Oskar Klein centre

Funded by the European Union

European Research Council
Established by the European Commission

pop-cosmos:

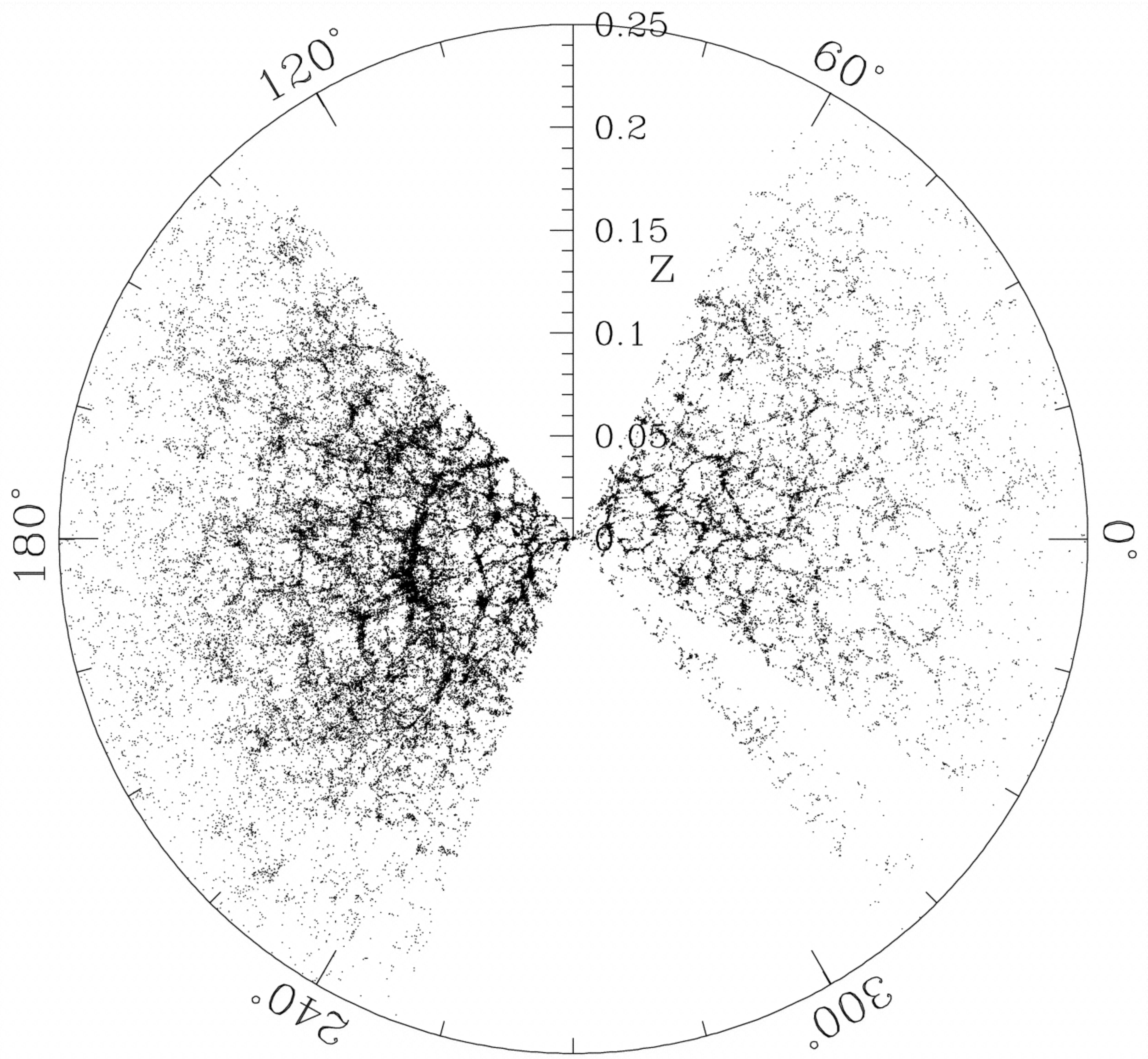
Comprehensive forward modelling of photometric galaxy surveys

<u>Stephen Thorp</u>, Justin Alsing, Sinan Deger, Hiranya Peiris, Boris Leistedt, Daniel Mortlock, Joel Leja, Arthur Loureiro

arXiv 2402.00935

Large Scale Structure Cosmology

- Map the matter distribution in the universe using spectra or photometry of galaxies
- Galaxy clustering (positions), or weak lensing (shapes) sensitive to cosmological parameters
- But… need redshifts



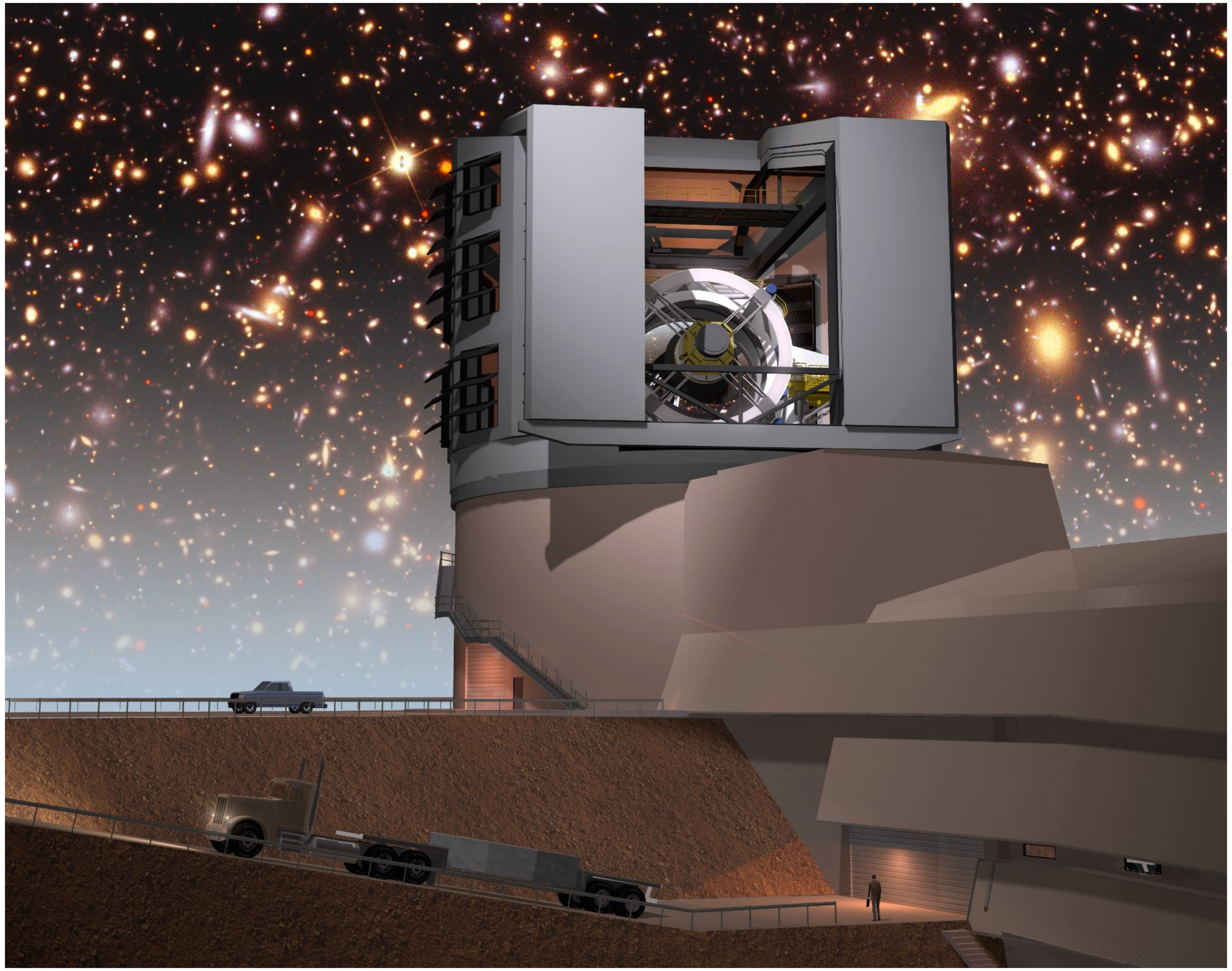
SDSS: *Blanton et al. (2003)*

Coming Soon…

VERA C. RUBIN OBSERVATORY

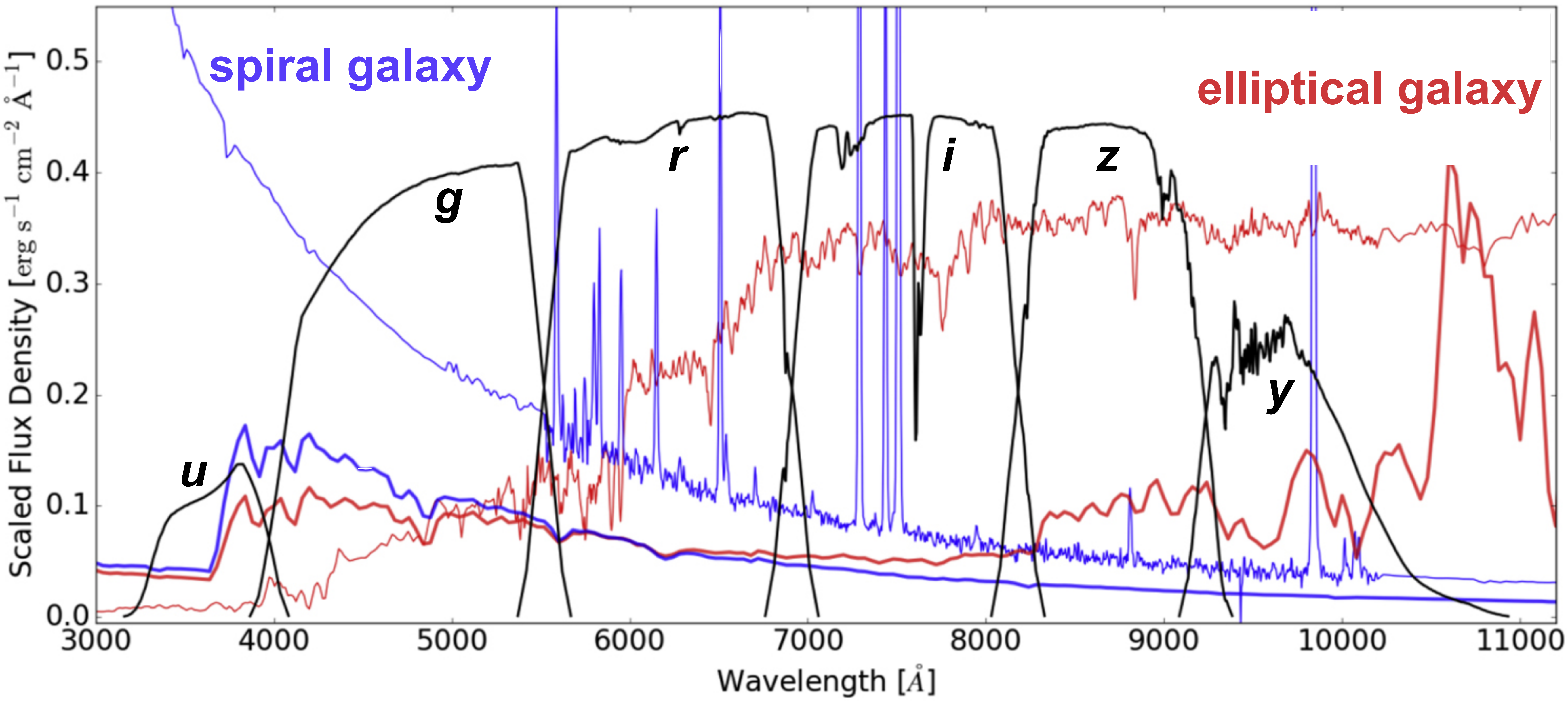
LSST Legacy Survey of Space and Time

- 18,000 deg^2
- Deep imaging in $ugrizy$
- 10 year LSST survey
- Single epoch: $r \lesssim 24\,\mathrm{mag}$;
  10 year co-add: $r < 26.9\,\mathrm{mag}$
- Billions of galaxies: impossible to get spectra
- Need *photometric* redshifts



Credit: Rubin Observatory

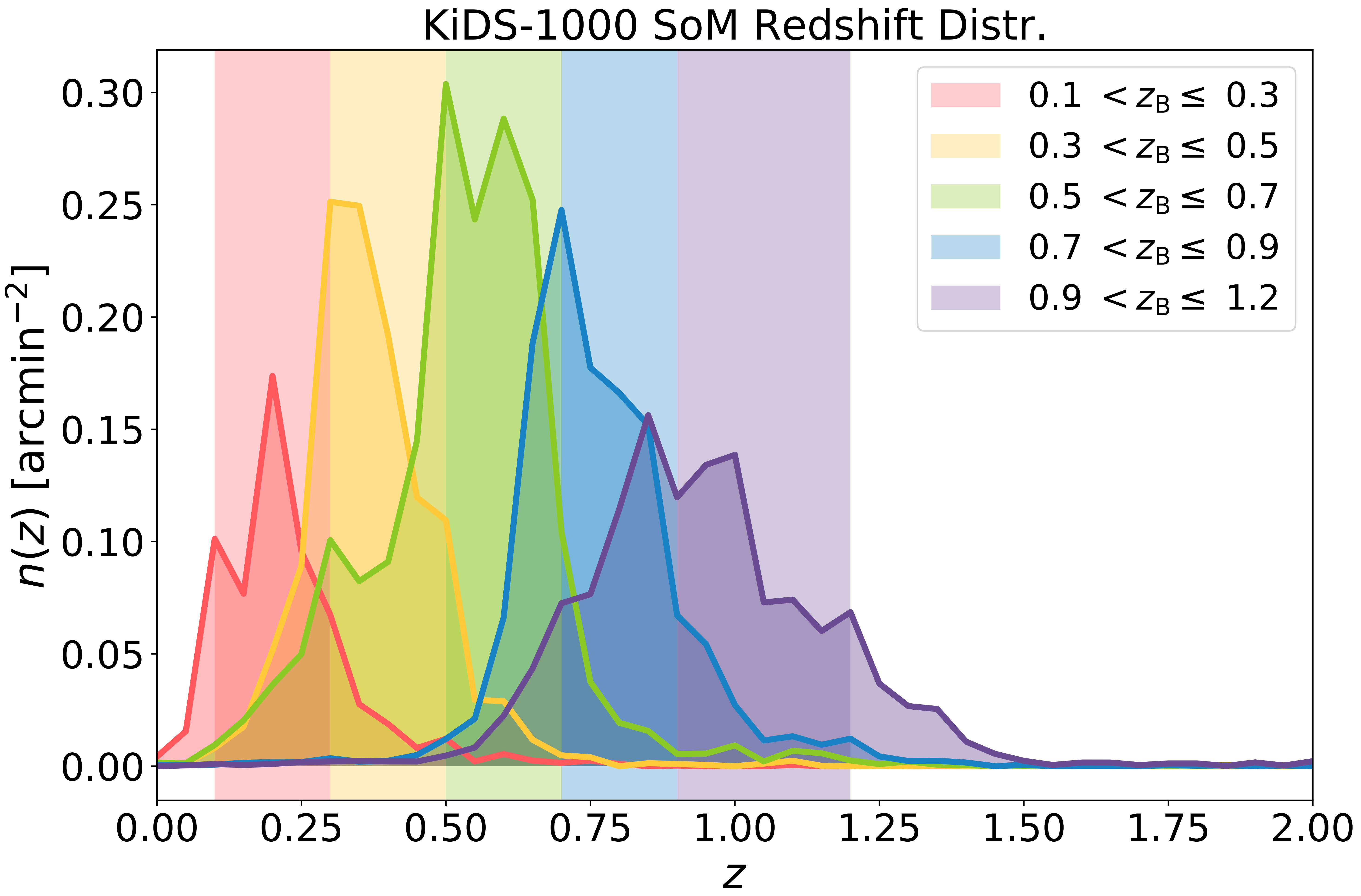
Photometry



Graham et al. (2017)

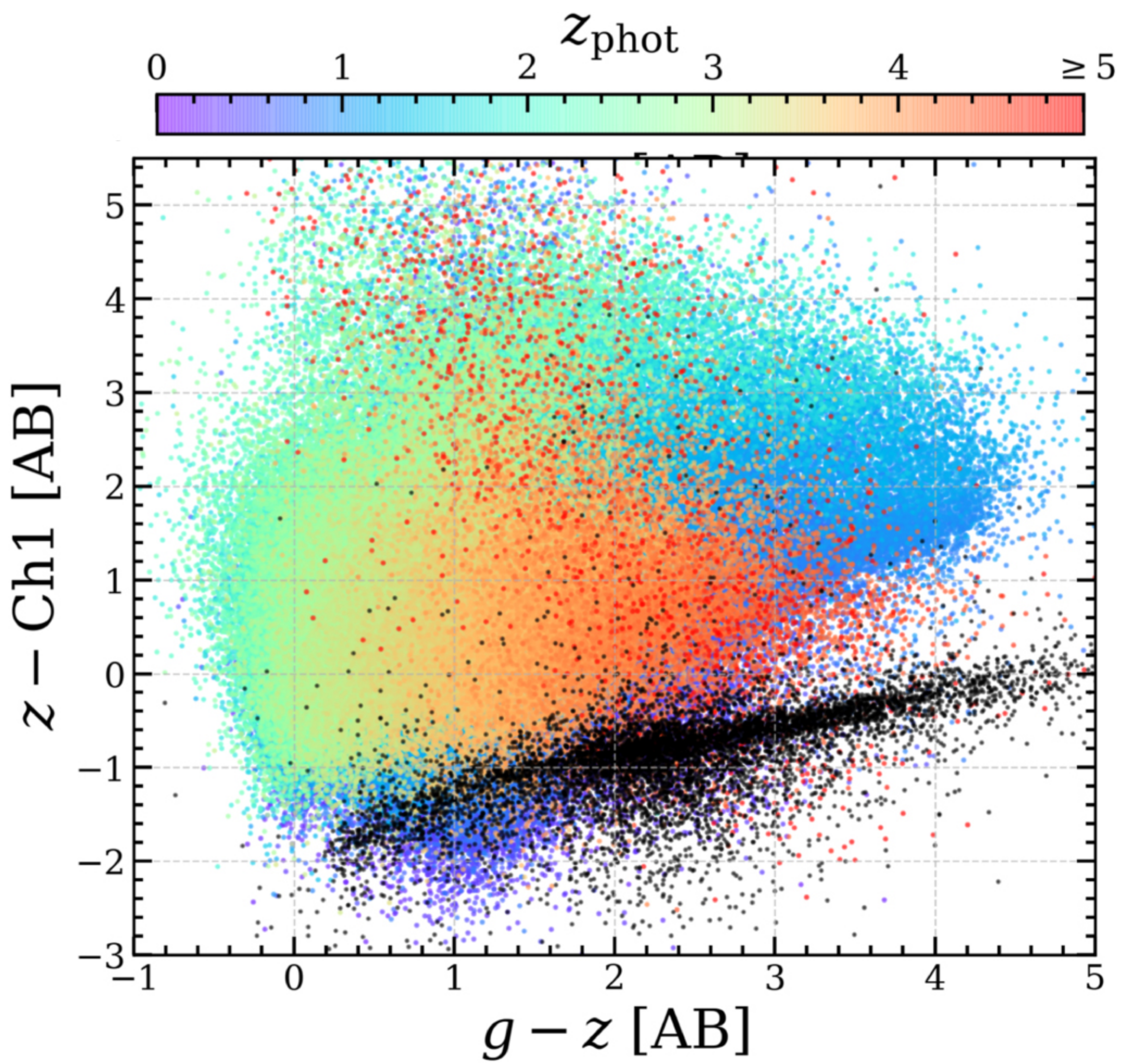
Photometric Redshifts

We need accurate characterisation of $n(z)$ in tomographic bins



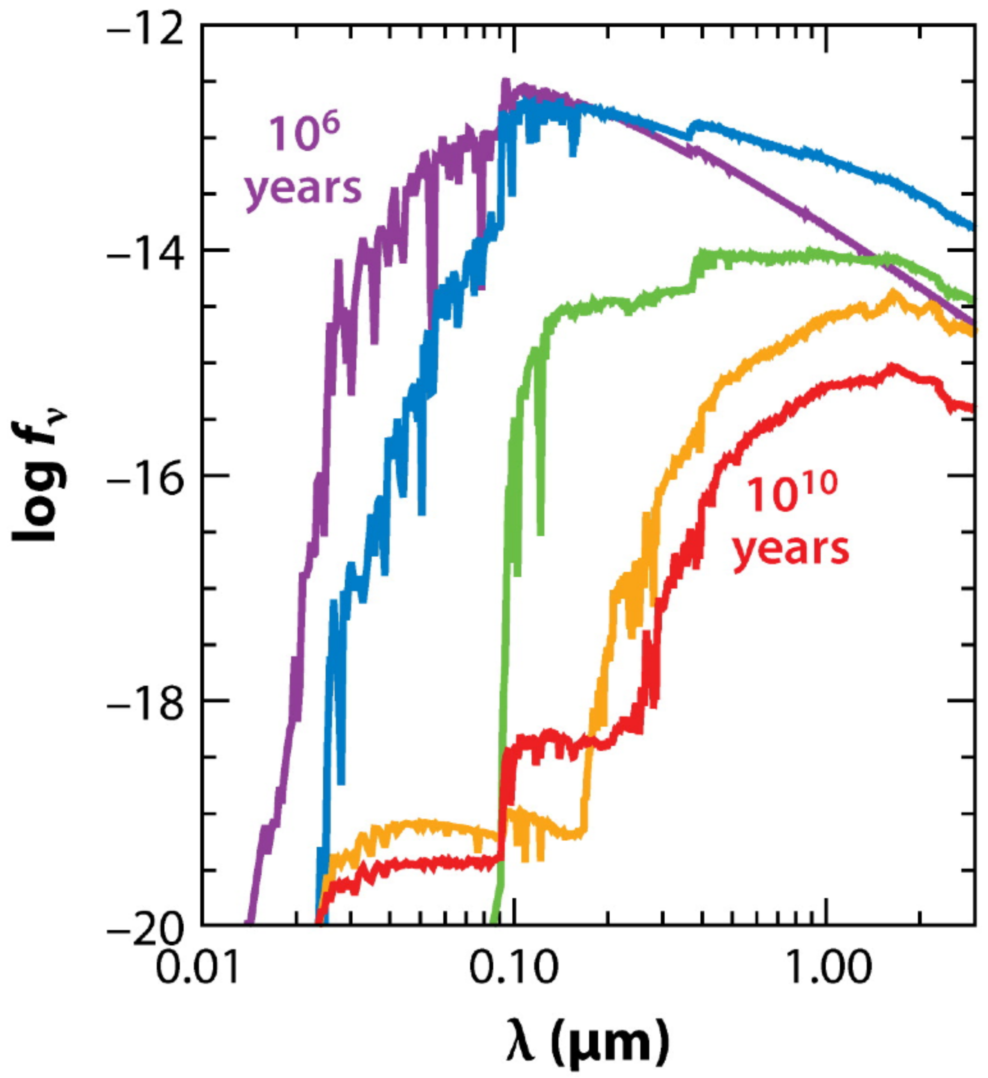
KiDS: Loureiro et al. (2022)

