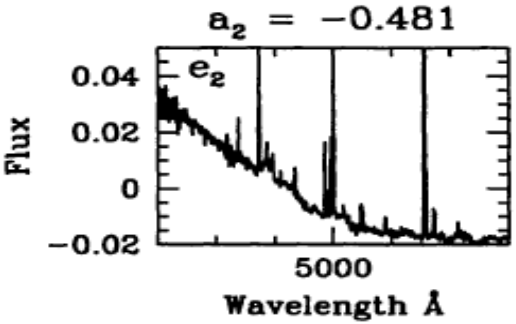
- Compare reference spectra to measurements
  - At different redshifts – MLE with Gaussian

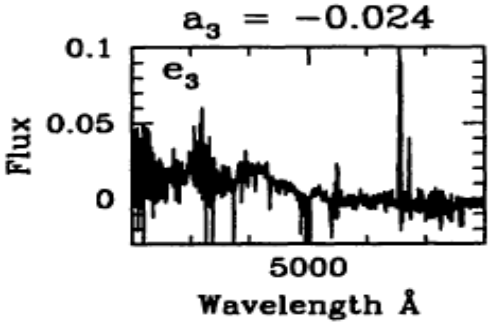


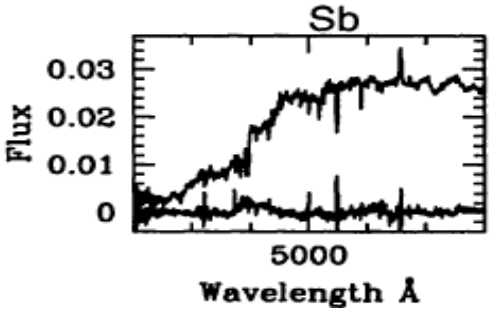
Eigenspectra









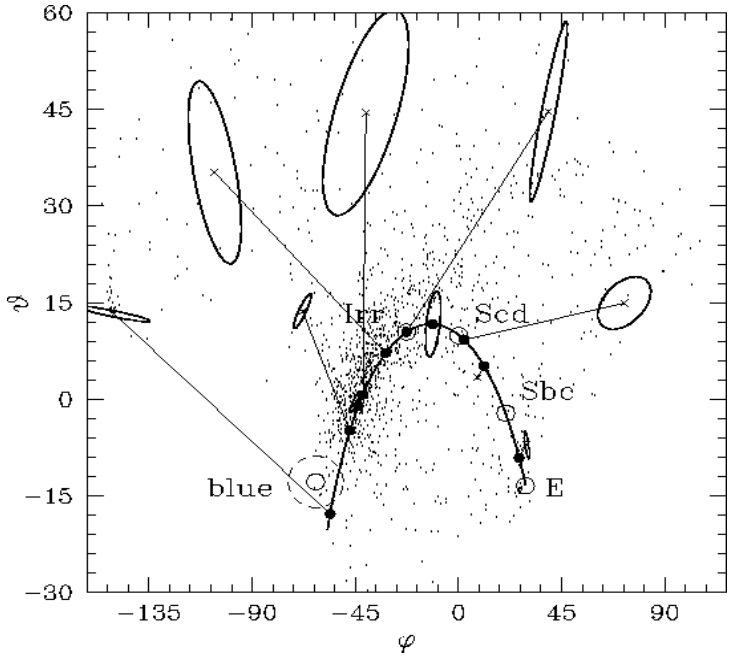


Interpolation

- Two mixing angles encode the spectral type when using three eigenspectra

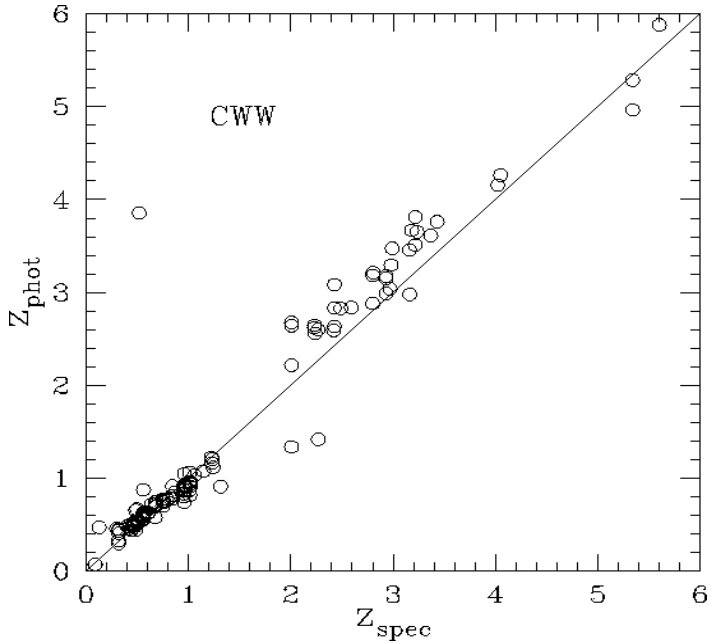
1D type parameter connects the CWW templates:

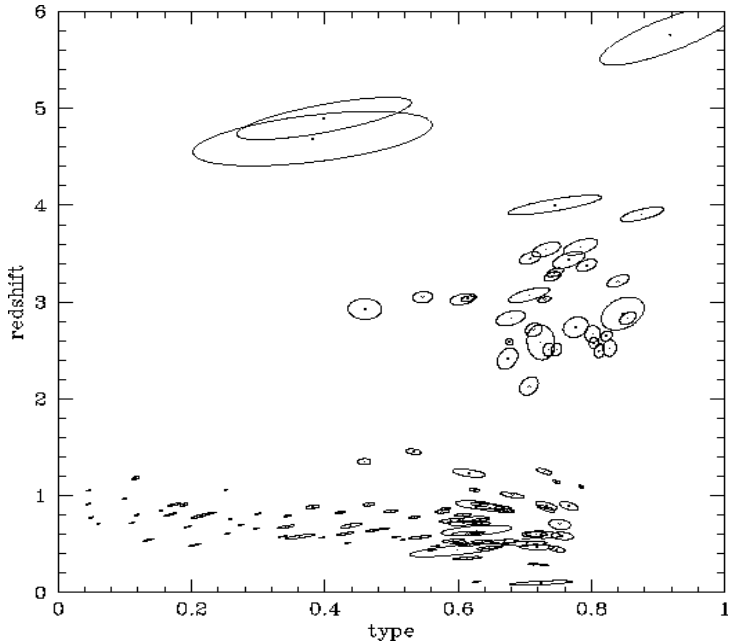
Ell	0
Sbc	0.13
Scd	0.35
Irr	0.56
blue	1



Consistent Redshift and Type

Tamás Budavári





Traditional Methods

Empirical Methods

Pure fit for *z* over the color hyper-plane

+ fast redshift prediction

- fitting functions and training sets needed for every survey

- unreliable extrapolation

Template Fitting

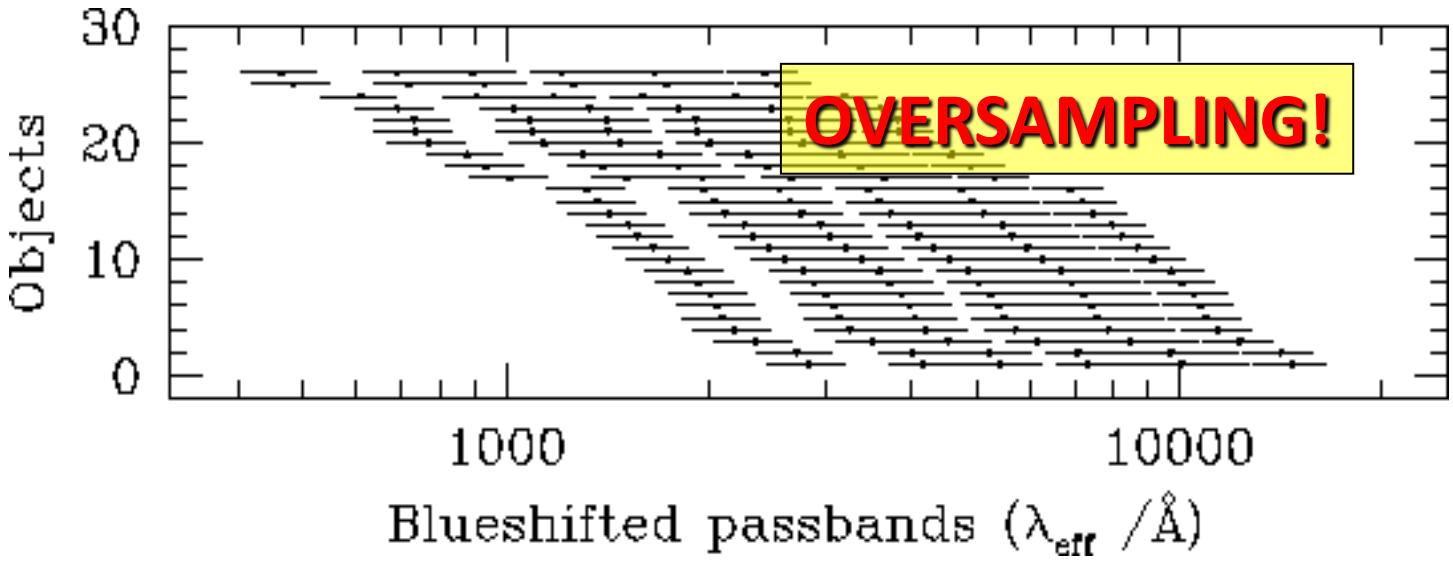
Comparison of known SEDs to photometry

+ physics is in templates, no training set required

+ more outcome: spectral type, reddening, etc.

- templates used as come

Templates from Photometry?