Forward Modelling a Photometric Catalog

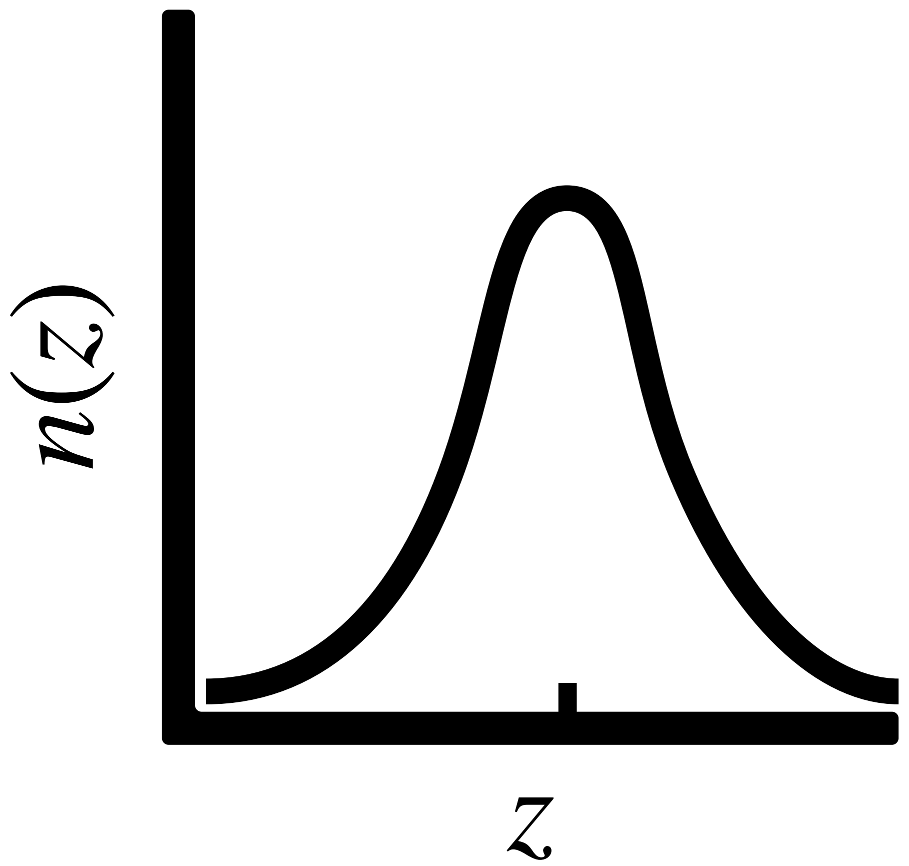


Conroy (2013); Weaver et al. (2022)

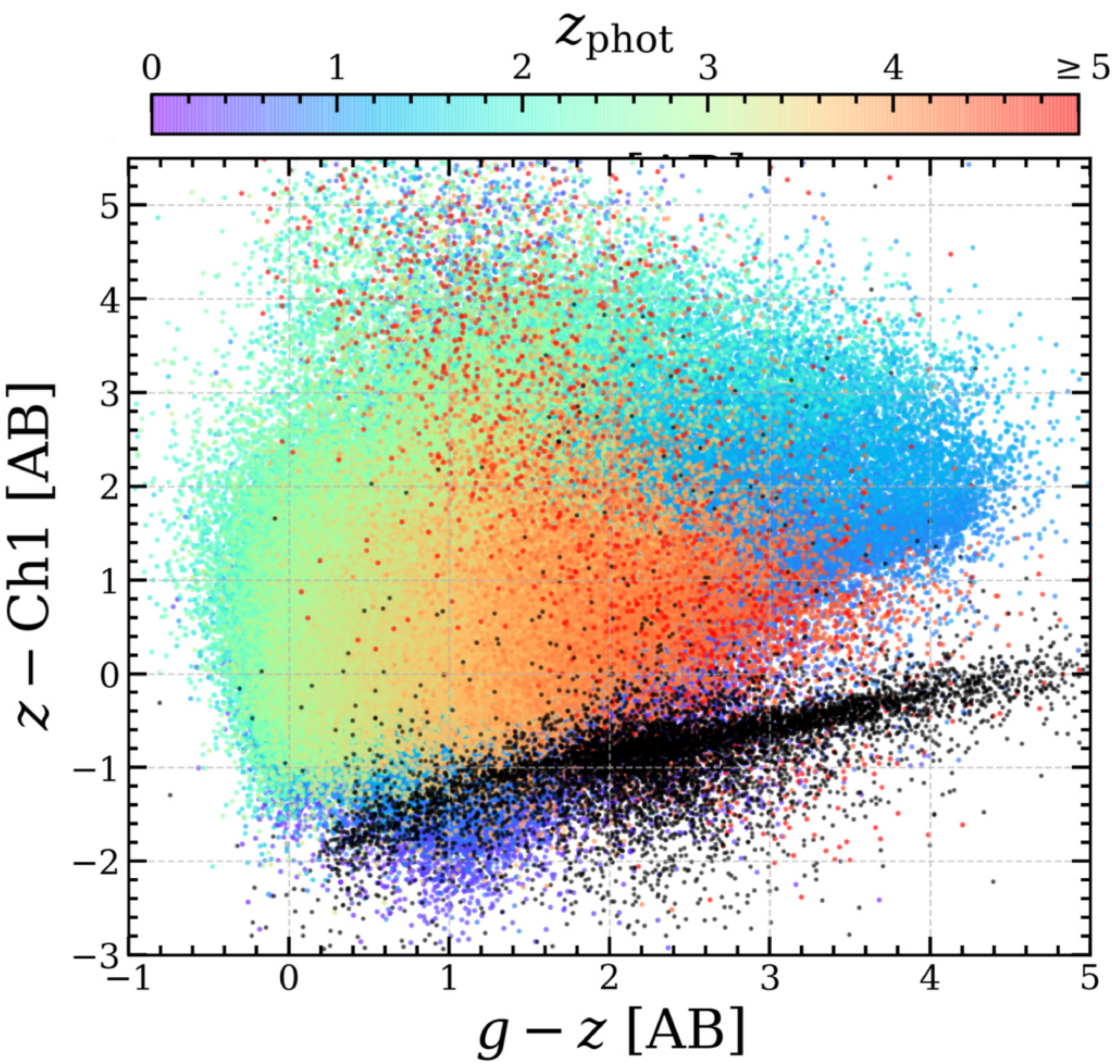
Forward Modelling a Photometric Catalog



+



=



Conroy (2013); Weaver et al. (2022)

Forward Modelling a Photometric Catalog

Forward Modelling a Photometric Catalog

1. Draw from galaxy population distribution over physical properties and z

Forward Modelling a Photometric Catalog

1. Draw from galaxy population distribution over physical properties and z
2. Compute fluxes via stellar population synthesis

Forward Modelling a Photometric Catalog

1. Draw from galaxy population distribution over physical properties and z
2. Compute fluxes via stellar population synthesis
3. Draw realistic uncertainties from an uncertainty model

Forward Modelling a Photometric Catalog

1. Draw from galaxy population distribution over physical properties and z
2. Compute fluxes via stellar population synthesis
3. Draw realistic uncertainties from an uncertainty model
4. Add noise based on the uncertainties

Forward Modelling a Photometric Catalog

1. Draw from galaxy population distribution over physical properties and z
2. Compute fluxes via stellar population synthesis
3. Draw realistic uncertainties from an uncertainty model
4. Add noise based on the uncertainties
5. Apply selection

Forward Modelling a Photometric Catalog

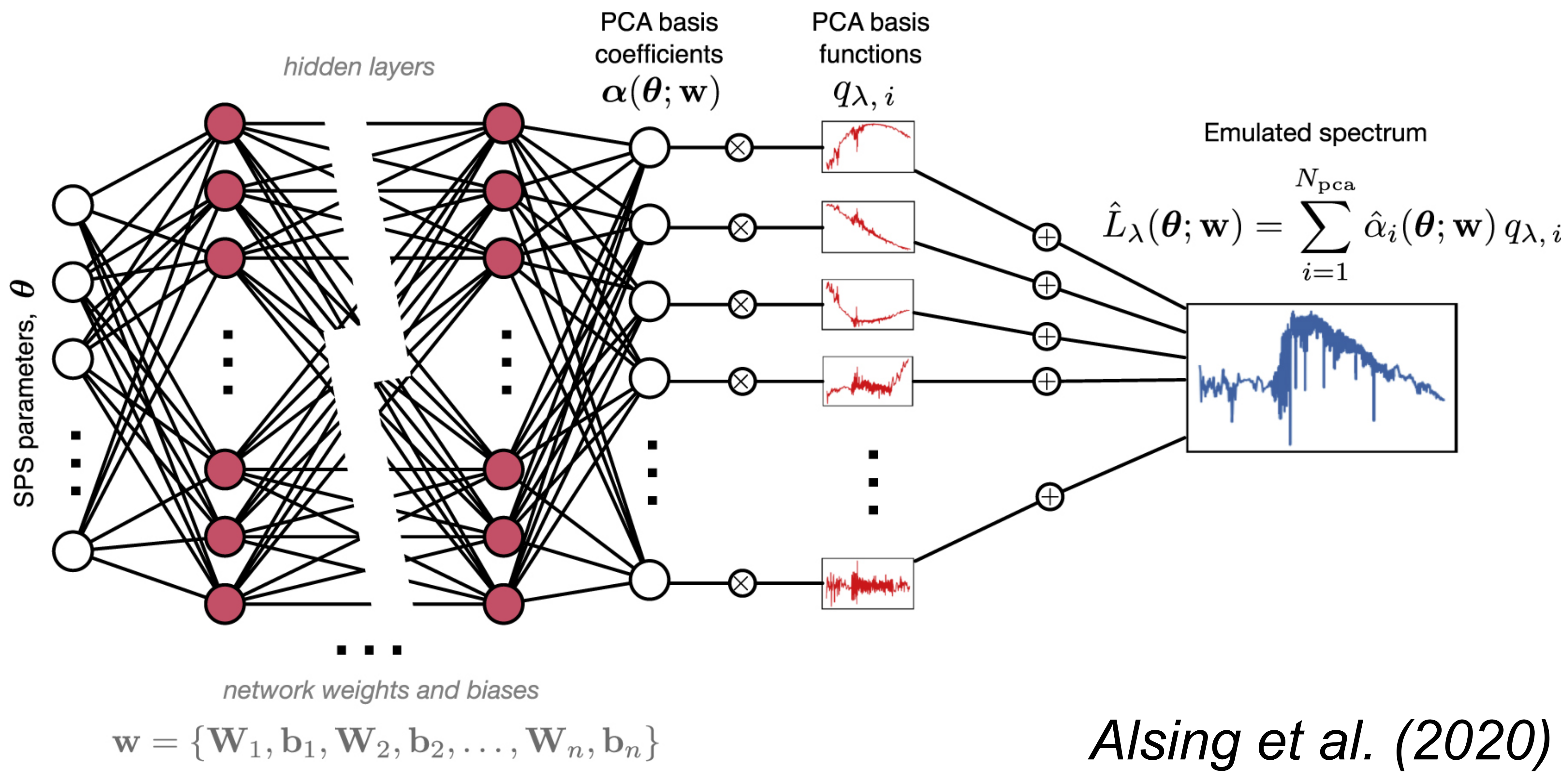
1. Draw from galaxy population distribution over physical properties and z
2. Compute fluxes via stellar population synthesis
3. Draw realistic uncertainties from an uncertainty model
4. Add noise based on the uncertainties
5. Apply selection
6. Compare model fluxes to data

pop-cosmos framework

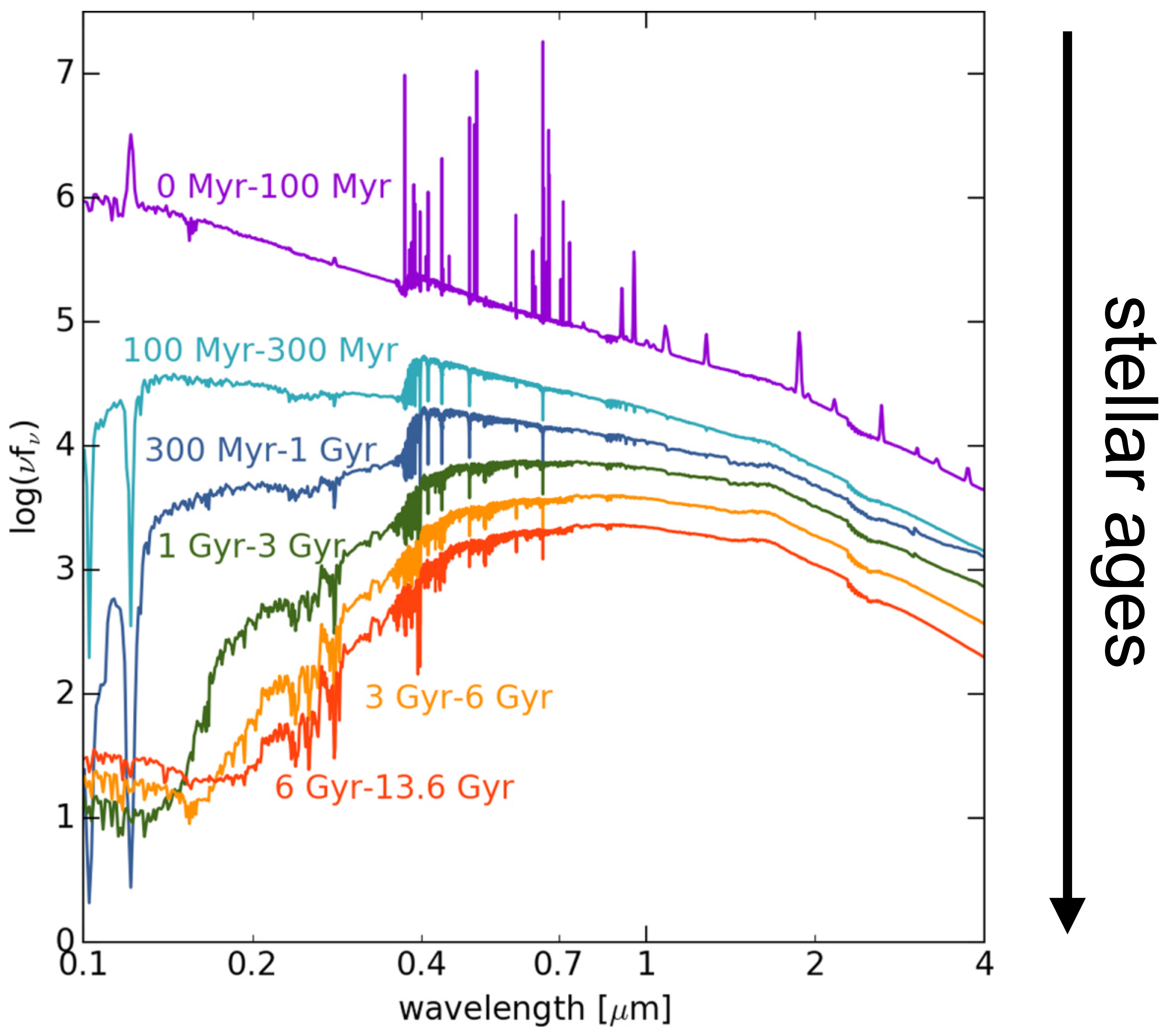
How do we represent a galaxy?

Stellar Population Synthesis (SPS)

- Builds galaxy SED from stars, gas, dust, etc.
- 16 parameters ($\approx$ Prospector-α)
- Emulated with Speculator



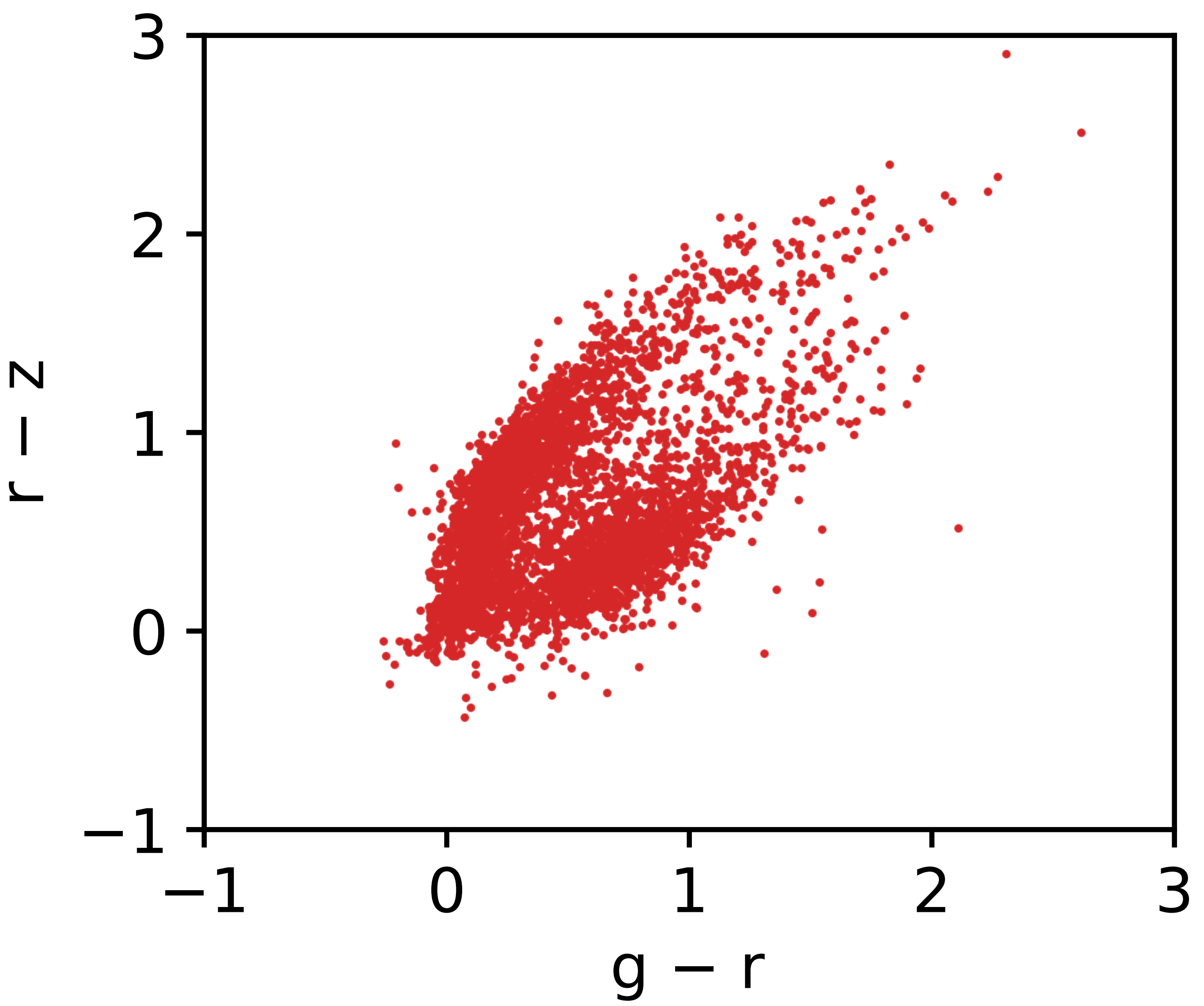
Alsing et al. (2020)



Leja et al. (2017)

What should our population model be?