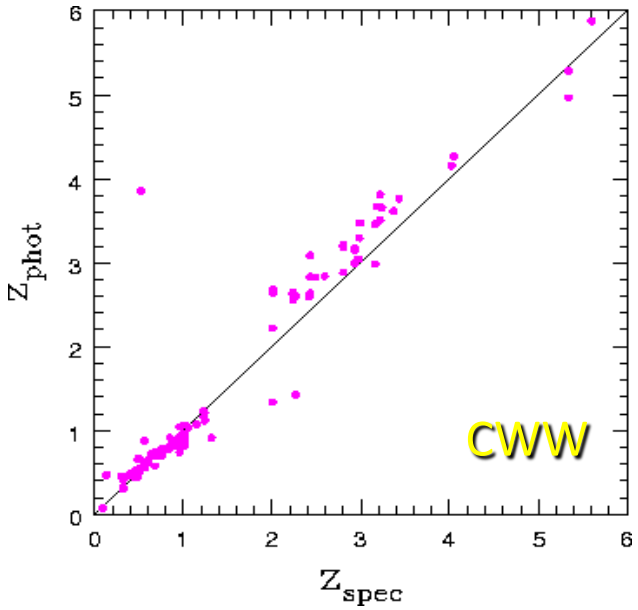
Simple Repair

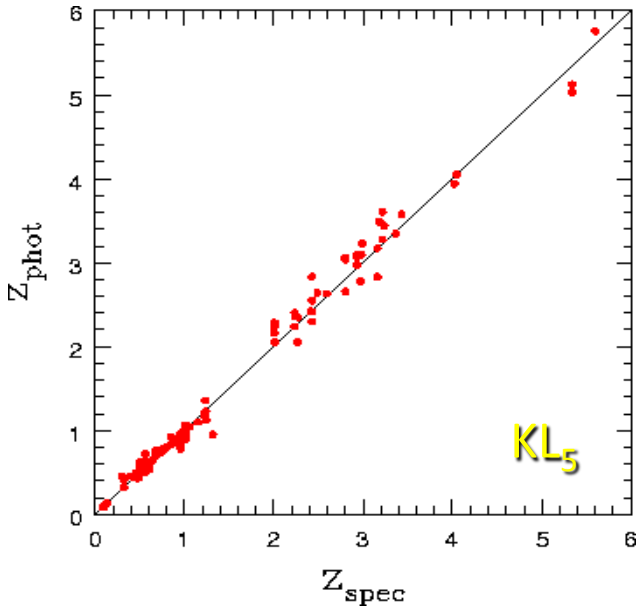
For each galaxy:

1. Compute type from known redshift
2. Derive estimate spectrum
3. Correct spectrum to photometry

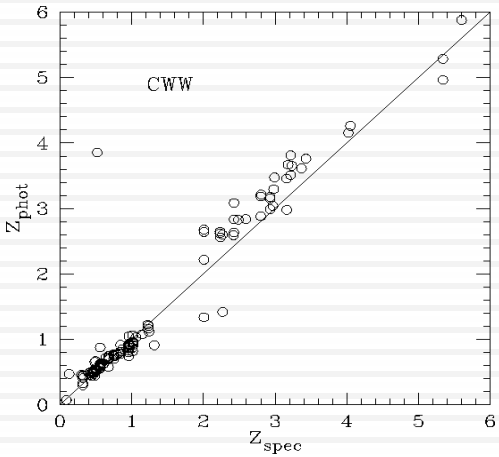
Build new basis from corrected spectra

HDF/NICMOS in 1999



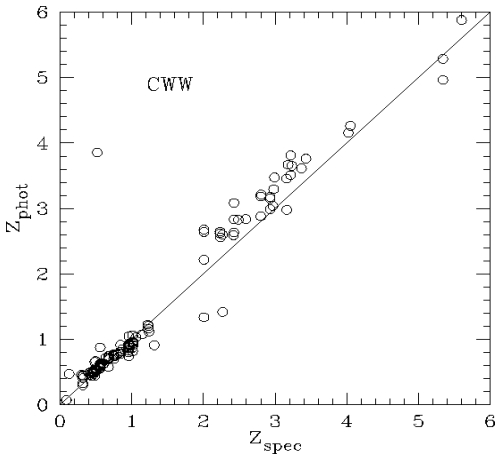


A Unified Framework



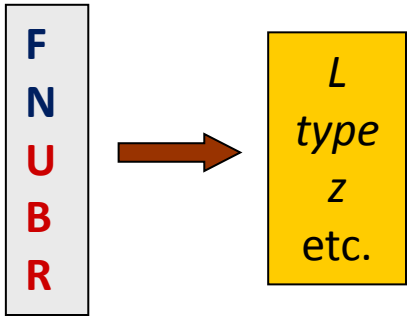
Photometric Redshifts

- Traditional ways
  - Empirical methods: fitting function, k-NN, neural nets
  - Template fitting: compare to model SEDs
- Very different techniques
  - Approx of low dimensional subspace
  - Single scalar estimates
- Many open questions

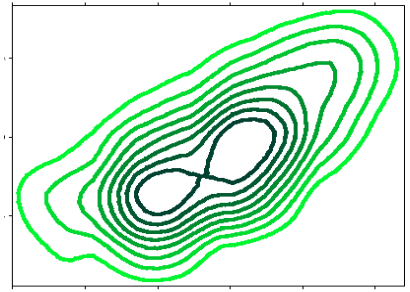


Photometric Inversion

- The general inversion problem
  - Constrain various properties consistently
  - Propagate uncertainties and correlations



- Assumption of functional relation is too simple
  - Probability density functions instead
  - Scientific analyses should use full PDFs