A score-based diffusion model



target distribution
$(t = 0)$

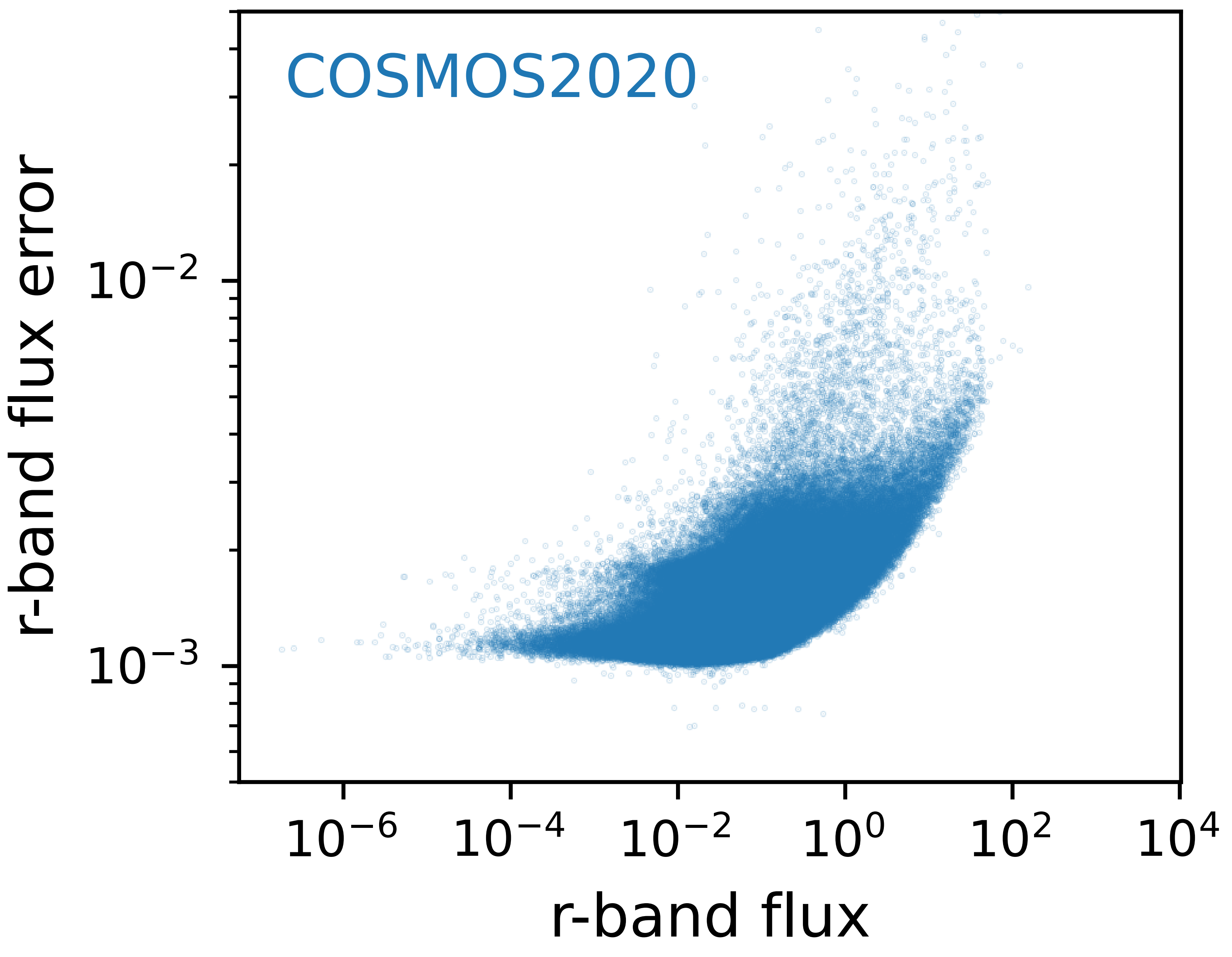
differential equation

iid Gaussian noise
$(t = T)$

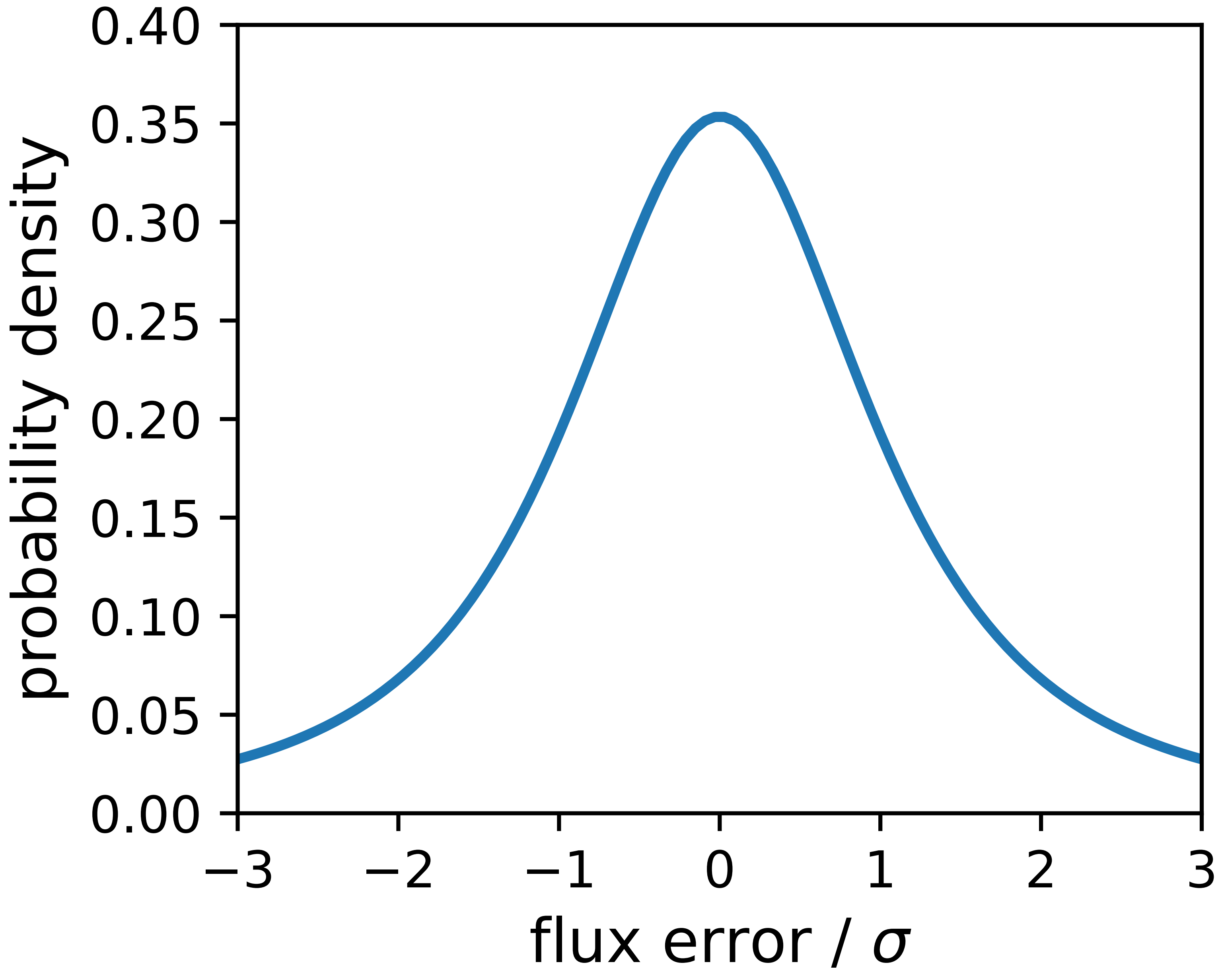
What goes into the data model?

Uncertainty, noise, and calibration

Uncertainty model
(mixture density network)

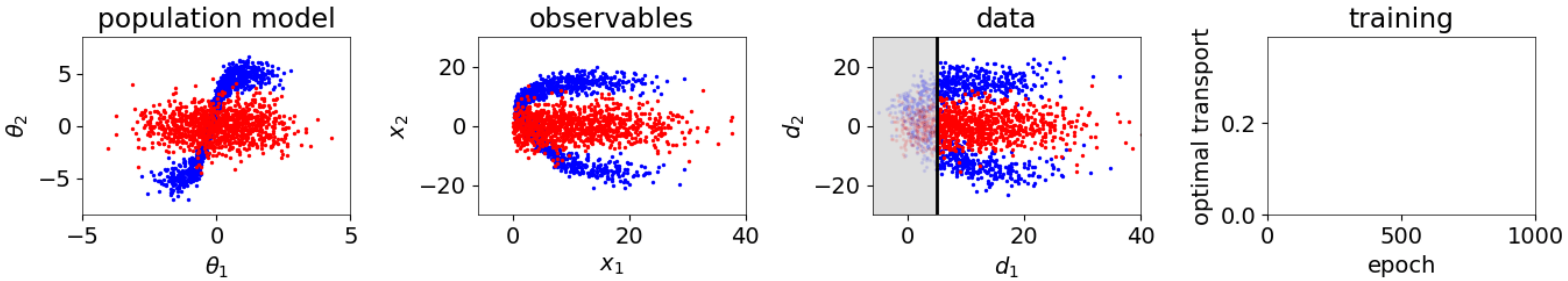


Student's-t error model



How do we fit a model like this?

Optimal Transport (a.k.a. the Wasserstein Distance)

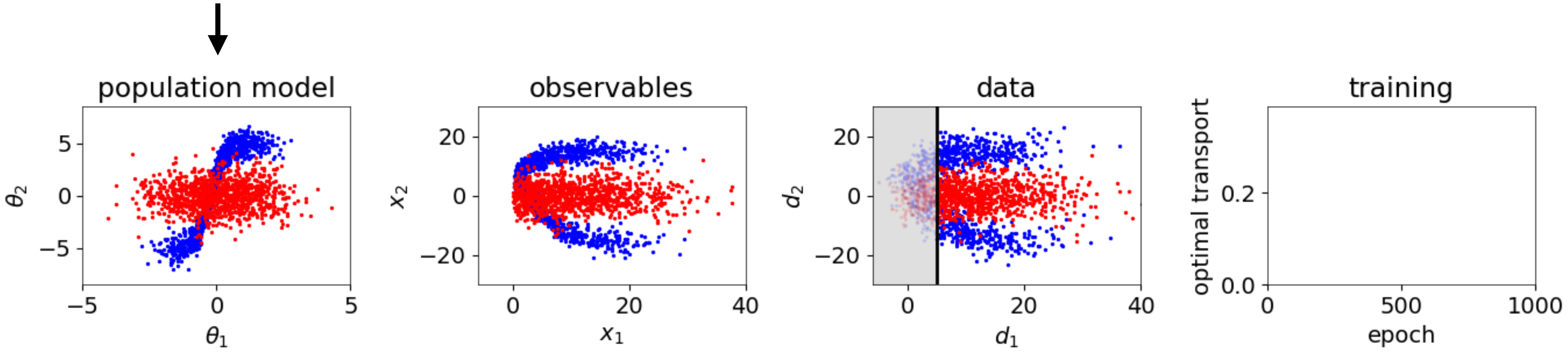


Credit: Justin Alsing

How do we fit a model like this?

Optimal Transport (a.k.a. the Wasserstein Distance)

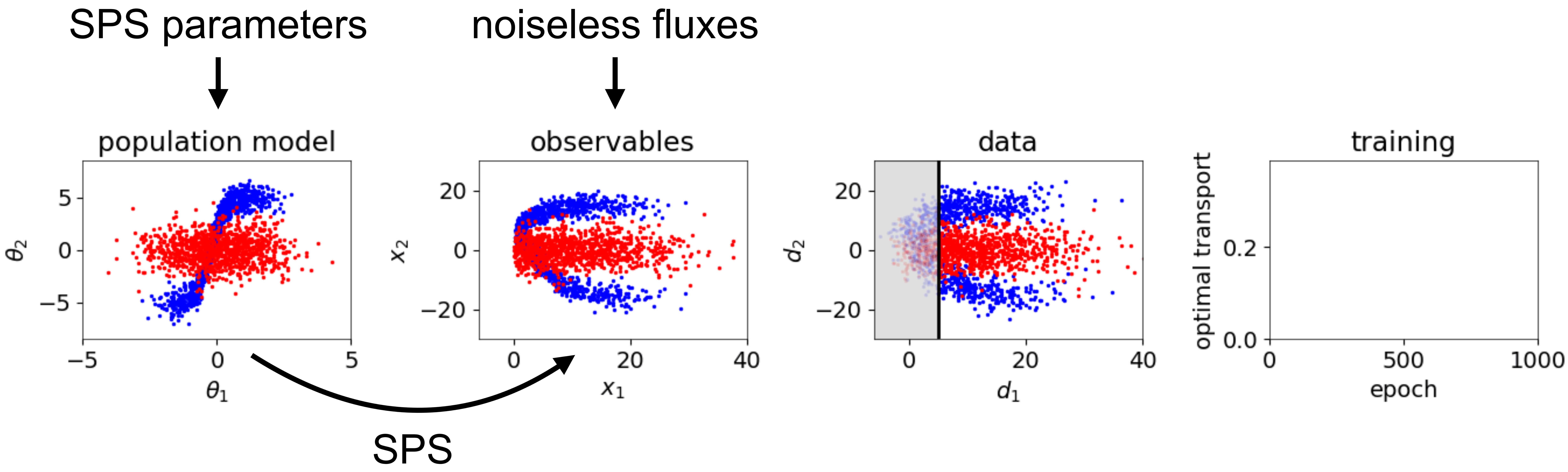
SPS parameters



Credit: Justin Alsing

How do we fit a model like this?

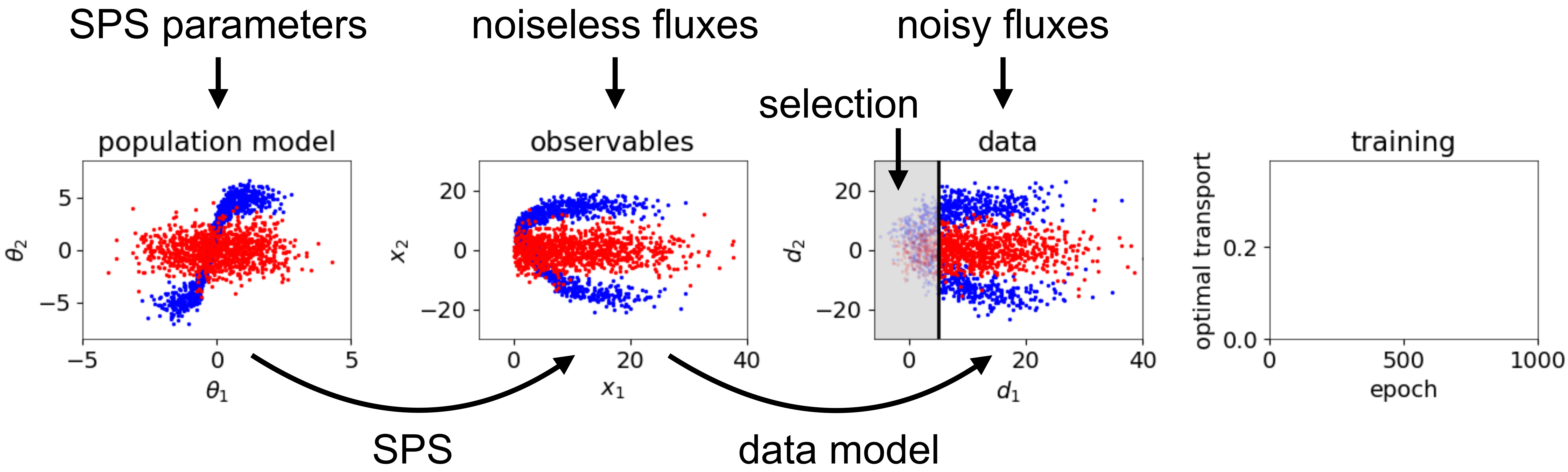
Optimal Transport (a.k.a. the Wasserstein Distance)



Credit: Justin Alsing

How do we fit a model like this?

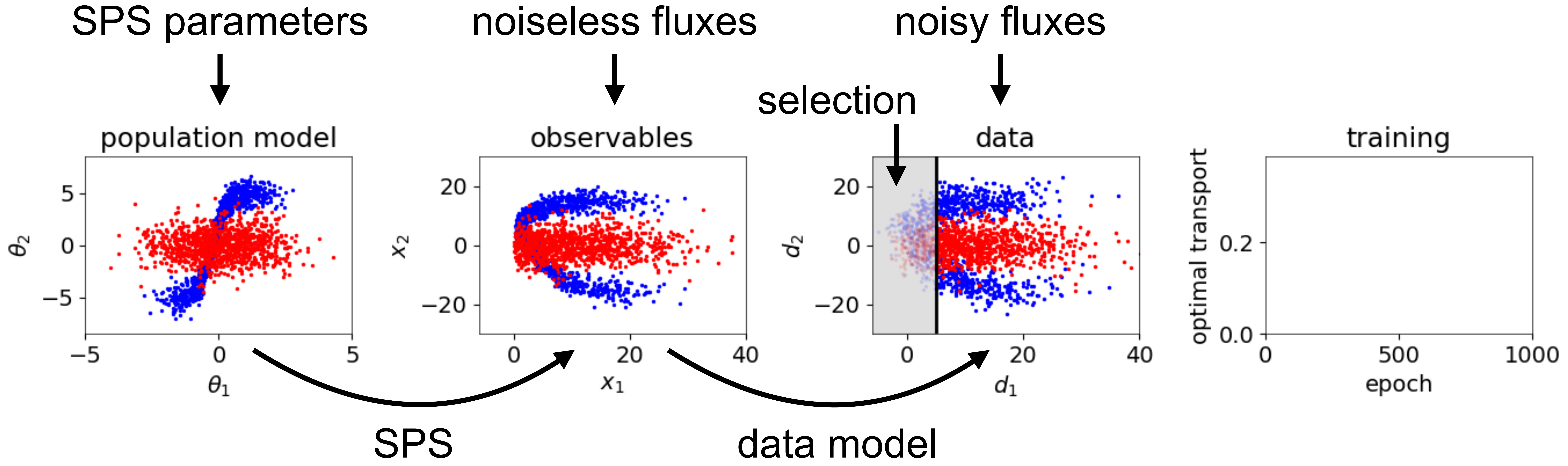
Optimal Transport (a.k.a. the Wasserstein Distance)



Credit: Justin Alsing

How do we fit a model like this?

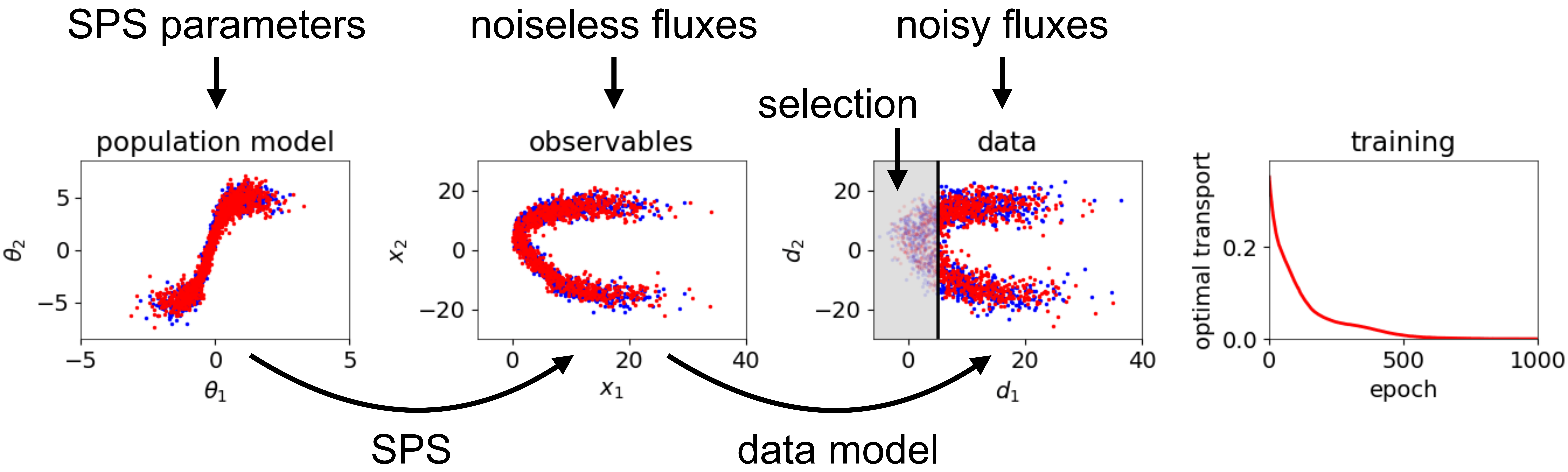
Optimal Transport (a.k.a. the Wasserstein Distance)



Credit: Justin Alsing

How do we fit a model like this?

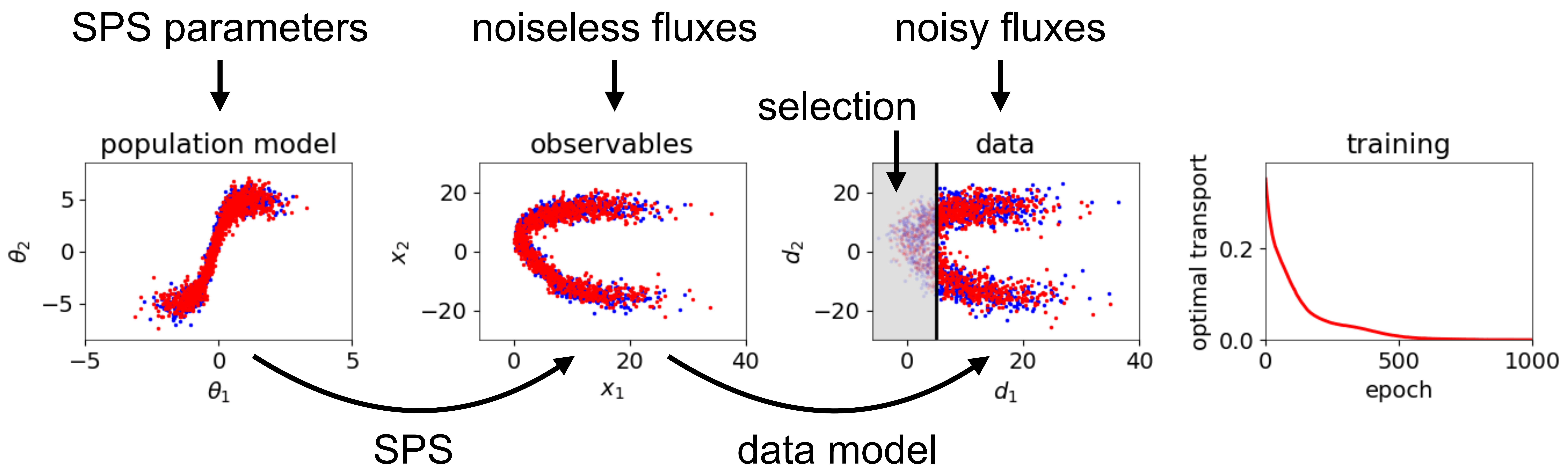
Optimal Transport (a.k.a. the Wasserstein Distance)



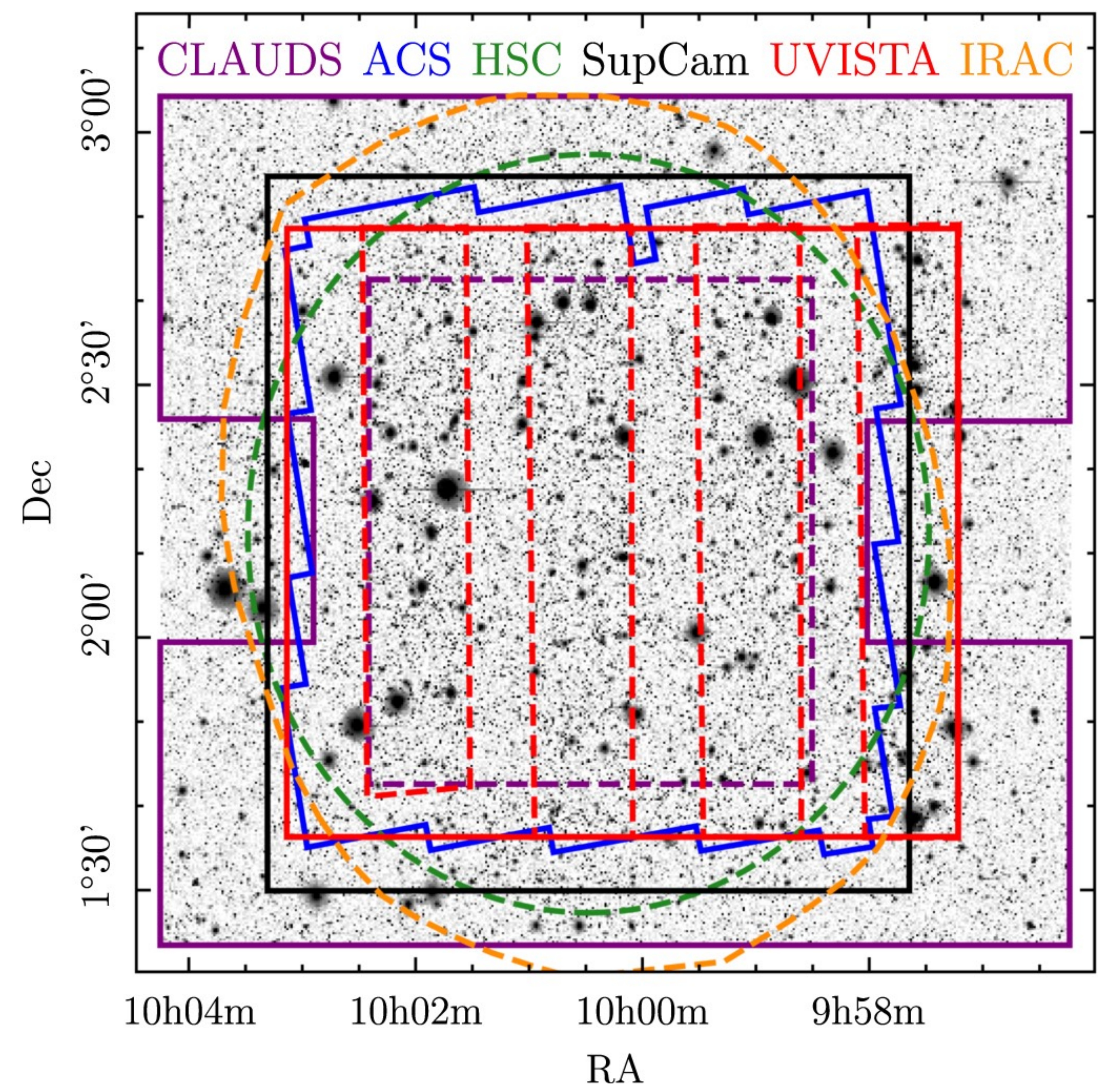
Credit: Justin Alsing

How do we fit a model like this?

Optimal Transport (a.k.a. the Wasserstein Distance)



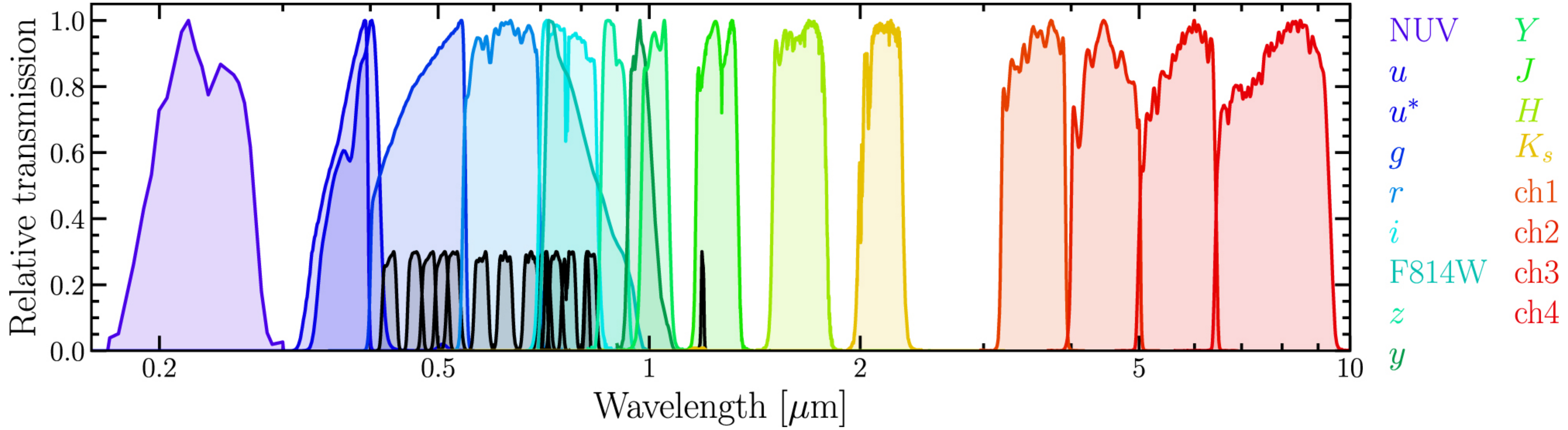
Credit: Justin Alsing



What data will we use?

COSMOS2020 (Weaver et al. 2022)

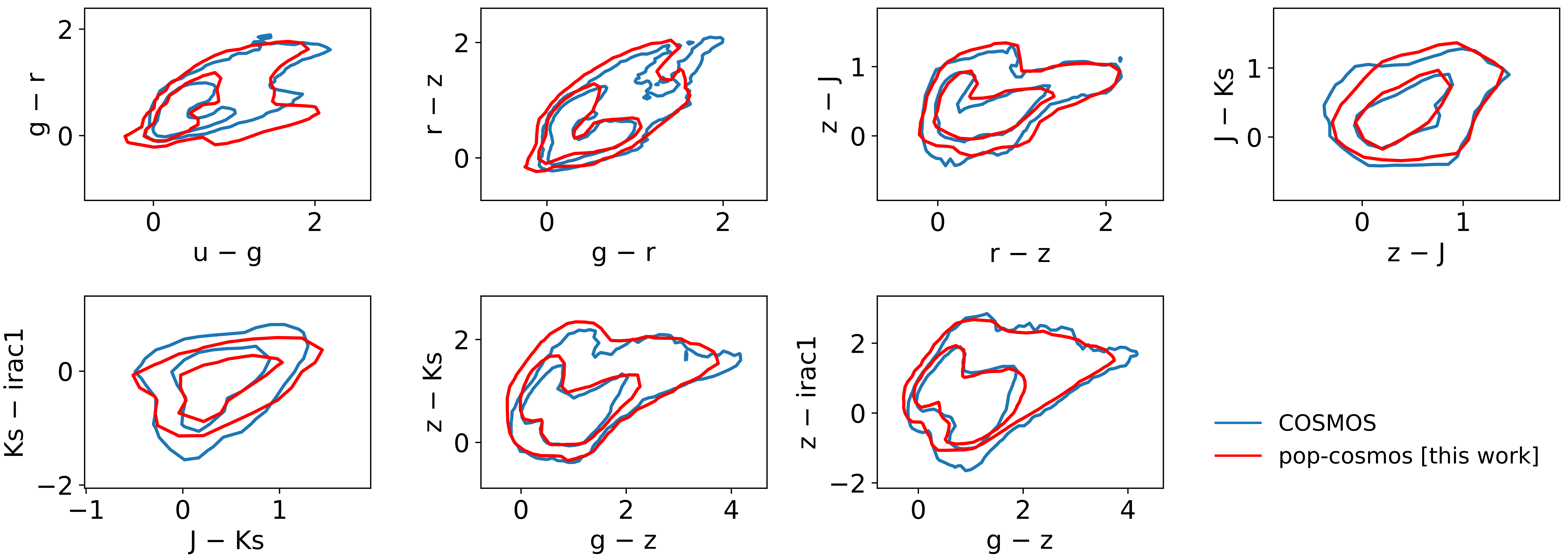
- Has ≈ 140,000 galaxies with $r < 25$
- Wide and narrow bands
- Coverage from near-UV to mid-IR



Results

Photometric Predictions

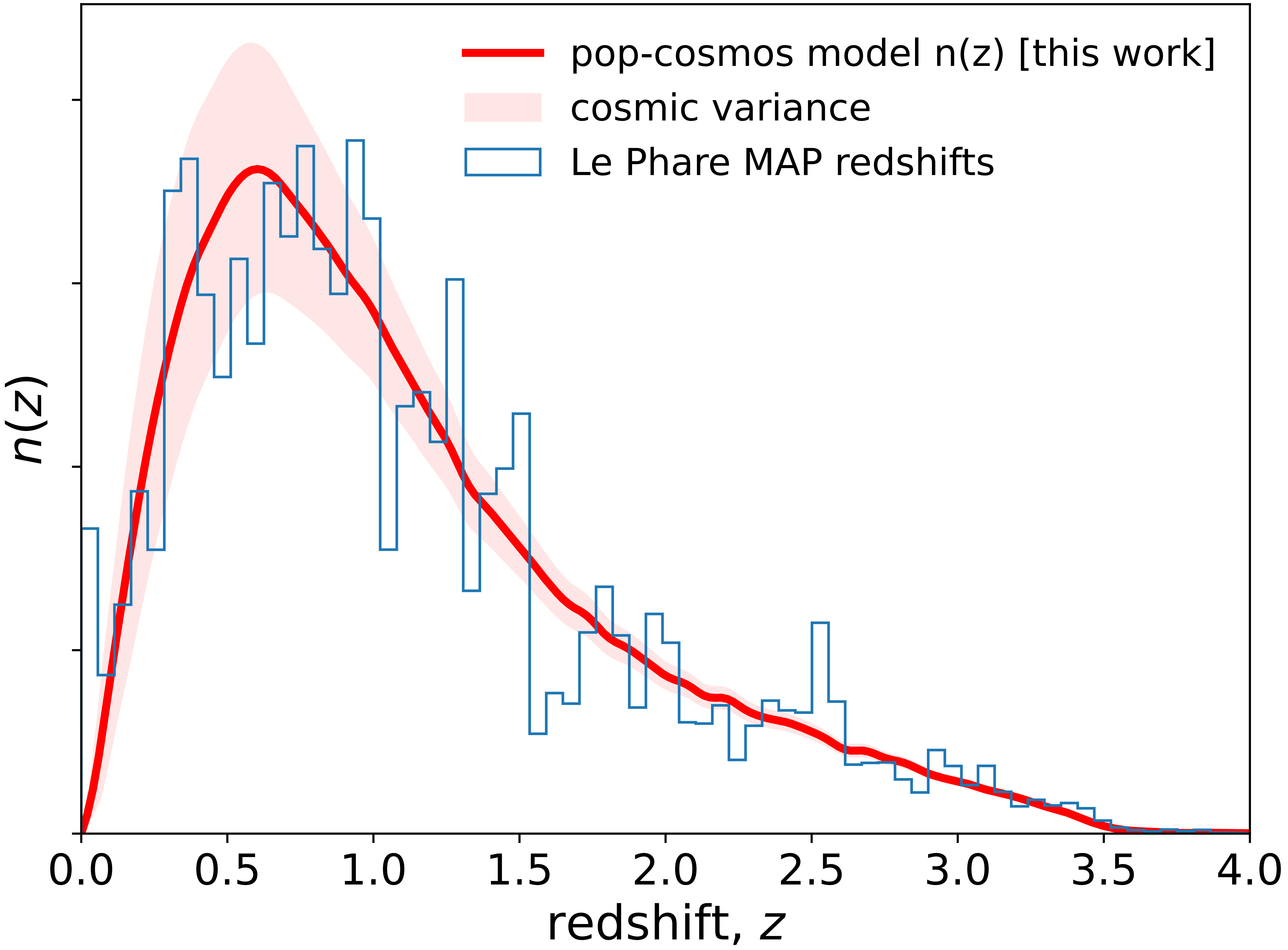
Colour-colour plots



Alsing et al. (2024)

Redshift Distribution

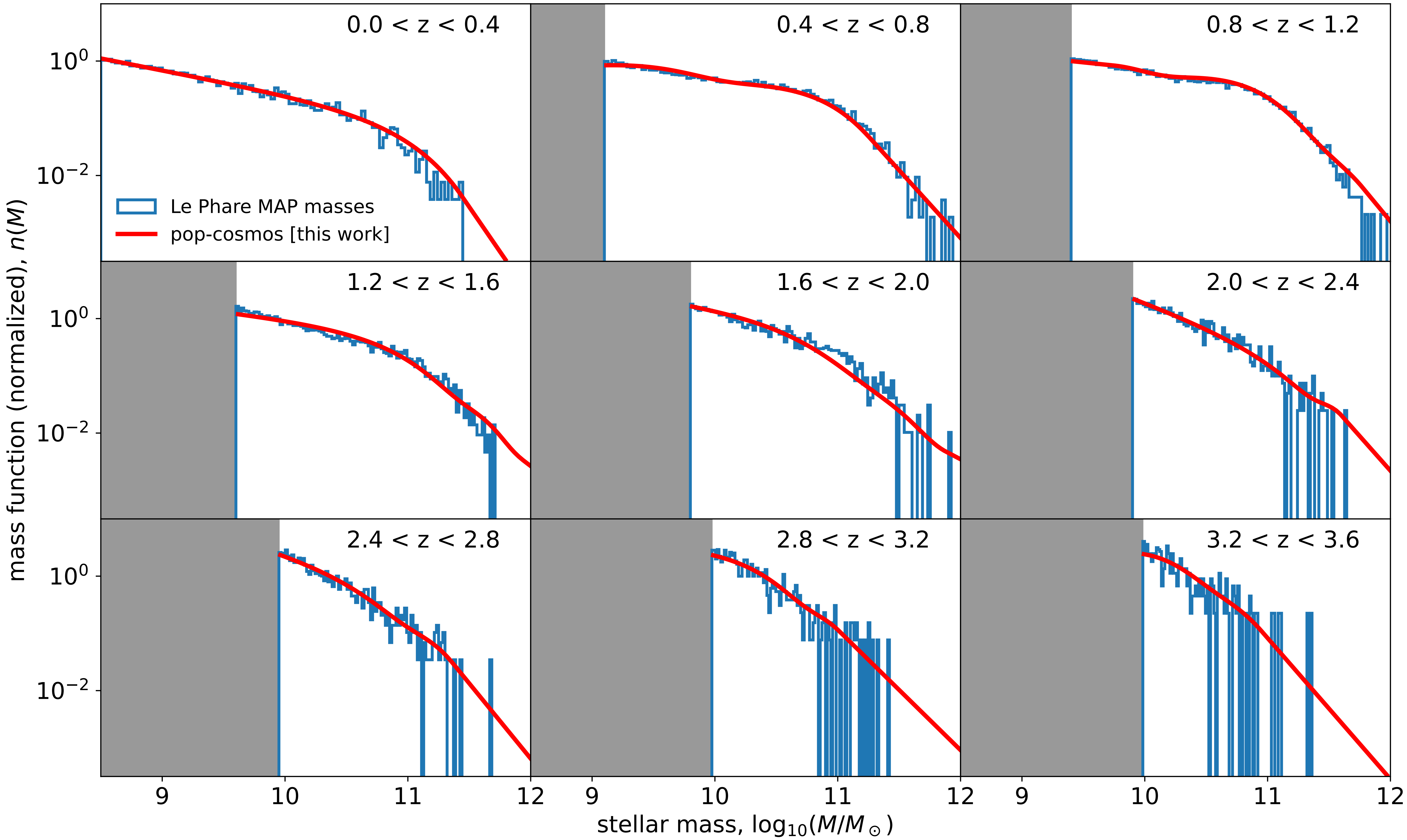
Predicted population-level $n(z)$



Alsing et al. (2024)

Stellar Mass Function

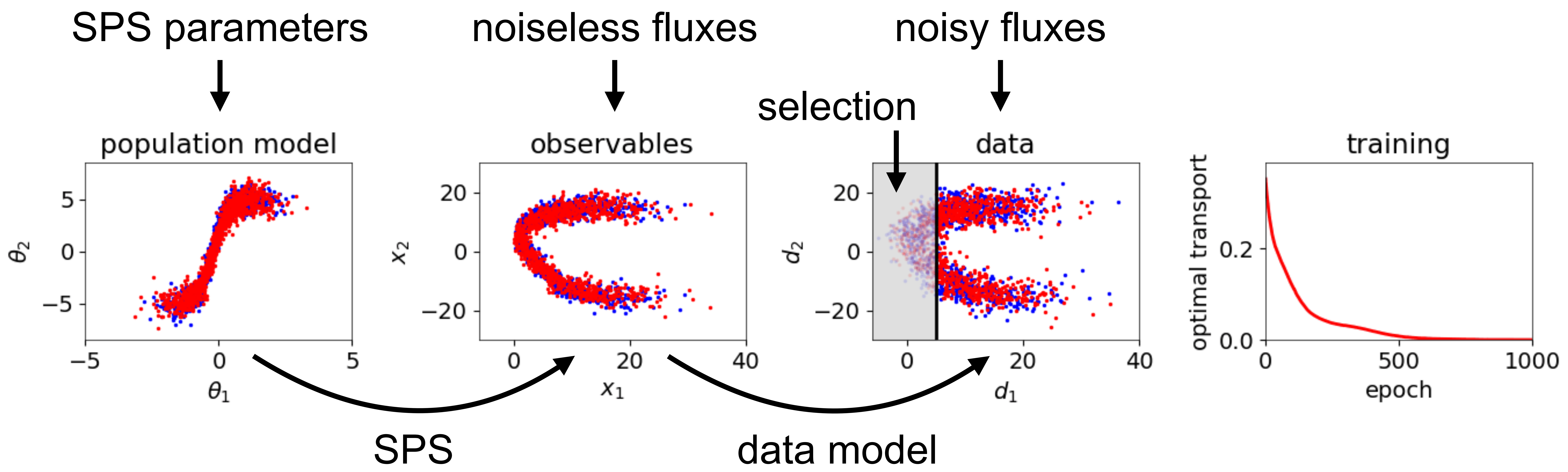
We get more than just redshifts!



Alsing et al. (2024)

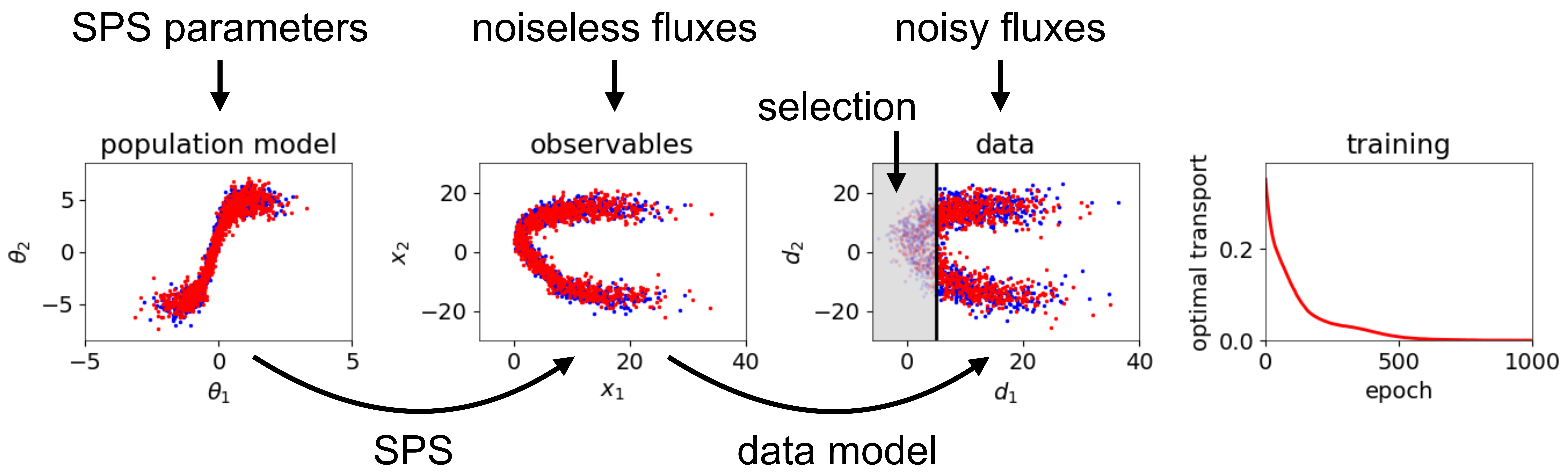
What's else can we do?

Making predictions for other surveys…



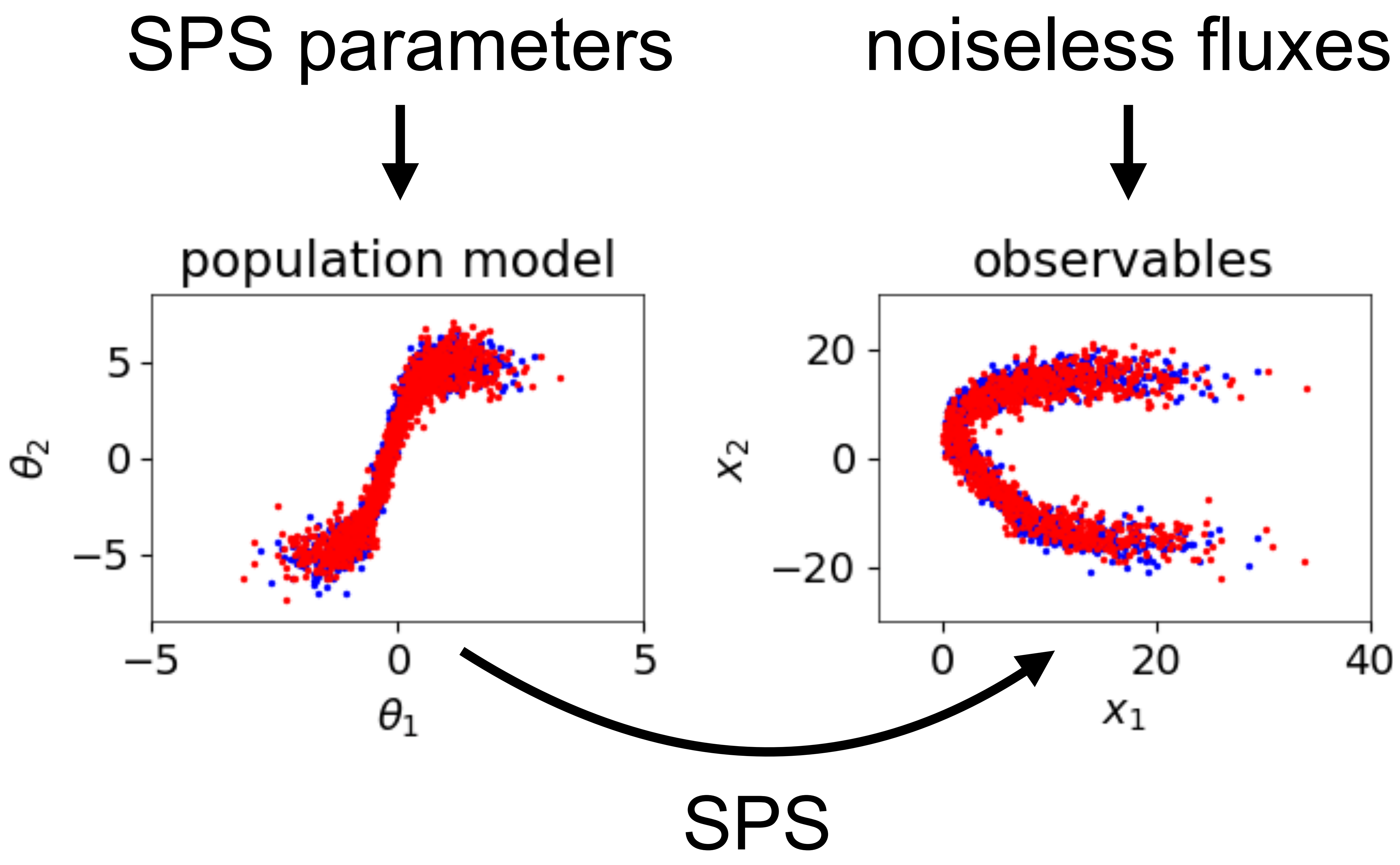
What’s else can we do?

Making predictions for other surveys…



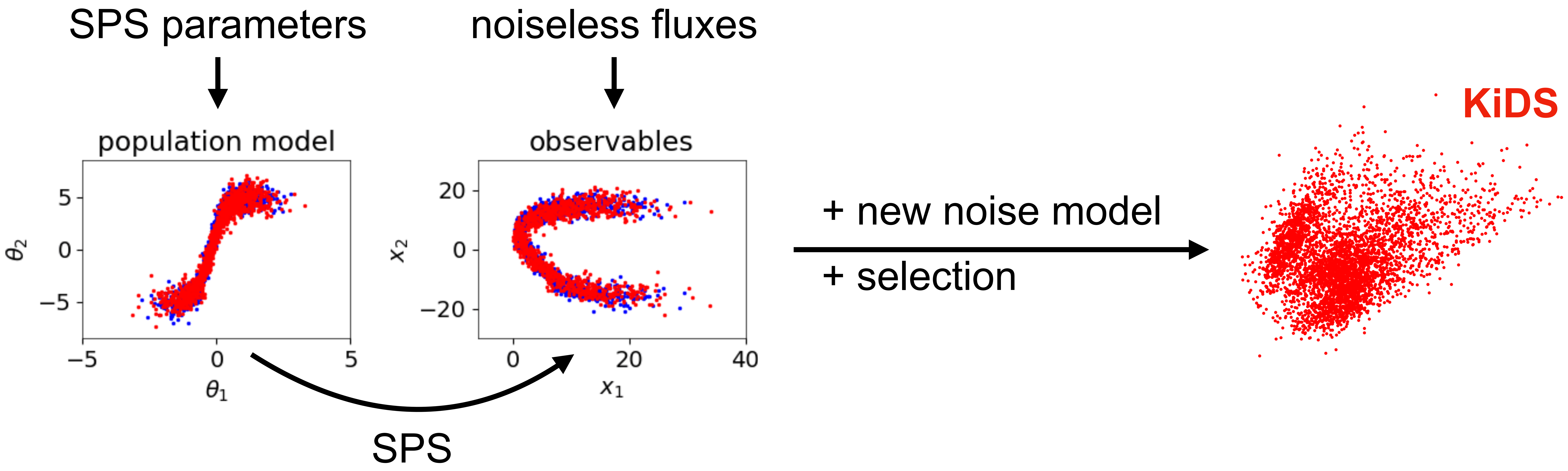
What’s else can we do?