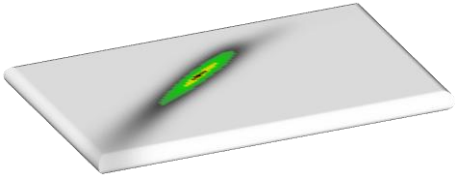
A Unified Framework

- Training and Query sets with different observables

$$T: \quad \{\boldsymbol{x}_t, \boldsymbol{\xi}_t\}_{t \in T}$$
$$Q: \quad \{\boldsymbol{y}_q\}_{q \in Q}$$
$$M: \quad \boldsymbol{\theta}$$

- Model yields observables for given parameter
  - Prediction via $p(\boldsymbol{x}, \boldsymbol{y} | \boldsymbol{\theta}, M)$ and has prior $p(\boldsymbol{\theta} | M)$
  - Also folds in the photometric accuracy

- We are after $p(\boldsymbol{\xi} | \boldsymbol{y}_q, M)$

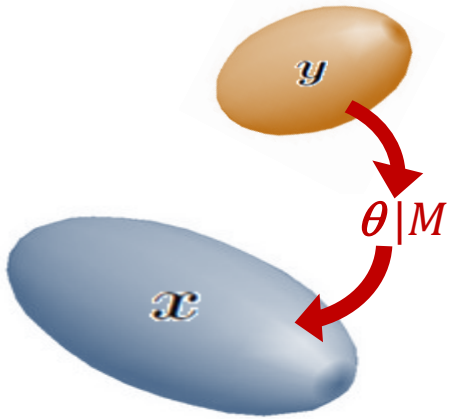


Mapping Observables

- The model provides the transformation rule

$$p(\boldsymbol{x}|\boldsymbol{y}_q, M) = \int d\boldsymbol{\theta}\; p(\boldsymbol{x}|\boldsymbol{\theta}, M)\, p(\boldsymbol{\theta}|\boldsymbol{y}_q, M)$$

with $$p(\boldsymbol{\theta}|\boldsymbol{y}_q, M) = \frac{p(\boldsymbol{\theta}|M)\, p(\boldsymbol{y}_q|\boldsymbol{\theta}, M)}{p(\boldsymbol{y}_q|M)}$$



- Think empirical conversion formulas but better
  - For example, from *UJFN* to *ugriz* with errors

Empirical Relation

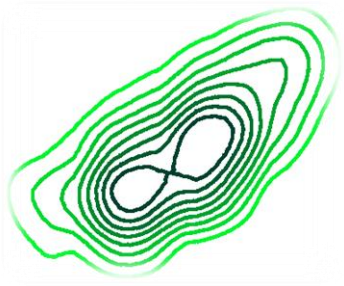
- Usually just assume a function $\boldsymbol{\xi} = \hat{\boldsymbol{\xi}}(\boldsymbol{x})$
  - Wrong! We know there are degeneracies: $p(\boldsymbol{\xi}|\boldsymbol{x})$
- There is a more general relation
  - Usual restriction is $p(\boldsymbol{\xi}|\boldsymbol{x}) = \delta(|\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}(\boldsymbol{x})|)$
  - A better estimation: $$p(\boldsymbol{\xi}|\boldsymbol{x}) = \frac{p(\boldsymbol{\xi}, \boldsymbol{x})}{p(\boldsymbol{x})}$$

Properties of Interest

- The final distribution is

$$p(\boldsymbol{\xi}|\boldsymbol{y}_q, M) = \int d\boldsymbol{x}\; p(\boldsymbol{\xi}|\boldsymbol{x})\, p(\boldsymbol{x}|\boldsymbol{y}_q, M)$$

- Estimate by the mean
  - If the result is unimodal (no guarantee)



Properties of Samples

- Often we need subsamples
  - Redshift slices and/or color cuts
  - Complicated interaction with photo-z errors
  - Variable kernel deconvolution
- Natural with the PDFs
  - $p(\boldsymbol{\xi}|Q,M) = \langle p(\boldsymbol{\xi}|\boldsymbol{y}_q,M)\rangle_{q\in Q}$
  - Selection boundaries in *Q*