Making predictions for other surveys…



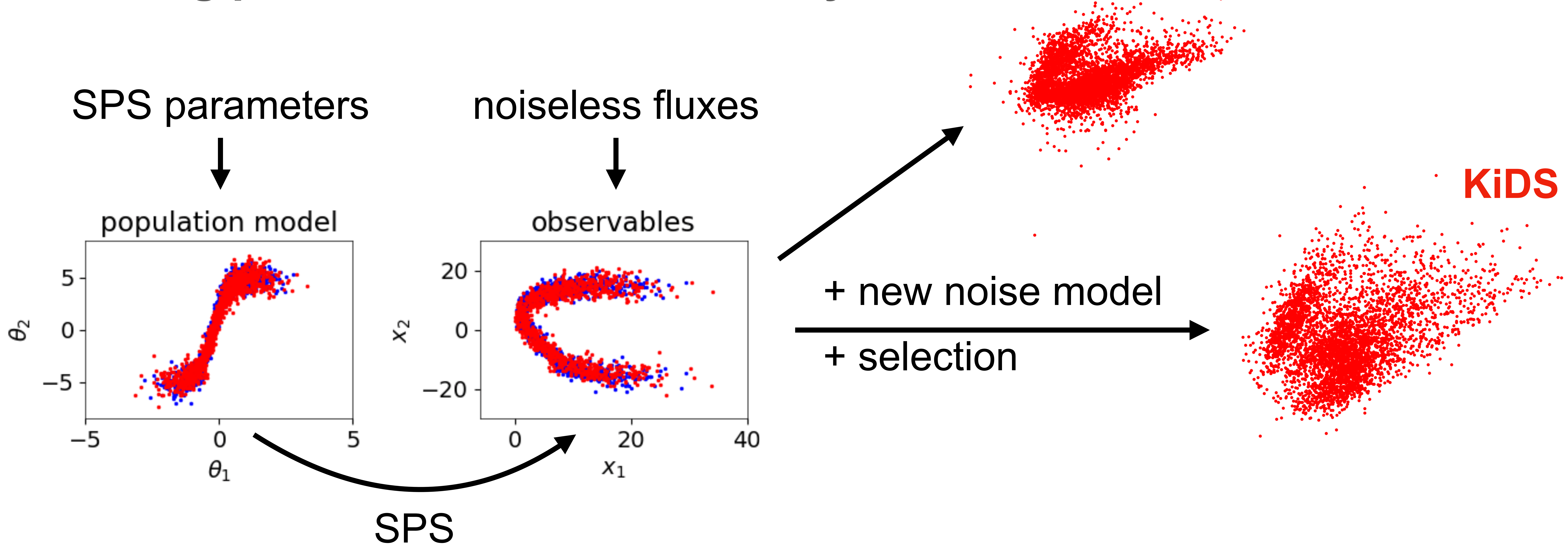
What’s else can we do?

Making predictions for other surveys…



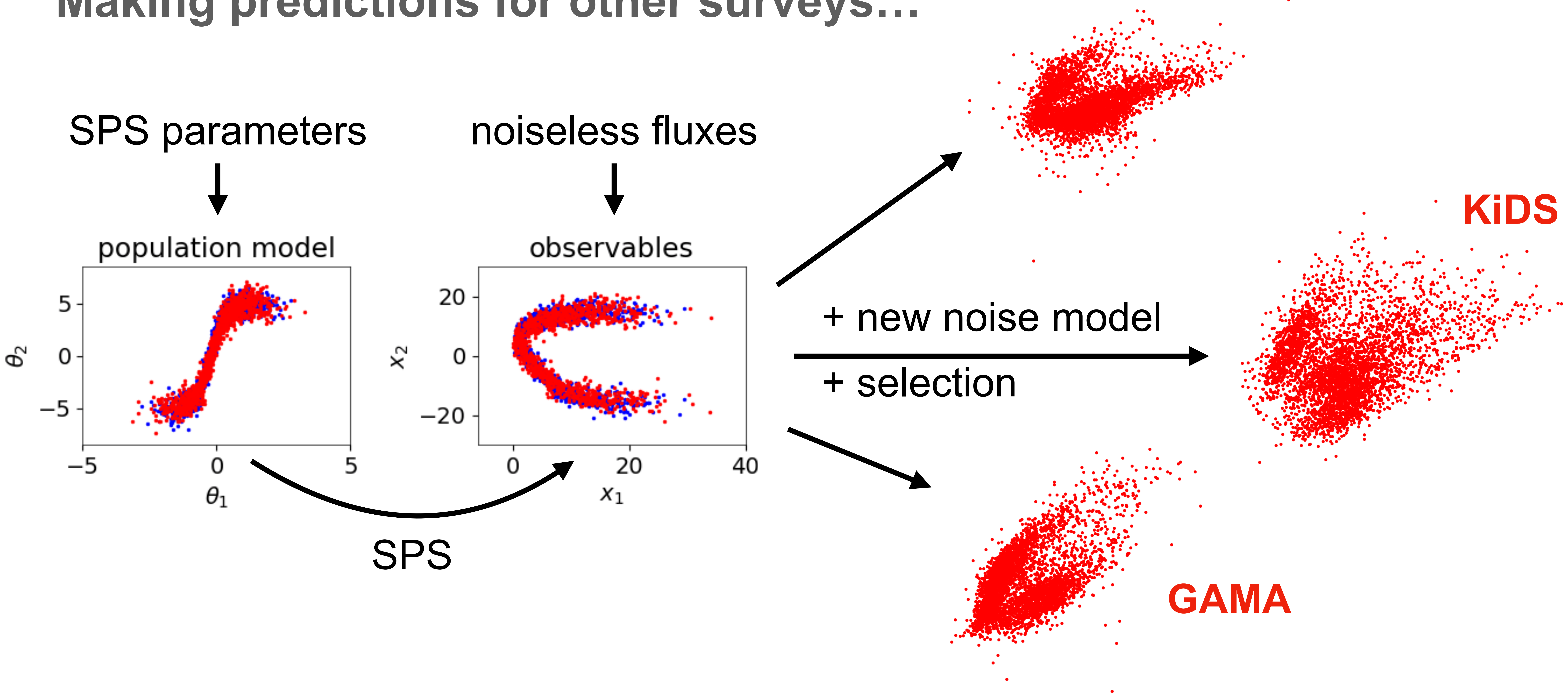
What’s else can we do?

Making predictions for other surveys…



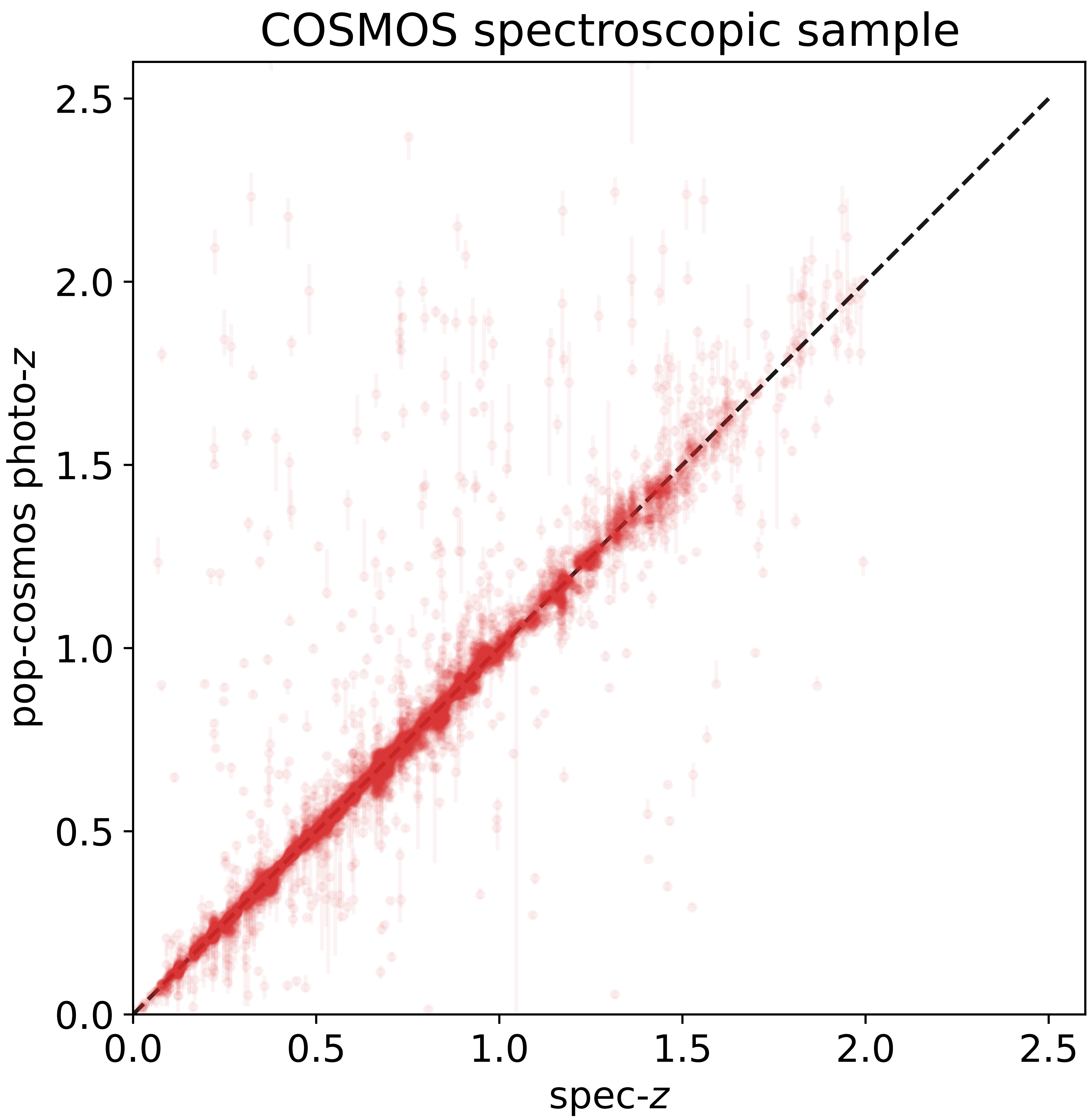
What’s else can we do?

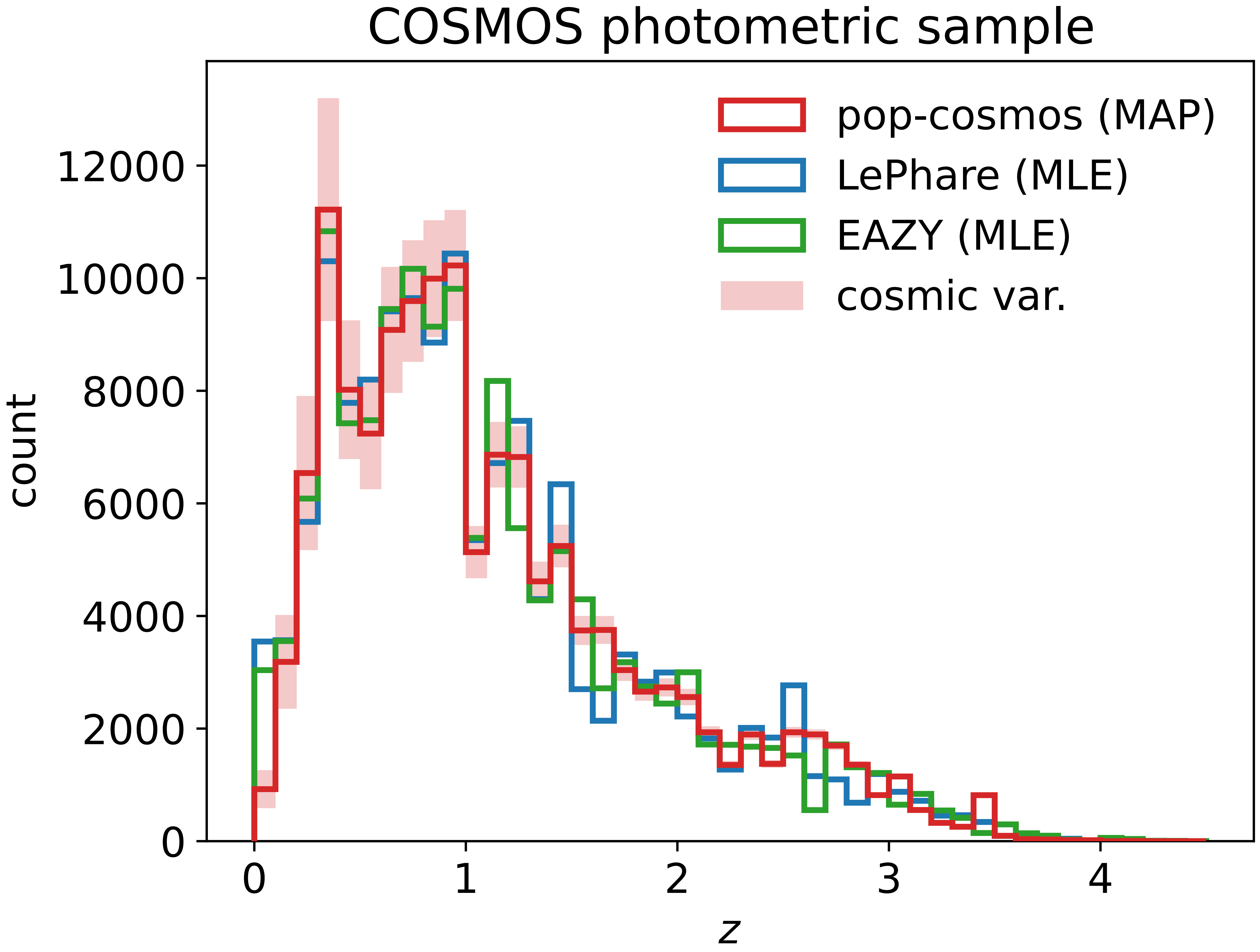
Making predictions for other surveys…



What's else can we do?

Using pop-cosmos as a prior for Bayesian photo-z estimation…





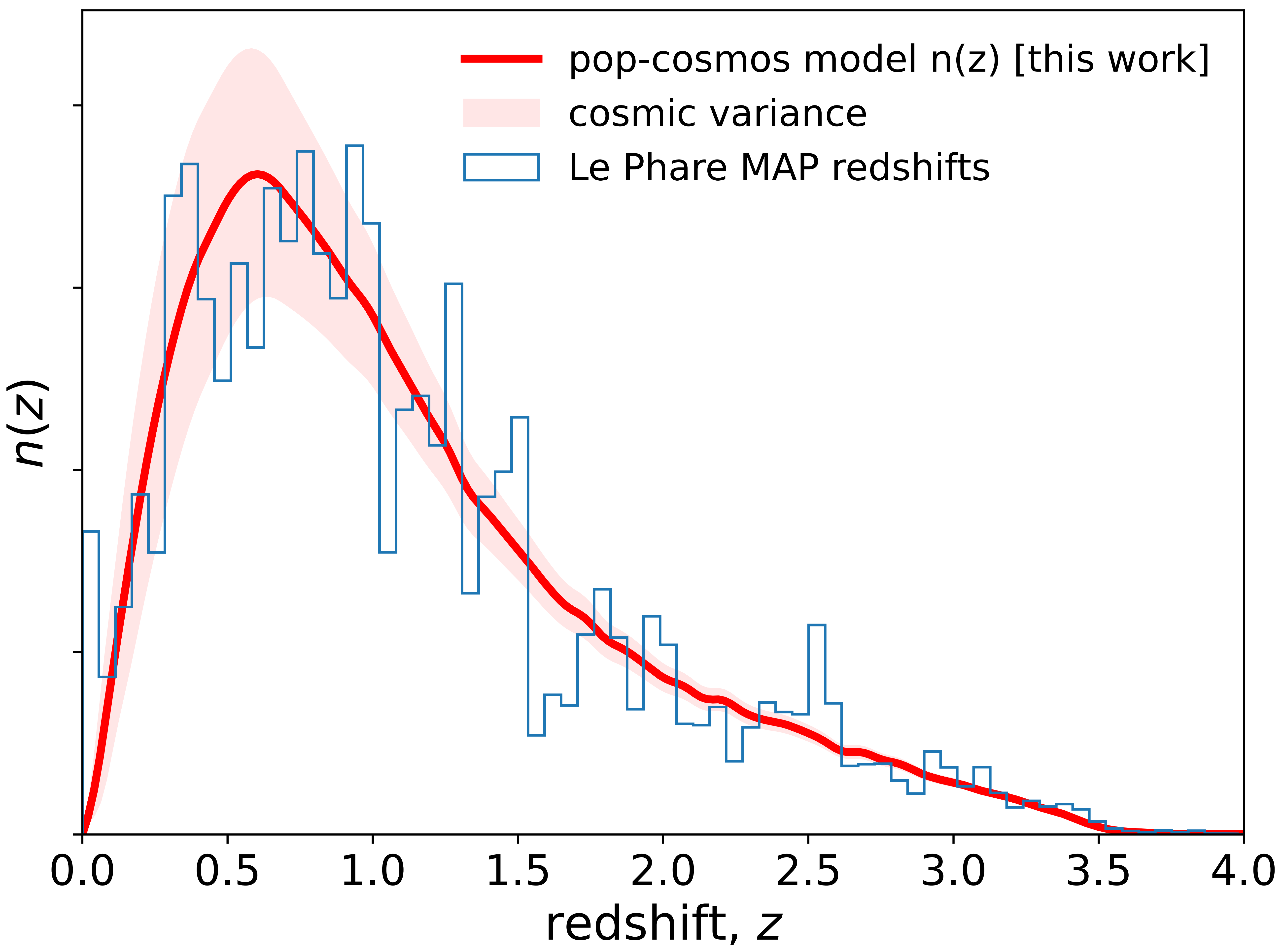
Thorp et al. (in prep.)

Summary

pop-cosmos

2402.00935

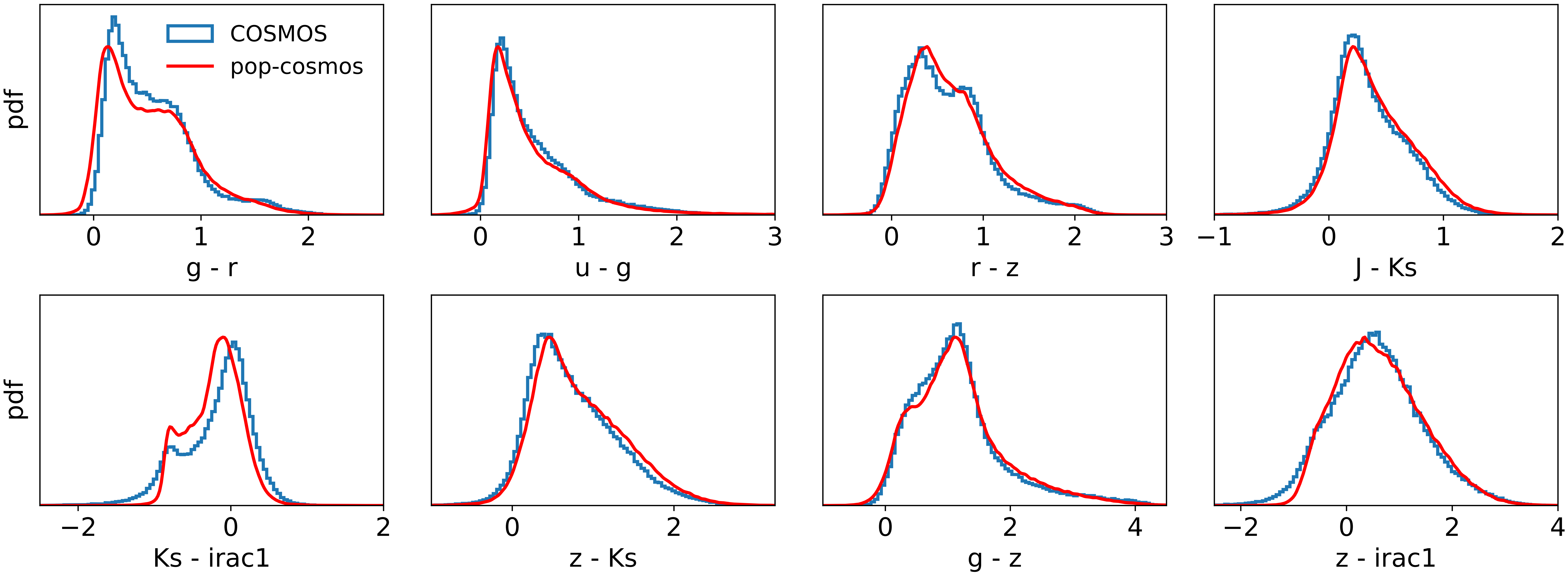
- Comprehensive forward model for galaxy photometry
- Flexible non-parametric population model
- Gives us redshift distribution, and can be used to make predictions for other surveys
- Tons of information about galaxy demographics on a huge sample with minimal selection (**see the talk + poster by Sinan Deger!**)



Extra Slides

Photometric Predictions

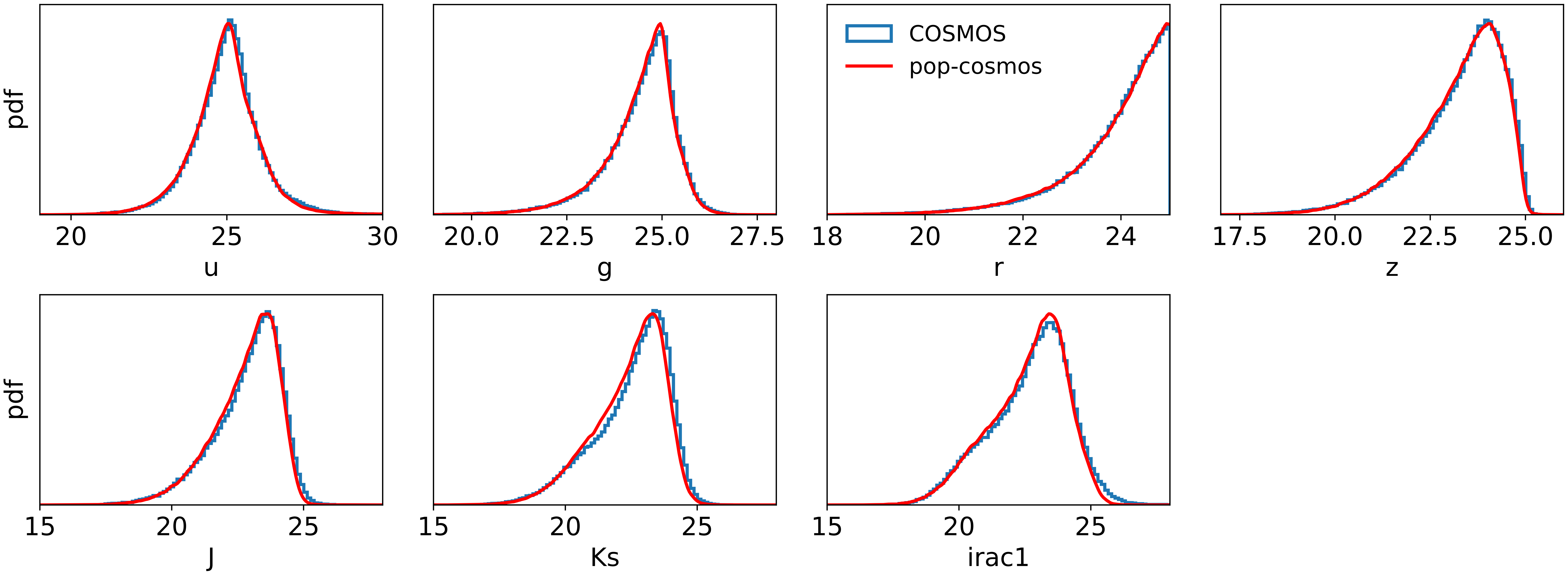
Colour marginals



Alsing et al. (2024)

Photometric Predictions

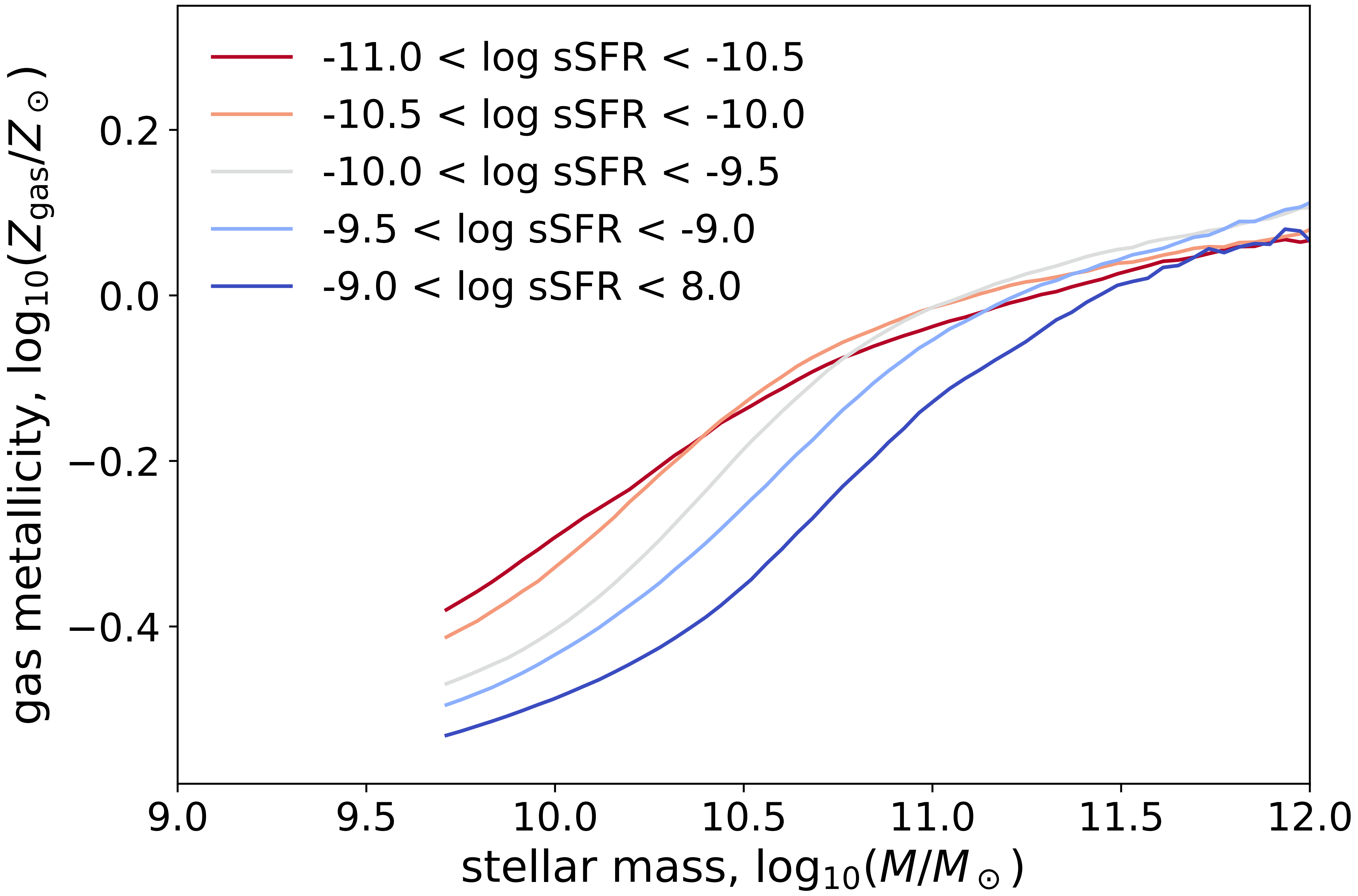
Magnitude marginals



Alsing et al. (2024)

Fundamental Metallicity Relation

Gas metallicity vs. mass vs. SFR

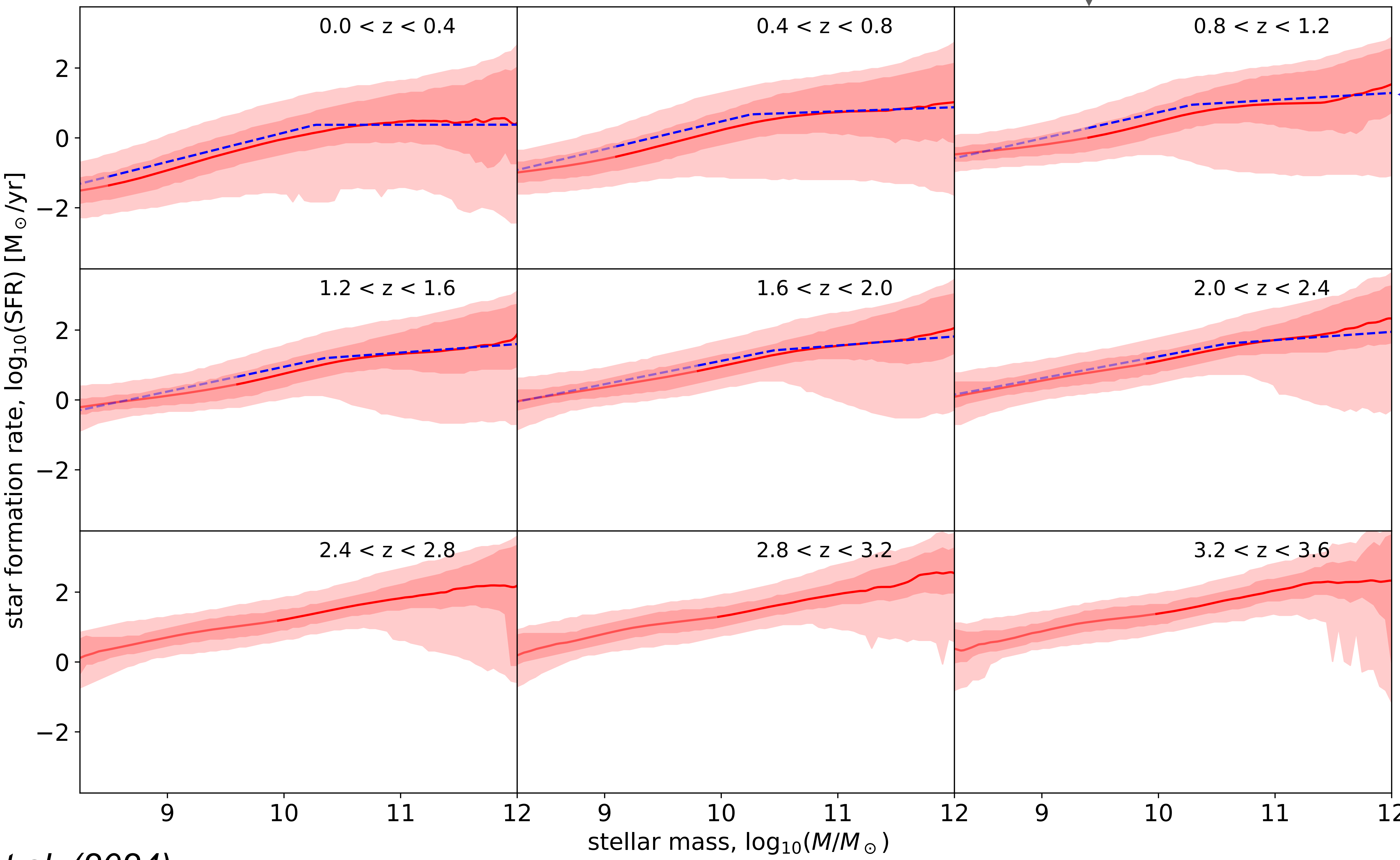


Alsing et al. (2024)

Star Forming Sequence

SFR vs. mass in redshift slices

estimated completeness limit



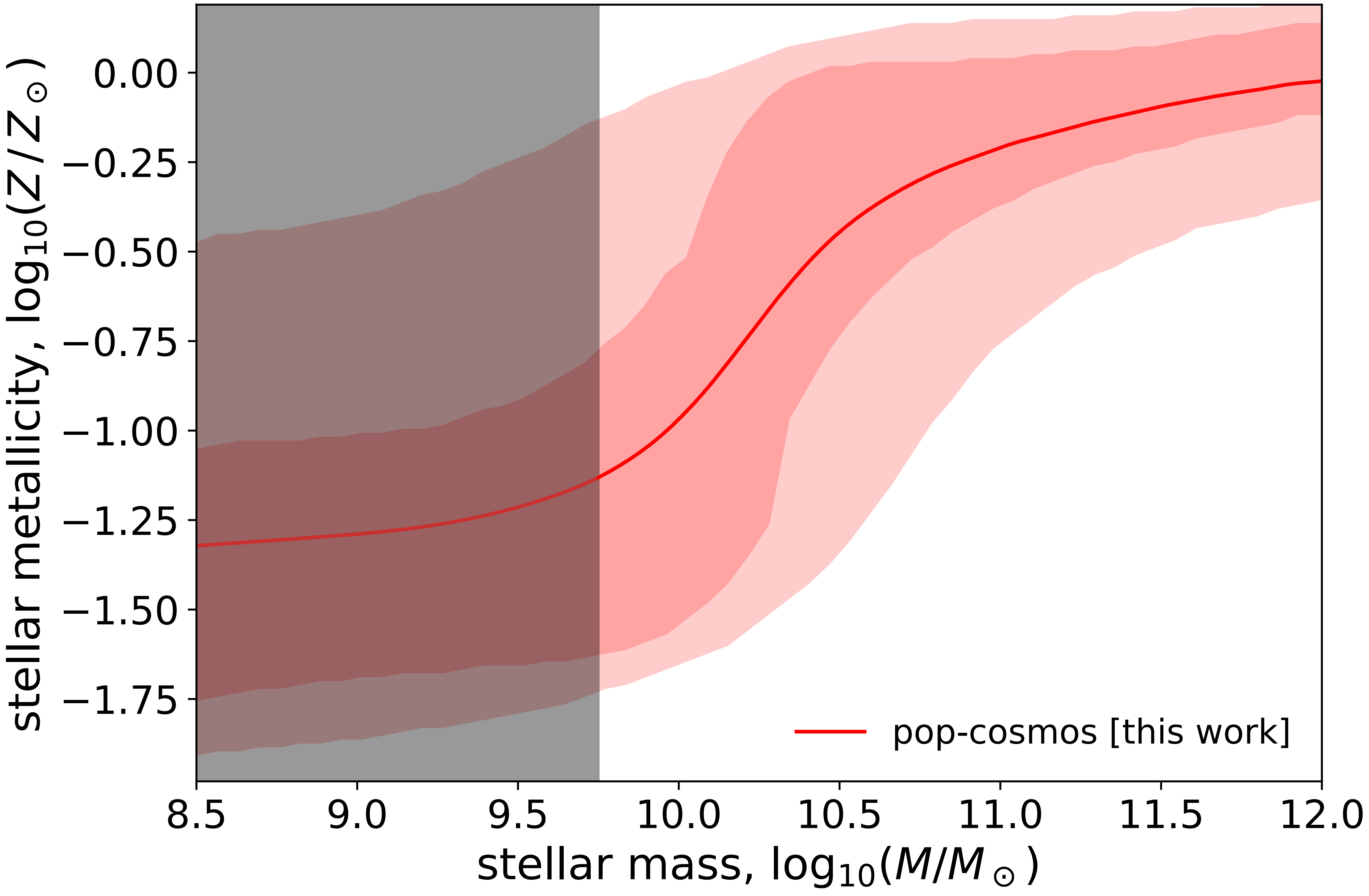
predictions from Leja et al. (2022)

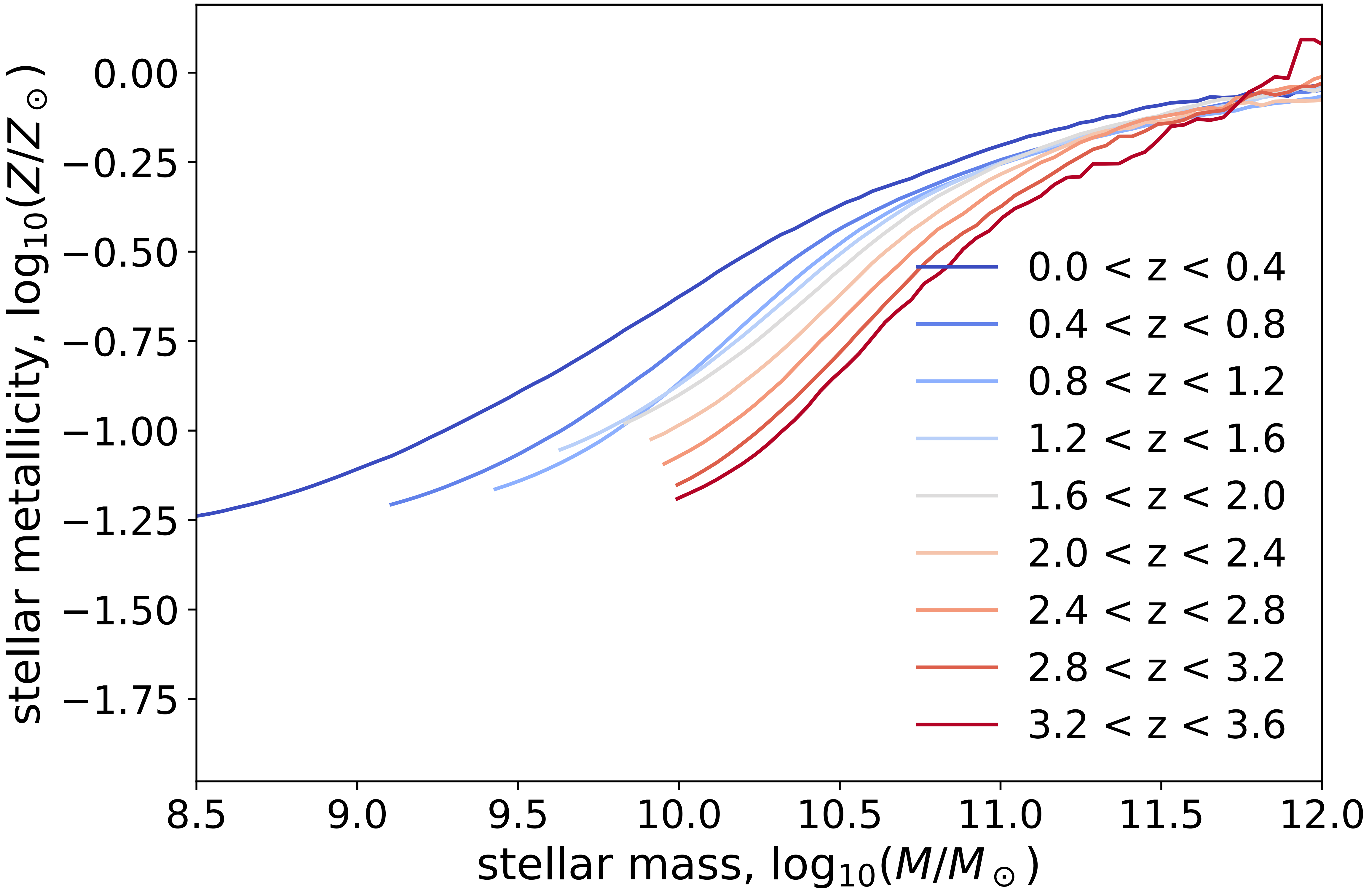
our predictions (distribution of SFRs as a function of mass)

Alsing et al. (2024)

Mass-Metallicity Relation

Stellar metallicity vs. mass

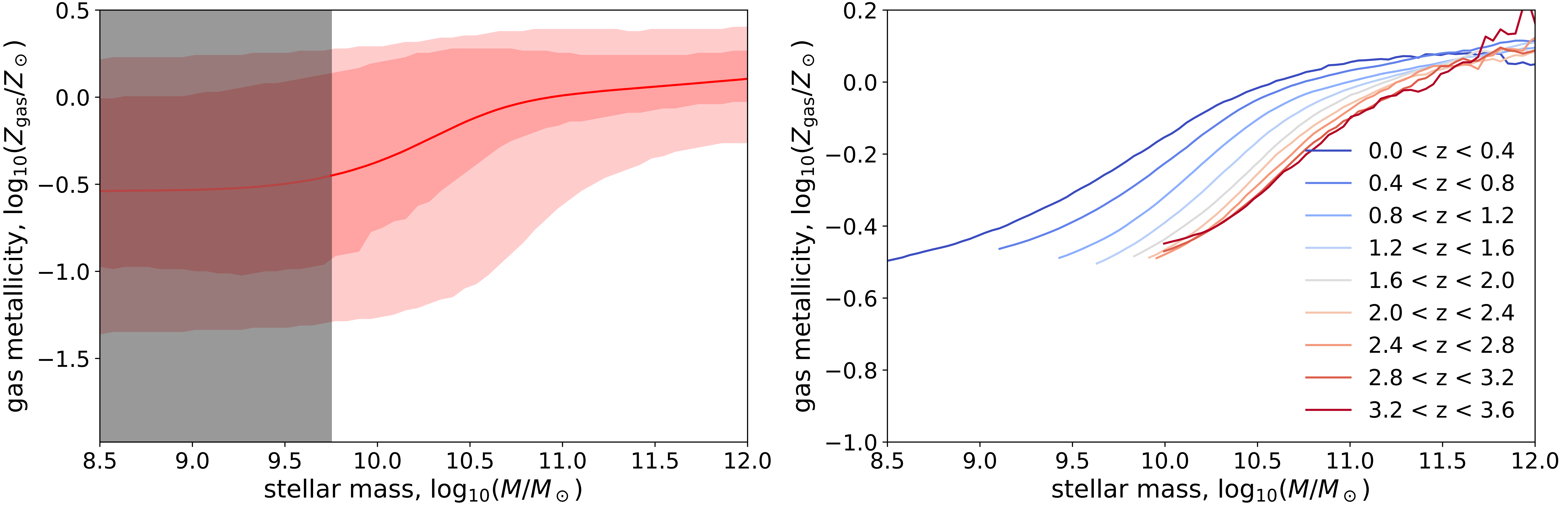




Alsing et al. (2024)