Selection Effects

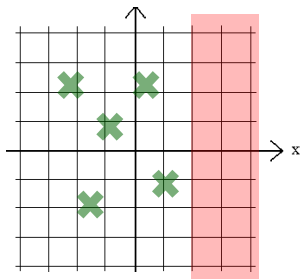
- Measure of reliability is the prob of *q* making the cuts
  - Using the window function

$$P(W|\boldsymbol{y}_q, M) = \int dx\; P(W|\boldsymbol{x})\, p(\boldsymbol{x}|\boldsymbol{y}_q, M)$$

- Measured relation may be biased
  - Include all observables in the selection!

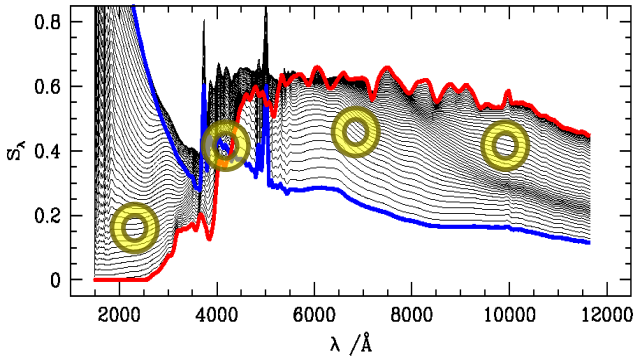
$$p(\boldsymbol{\xi}|\boldsymbol{x}, T) = \frac{p(\boldsymbol{\xi}|\boldsymbol{x})\, P(T|\boldsymbol{x}, \boldsymbol{\xi})}{P(T|\boldsymbol{x})}$$

- For instance
  - Cannot use just colors, if there was a magnitude cut
  - Cannot use just fluxes, if cut on morphology



Template Fitting

- Artificial training set $\{x_t, \xi_t\} = \{\bar{x}(\theta_t), \bar{\xi}(\theta_t)\}$
  - From a grid of model points
  - No errors $p(x|\theta, M) = \delta(|x - \bar{x}(\theta)|)$



- Analytic result

$$p(\xi|y_q, M) \propto \sum_{t \in T} \delta(|\xi - \xi_t|)\, p(\theta_t|M)\, N(y_q|\bar{y}(\theta_t), C_q)$$

GRID PRIOR LIKELIHOOD

Empirical Estimates

- Simplest assumption

$$p(z|x_t, M) = \delta(z - z_t)$$

- Generalized kernel regression
  - Average
  - Weighting
    - Local lin, RF, Wz

$$\bar{z}_q = \int dz\ z \int dx\ p(z|x)\, p(x|y_q, M)$$

$$\bar{z}_q \approx \sum_{t \in T} z_t\, p(x_t|y_q, M)\, \Delta x_t = \sum_{t \in T} z_t P_{tq}$$

- Need representative training sets without Δx

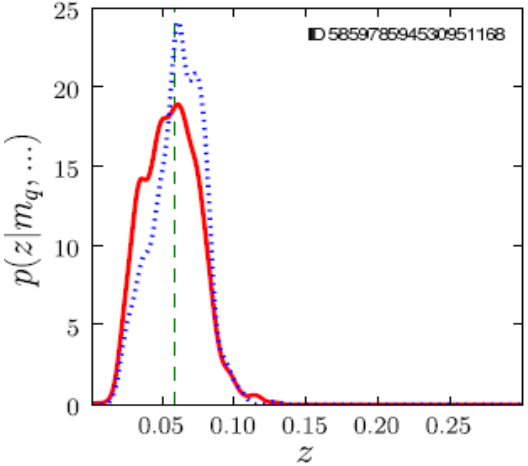
Improved Empirics

- Minimalist model
  - Normal distributions, same quantities: $\bar{x}(\theta) = \theta$ and $\bar{y}(\theta) = \theta$
  - With simple prior, mapping is analytic , e.g., for flat