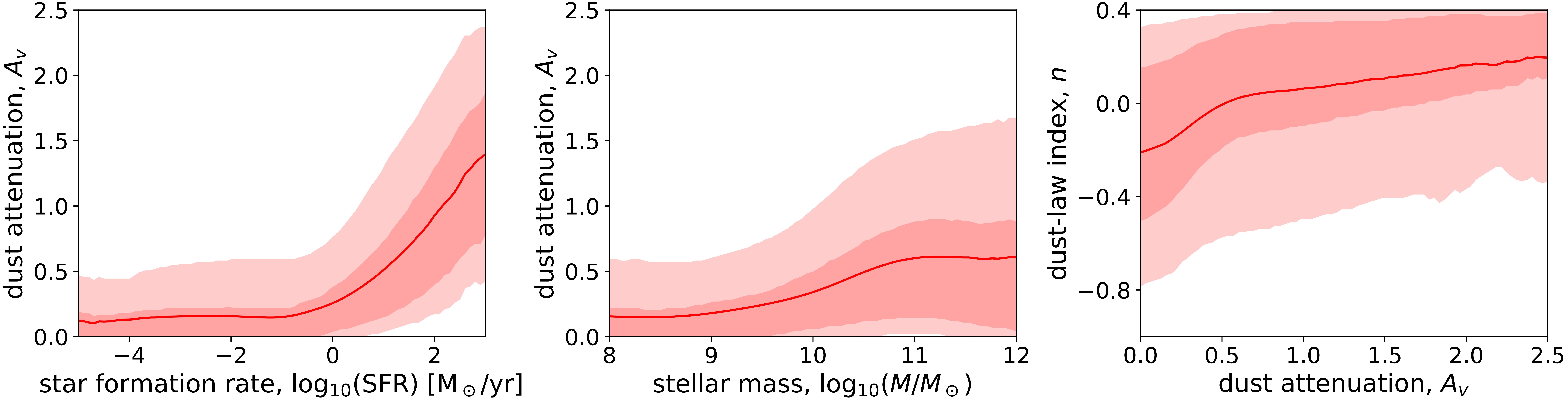
Mass-Metallicity Relation

Gas metallicity vs. mass



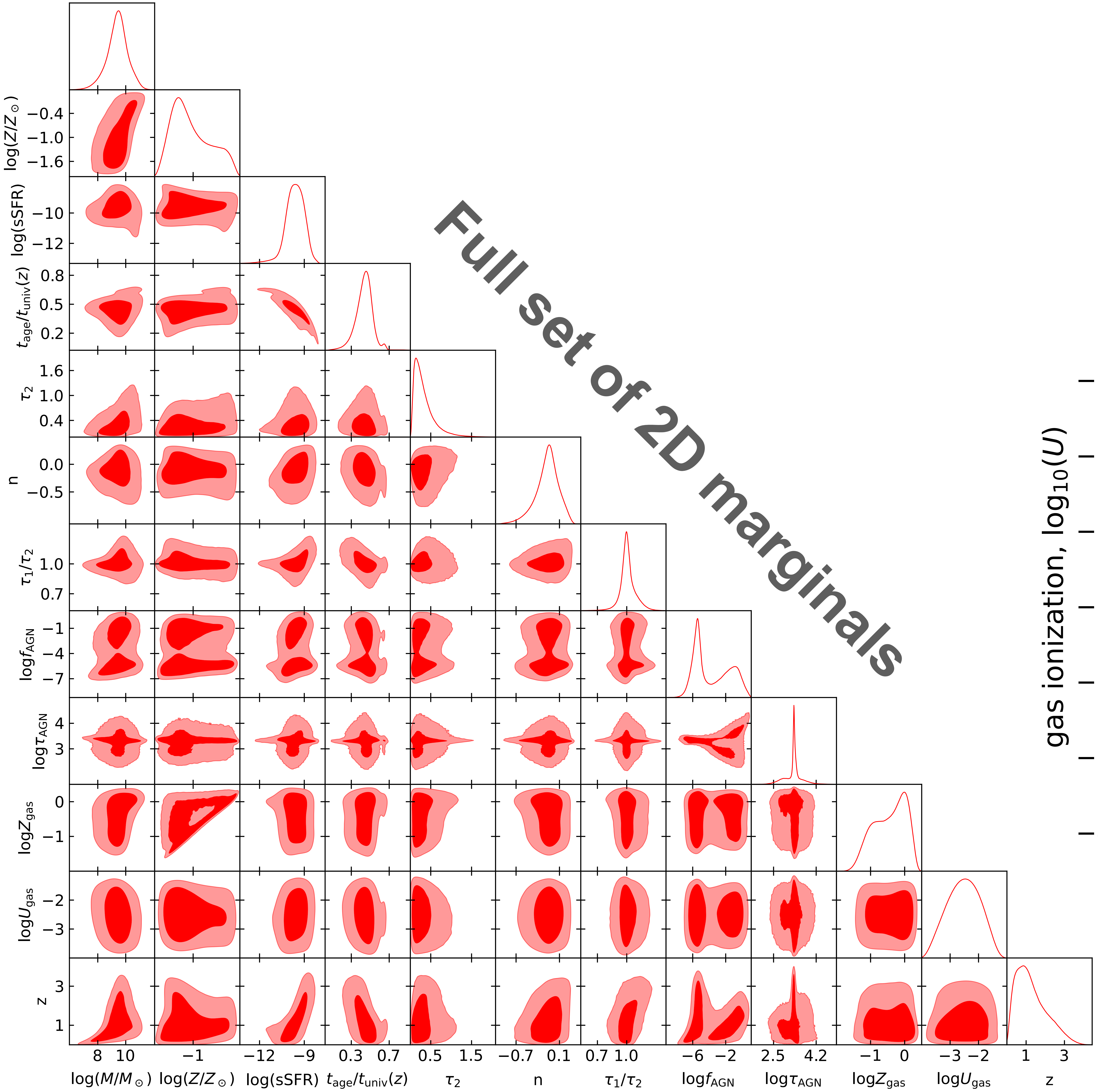
Alsing et al. (2024)

Dust

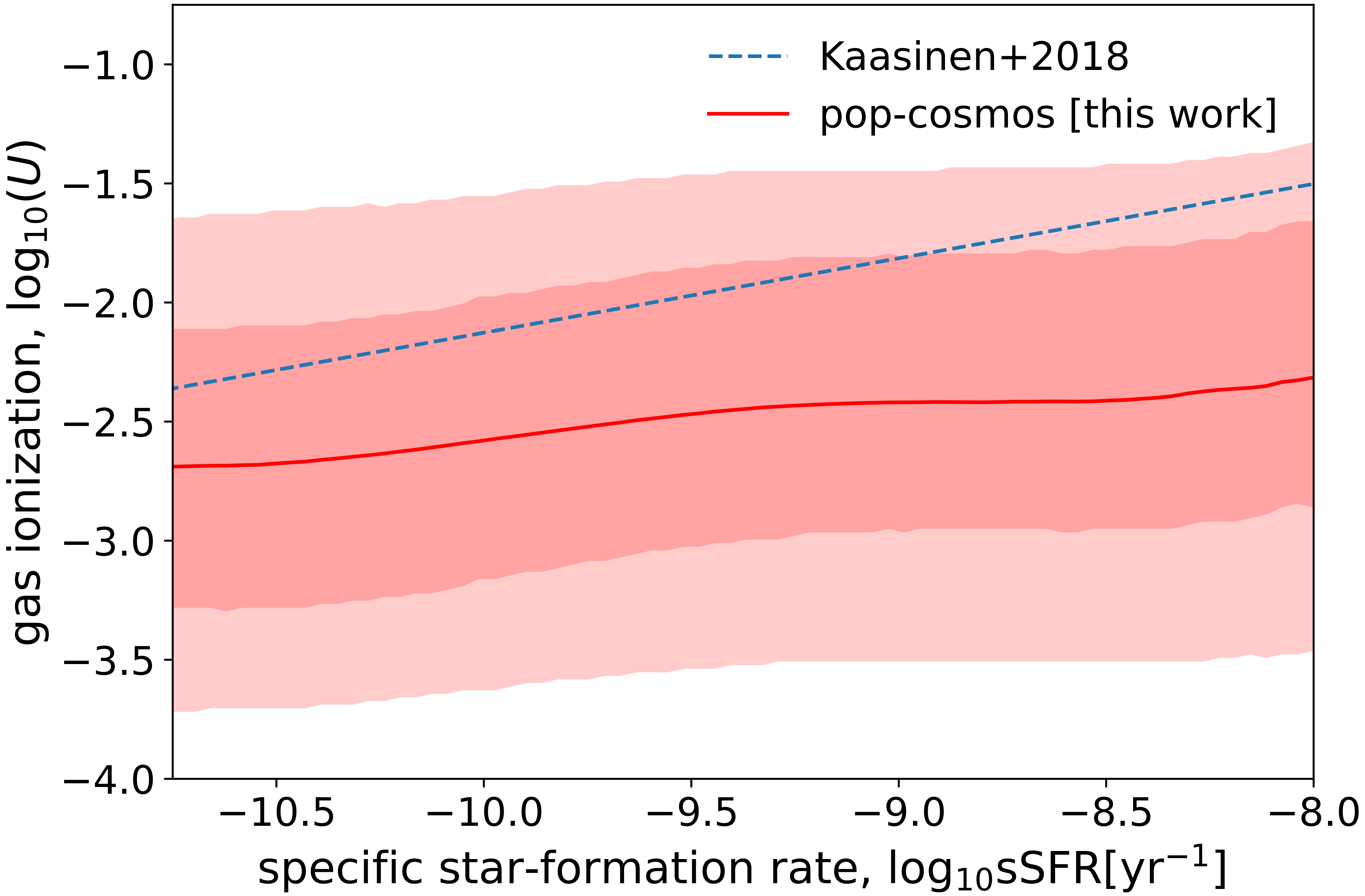


Alsing et al. (2024)

Even More Plots!



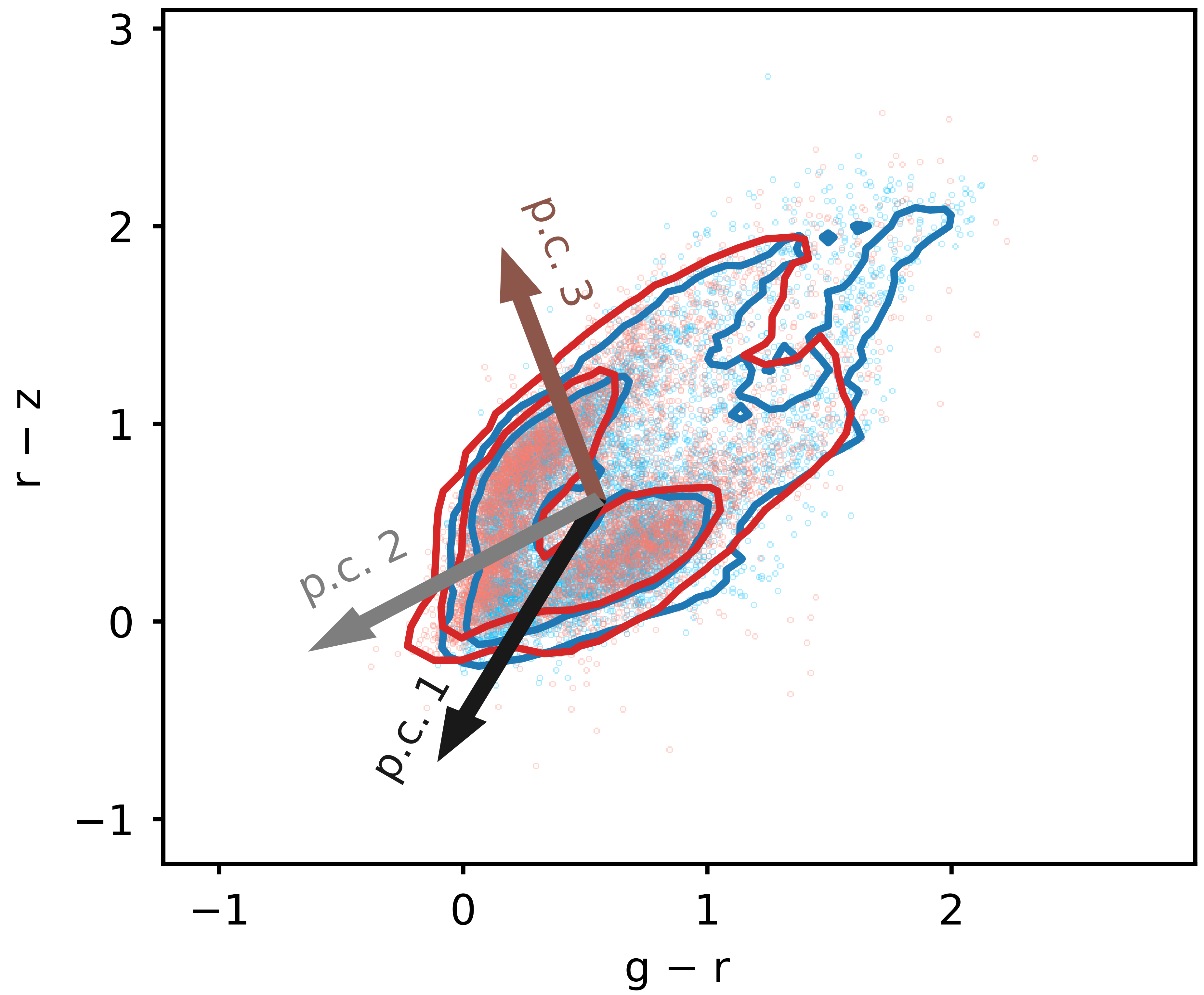
Gas ionisation

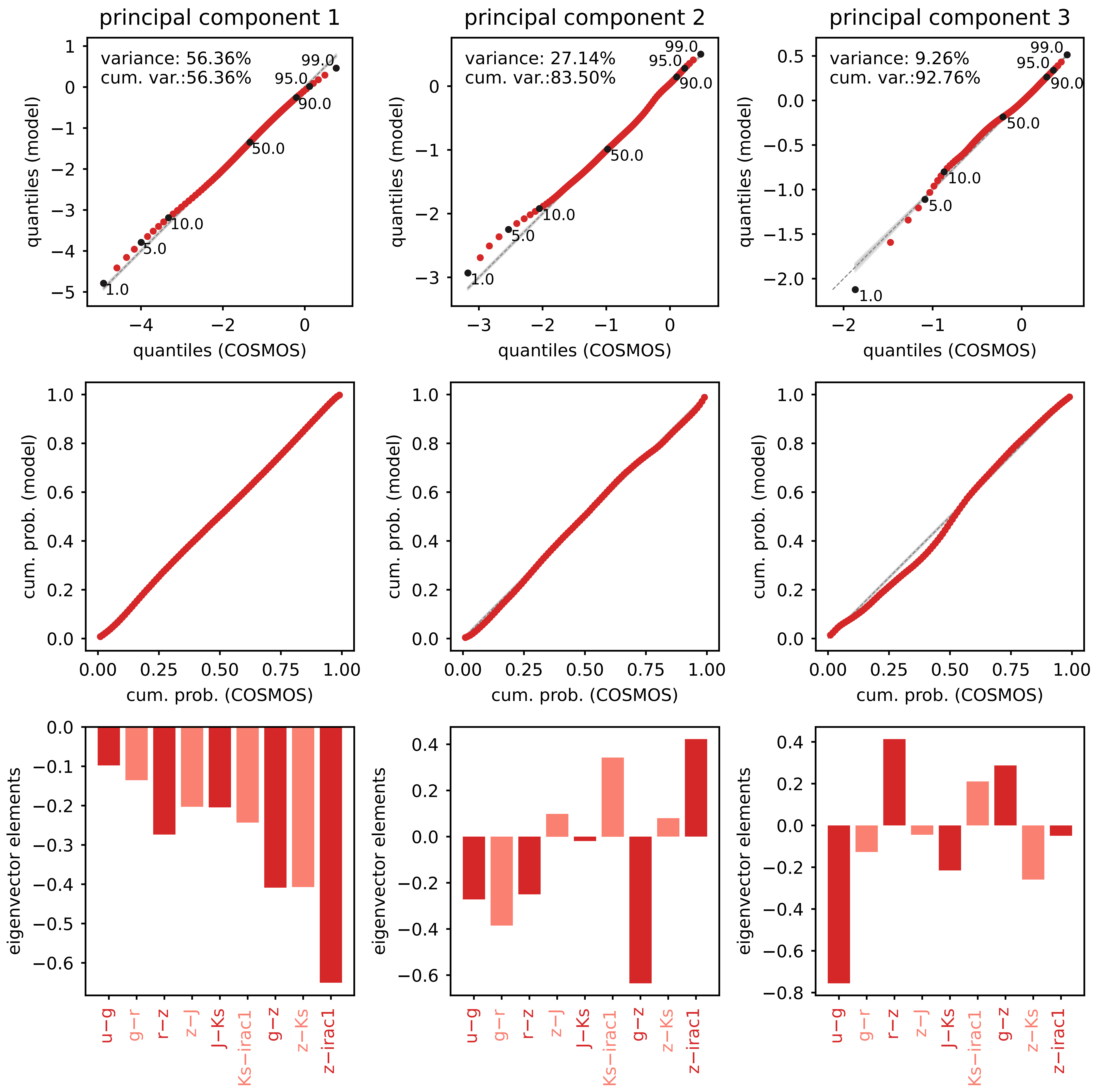


Alsing et al. (2024)

Validation

Data-space validation using Q-Q and P-P plots

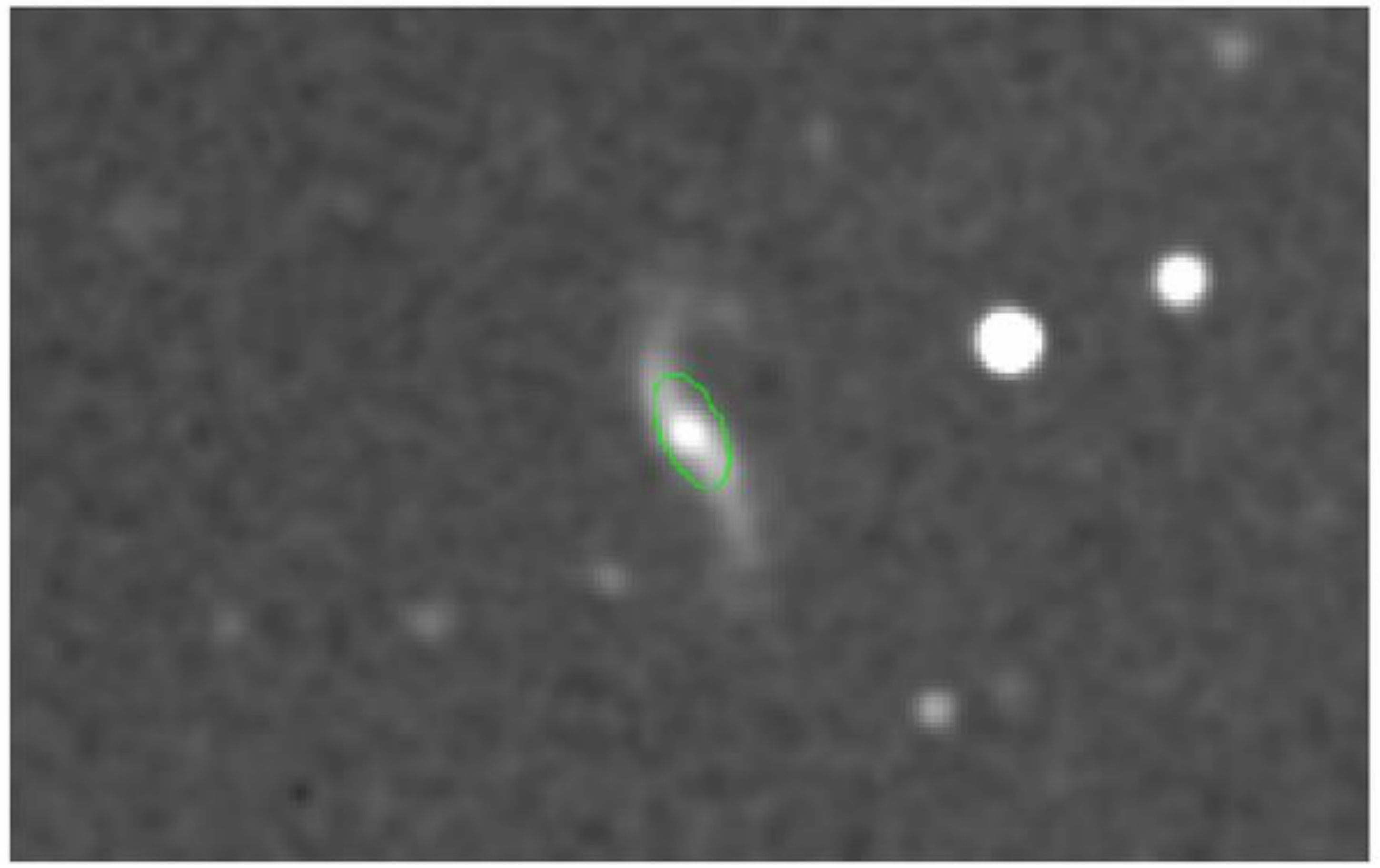




Thorp et al. (2024)

Image Based Selection

- Some galaxy surveys have image-level selection (e.g. based on shape, PSF)
- We can train a conditional density estimator $P(\text{image property} \mid \varphi)$ to learn a model for the relevant image-level summary statistics given SPS parameters
- Allows us to forward model these catalogs without simulating full images



KiDS: *Kuijken et al. (2015)*

Our stellar population synthesis parameters

$\approx$ Prospector-α

$$\varphi = [\log_{10}(M/M_\odot),$$ stellar mass

$$\Delta \log_{10}(\mathrm{SFR}),$$ star forming history ($\times$ 7)

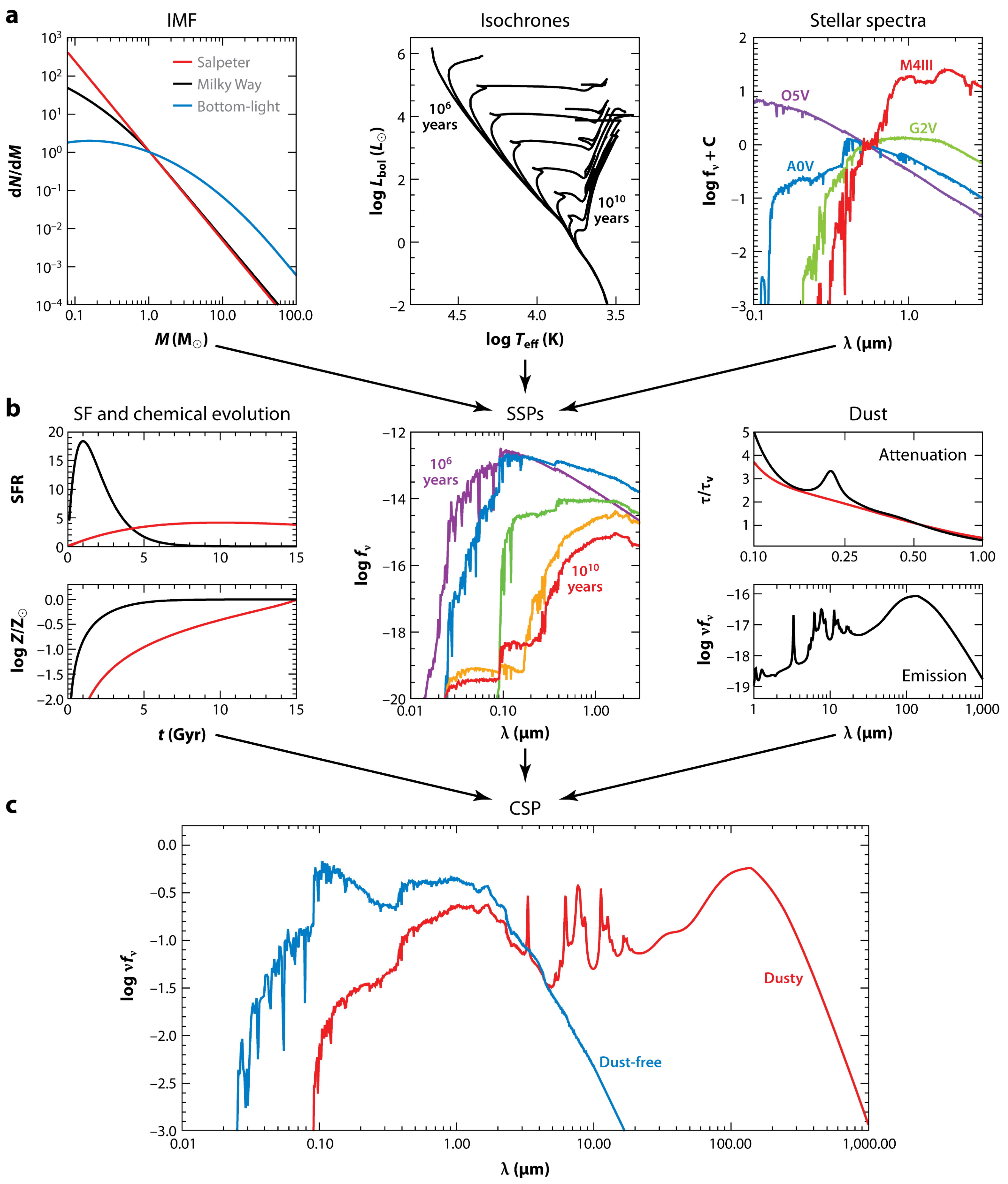
$$\tau_1, \tau_2, n,$$ dust attenuation

$$\ln(f_{\mathrm{AGN}}), \ln(\tau_{\mathrm{AGN}}),$$ active galactic nuclei

$$\log_{10}(Z/Z_\odot),$$ stellar metallicity

$$\log_{10}(Z_{\mathrm{gas}}/Z_\odot), \log_{10}(U_{\mathrm{gas}})]$$ gas metallicity + ionisation