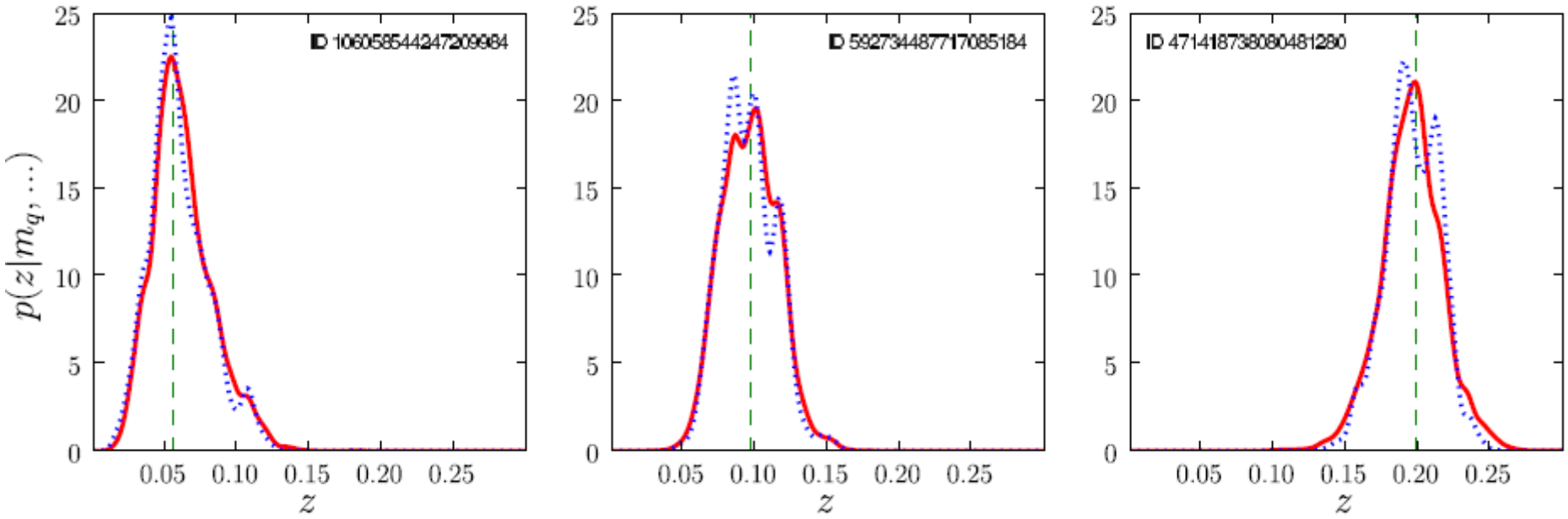
$$p(\boldsymbol{x}_t|\boldsymbol{y}_q, M) = \int d\boldsymbol{\theta}\; N(\boldsymbol{x}_t|\boldsymbol{\theta}, \mathbf{C}_t)\, N(\boldsymbol{\theta}|\boldsymbol{y}_q, \mathbf{C}_q)$$

- Empirical relation
  - Ratio: KDE of joint and marginalized
- Numerical summation