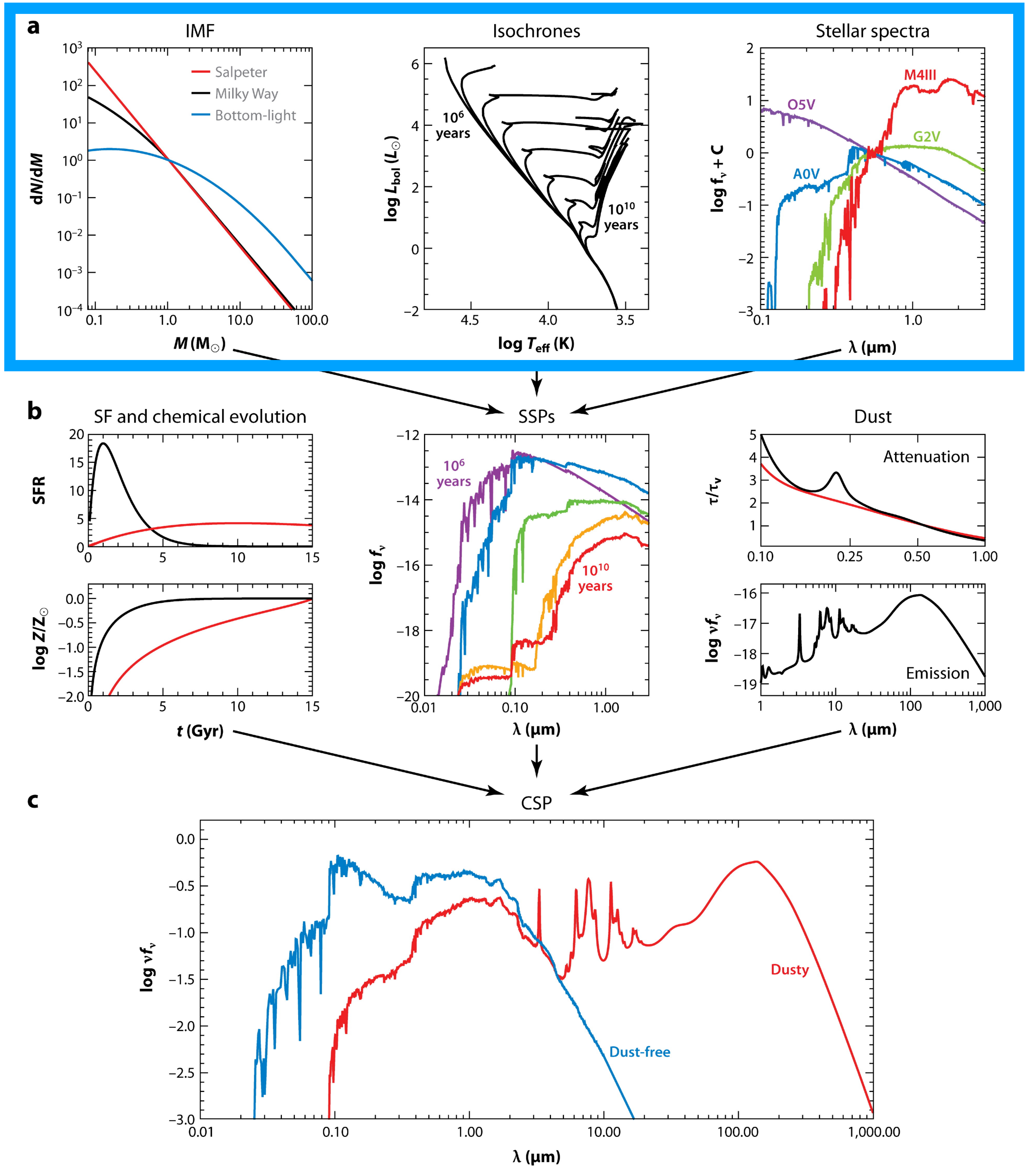
What goes into stellar population synthesis?



Conroy (2013)

What goes into stellar population synthesis?

Assembly of simple stellar populations from initial mass function, isochrones, and stellar spectra

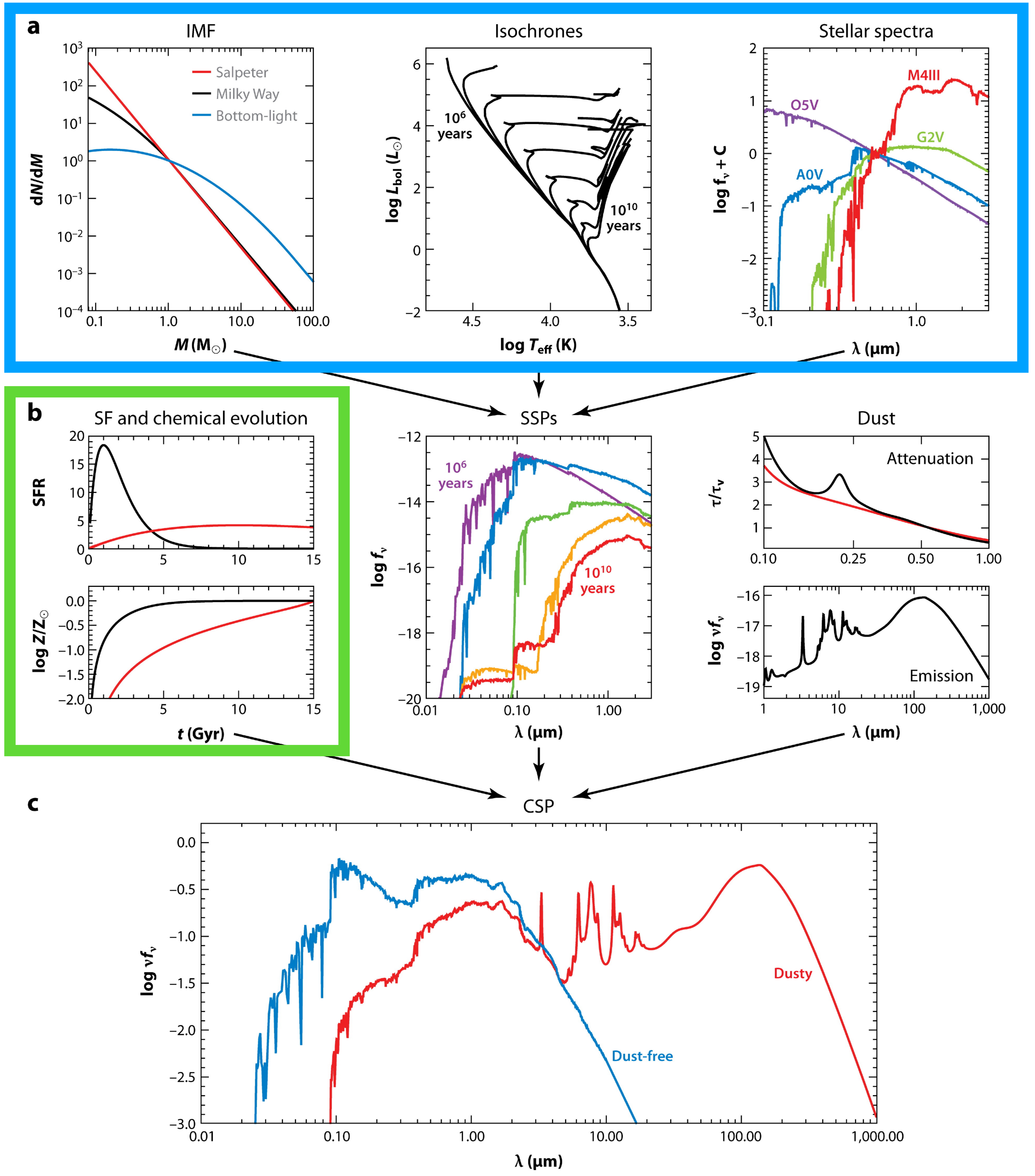


Conroy (2013)

What goes into stellar population synthesis?

Assembly of simple stellar populations from initial mass function, isochrones, and stellar spectra

Star formation history / evolution



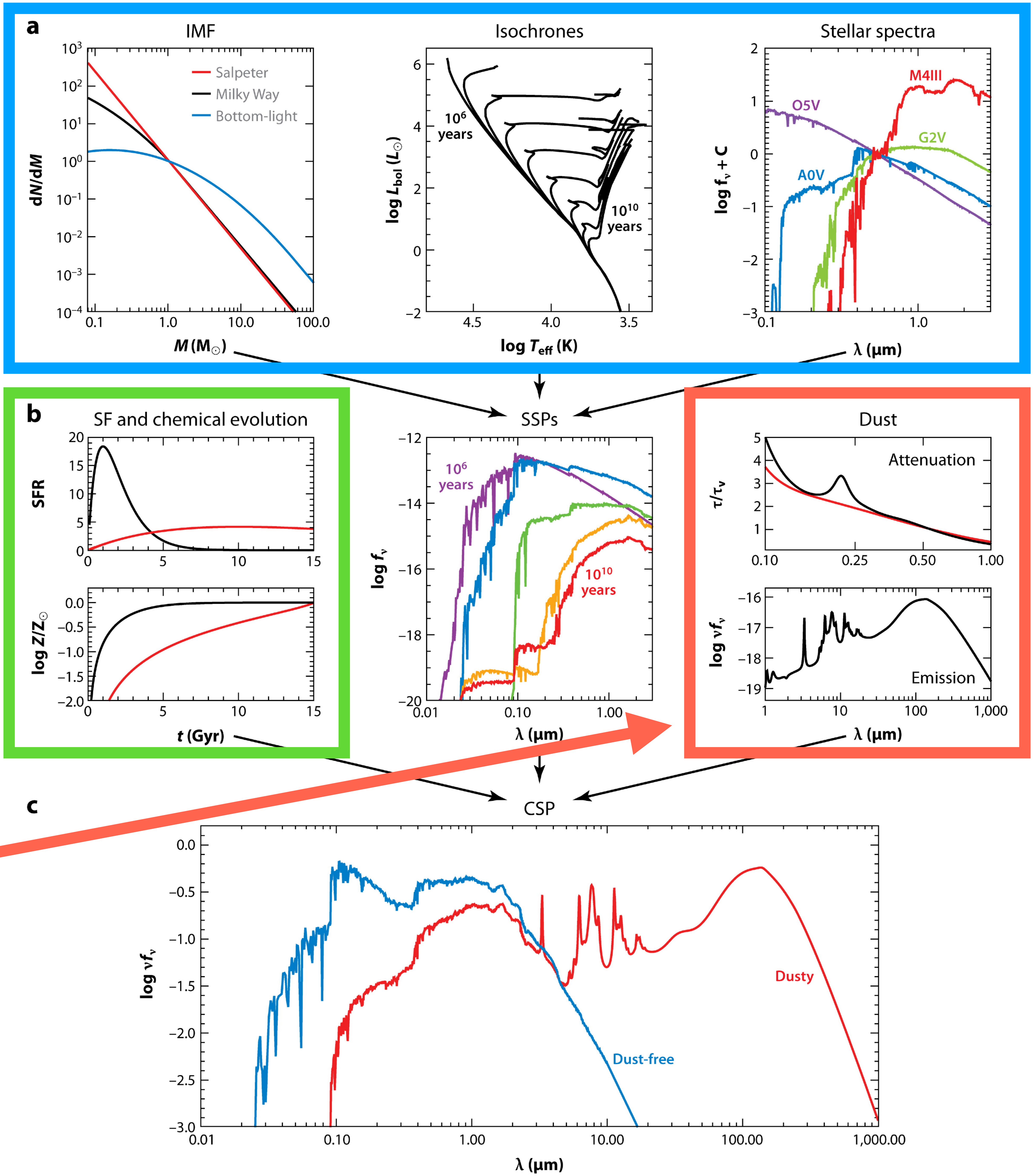
Conroy (2013)

What goes into stellar population synthesis?

Assembly of simple stellar populations from initial mass function, isochrones, and stellar spectra

Star formation history / evolution

Dust attenuation and emission



Conroy (2013)

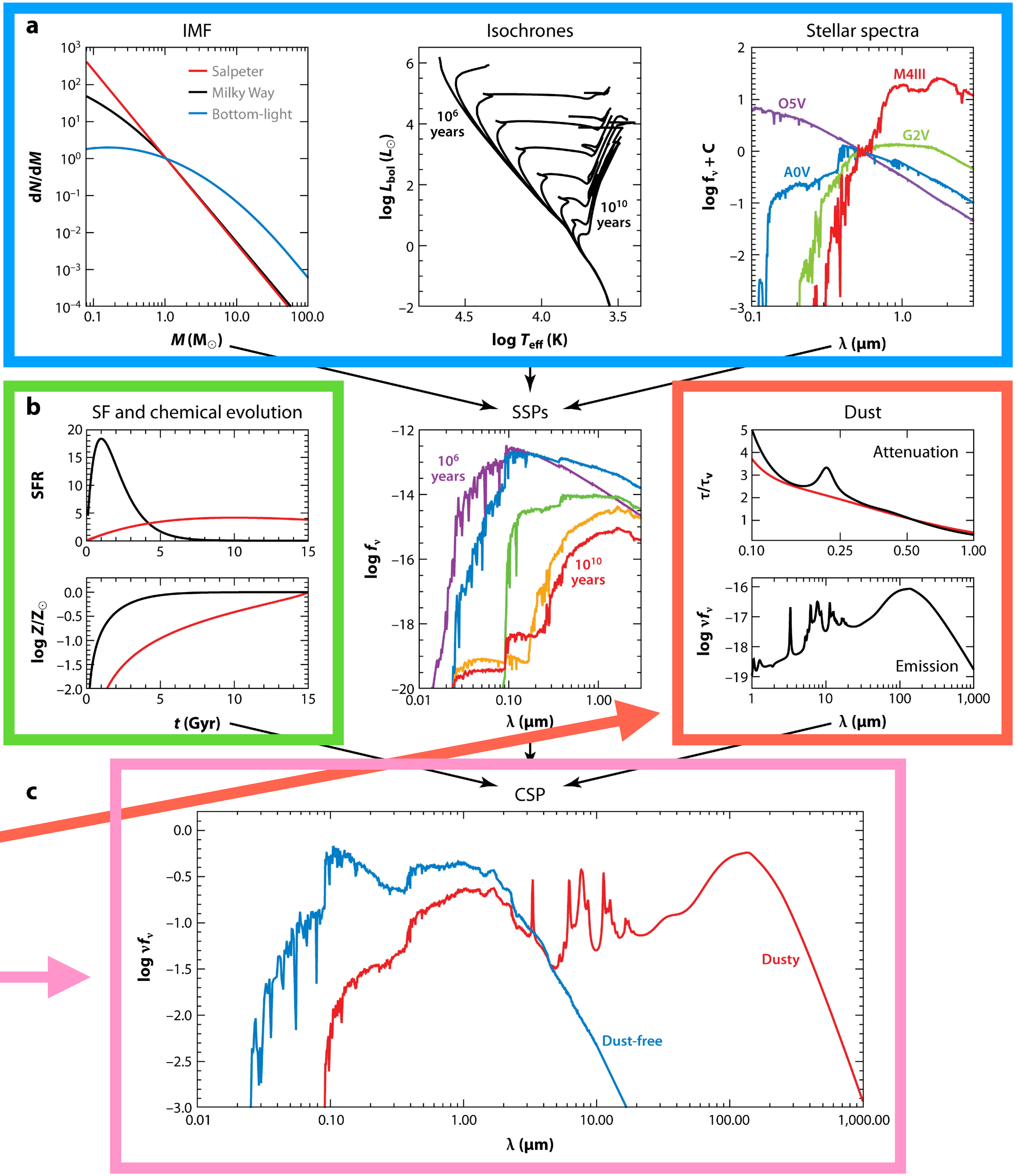
What goes into stellar population synthesis?

Assembly of simple stellar populations from initial mass function, isochrones, and stellar spectra

Star formation history / evolution

Dust attenuation and emission

Assembly of composite stellar populations



Conroy (2013)

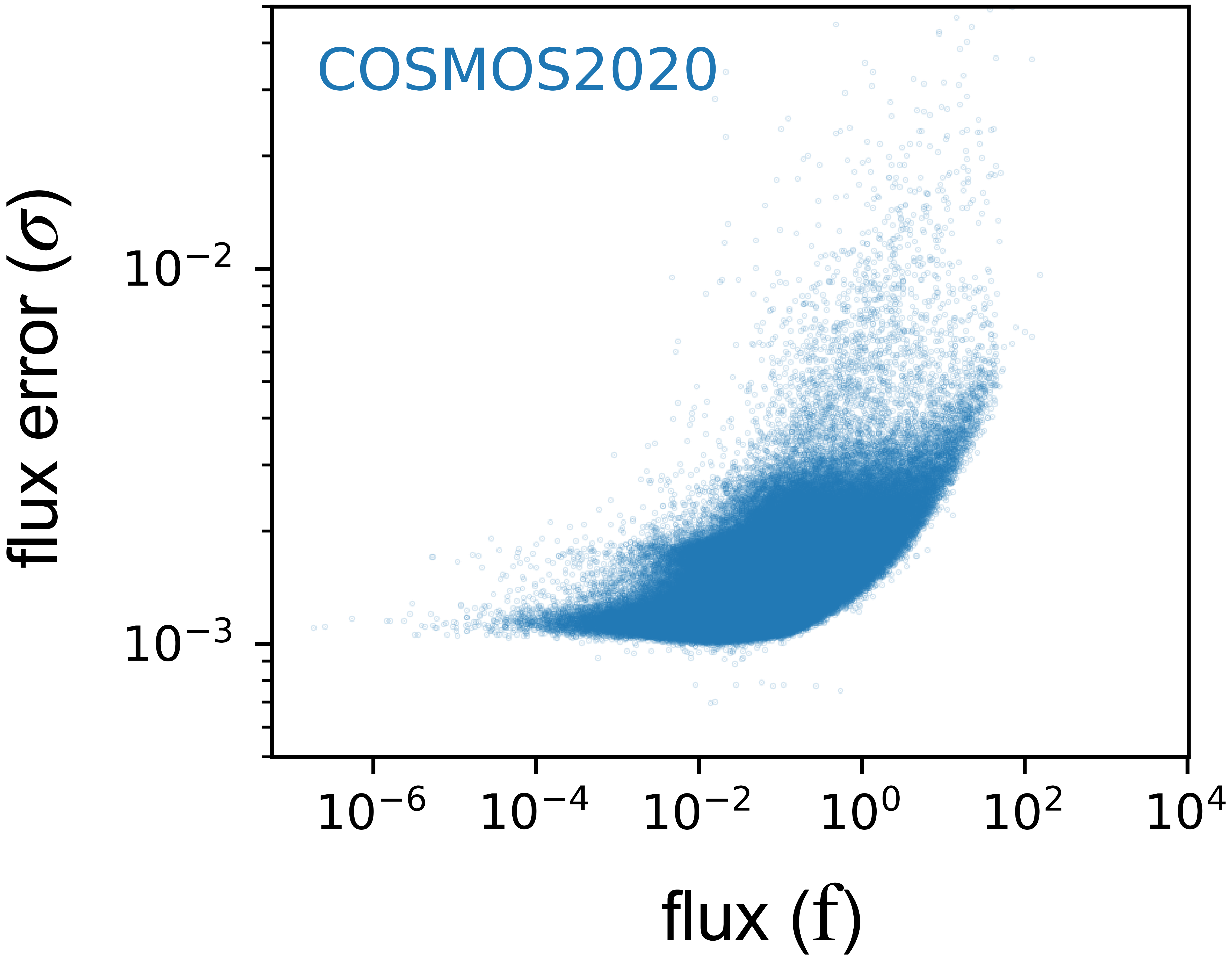
Our uncertainty model

Mixture density network (see Bishop 2006)

$$\mathbf{s} \sim N(0,1)$$

$$\sigma = \boxed{\mu(\mathbf{f})} + \boxed{\Sigma(\mathbf{f})} \odot \mathbf{s}$$

dense neural network (2 layers, 128 hidden units, tanh activation)



More on diffusion models

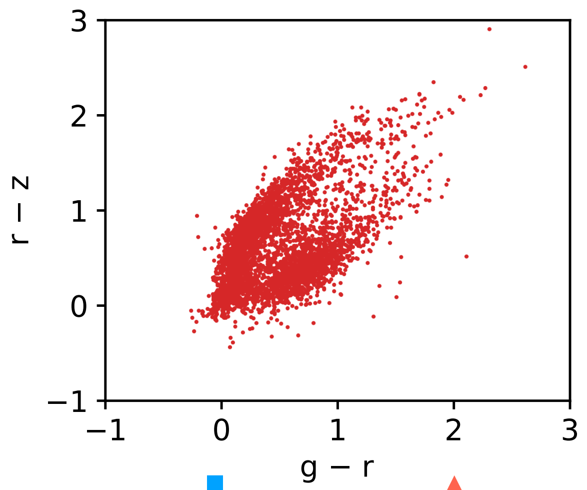
Variance-exploding SDE

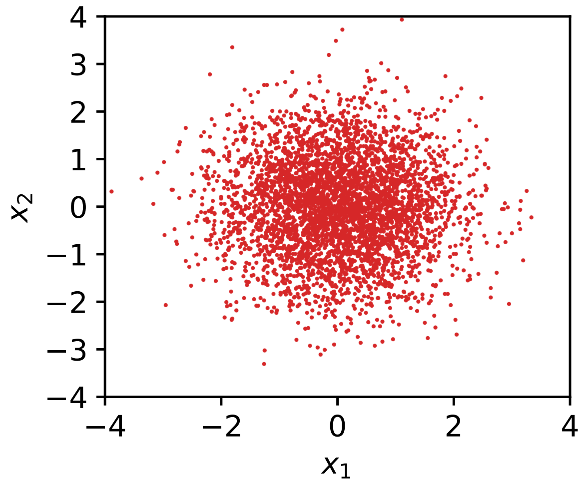
$$\mathbf{x}(t=0) \sim P_0(\mathbf{x}) \equiv P(\varphi, z)$$

$$\mathrm{d}\mathbf{x} = g(t)\,\mathrm{d}\mathbf{w}$$ forward

reverse $$\mathrm{d}\mathbf{x} = -\frac{1}{2} g^2(t)\,\nabla_{\mathbf{x}} P_t(\mathbf{x})\,\mathrm{d}t$$

$$\mathbf{x}(t=T) \sim P_T(\mathbf{x}) \equiv N(0,1)$$





$$s(\mathbf{x}, t) = \nabla_{\mathbf{x}} P_t(\mathbf{x})$$ score function

$$\mathrm{d}\mathbf{x} = g(t)\,\mathrm{d}\mathbf{w}$$ Brownian motion ($\mathbf{w}$)

$$\mathrm{d}\mathbf{x} = -\frac{1}{2} g^2(t)\,\nabla_{\mathbf{x}} P_t(\mathbf{x})\,\mathrm{d}t$$

$$g^2(t) = \frac{\mathrm{d}\sigma^2(t)}{\mathrm{d}t}$$

$$\sigma(t) = \sigma_0(\sigma_T/\sigma_0)^{t/T}$$

$$\sigma_0 = 0.01,\ \sigma_T = 6,\ T = 1$$