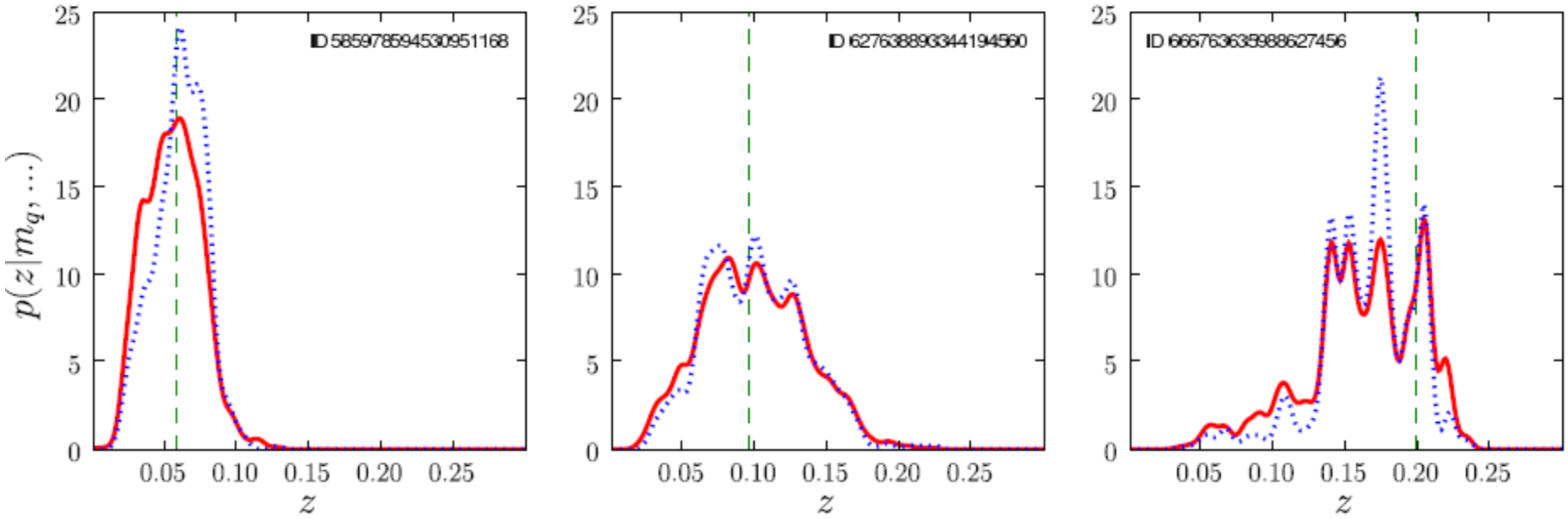
$$p(\boldsymbol{\xi}_r|\boldsymbol{y}_q, T, M) \propto \sum_{t\in T} p(\boldsymbol{\xi}_r|\boldsymbol{x}_t, T)\frac{p(\boldsymbol{x}_t|\boldsymbol{y}_q, M)}{p(\boldsymbol{x}_t|T)}$$



Red Galaxies



Blue Galaxies



Advanced Methods

- Mapping observables via models
  - Any complete basis on wavelength range
  - Physics is in the prior!
- Relation of properties
  - Conditional densities
- Empirical but with templates
  - Unified framework at its best

$$T: \quad \{\boldsymbol{x}_t, \boldsymbol{\xi}_t\}_{t \in T}$$
$$Q: \quad \{\boldsymbol{y}_q\}_{q \in Q}$$
$$M: \quad \boldsymbol{\theta}$$
$$p(\boldsymbol{\theta}|M)$$
$$p(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{\theta}, M)$$

Concerning Priors

- Measured densities in the Query set
  - KDE, Voronoi?

$$p(\boldsymbol{y}|Q)$$

- Consistent models should match that
  - Equation

$$p(\boldsymbol{y}|M) = \int d\boldsymbol{\theta}\, p(\boldsymbol{y}|\boldsymbol{\theta}, M)\, p(\boldsymbol{\theta}|M)$$

- Deconvolution yields empirical prior
  - Cf. naïve repair

$$p(\boldsymbol{y}|M) = p(\boldsymbol{y}|Q)$$

Summary

- **Bayesian approach** places former heuristics on a firm statistical basis
  - Existing photo-z methods are special cases of the more general approach
- Enables us to properly include physical priors
  - E.g., fluxes for xmatch, SEDs and positions for photoz
- Opens the door for optimal, next generation techniques
  - Missing ingredients: fast algorithms, indexing, new hardware, GPUs
- Full inversion of photometry is just ahead of us