Dissecting Cluster Cosmology, IFPU, Trieste
(Day 3: July 5, 2023)

Cosmological Mass Calibration of Galaxy Clusters with Weak Lensing: Gaussian-likelihood and Likelihood-Free Approaches

Keiichi Umetsu (ASIAA, Taiwan)

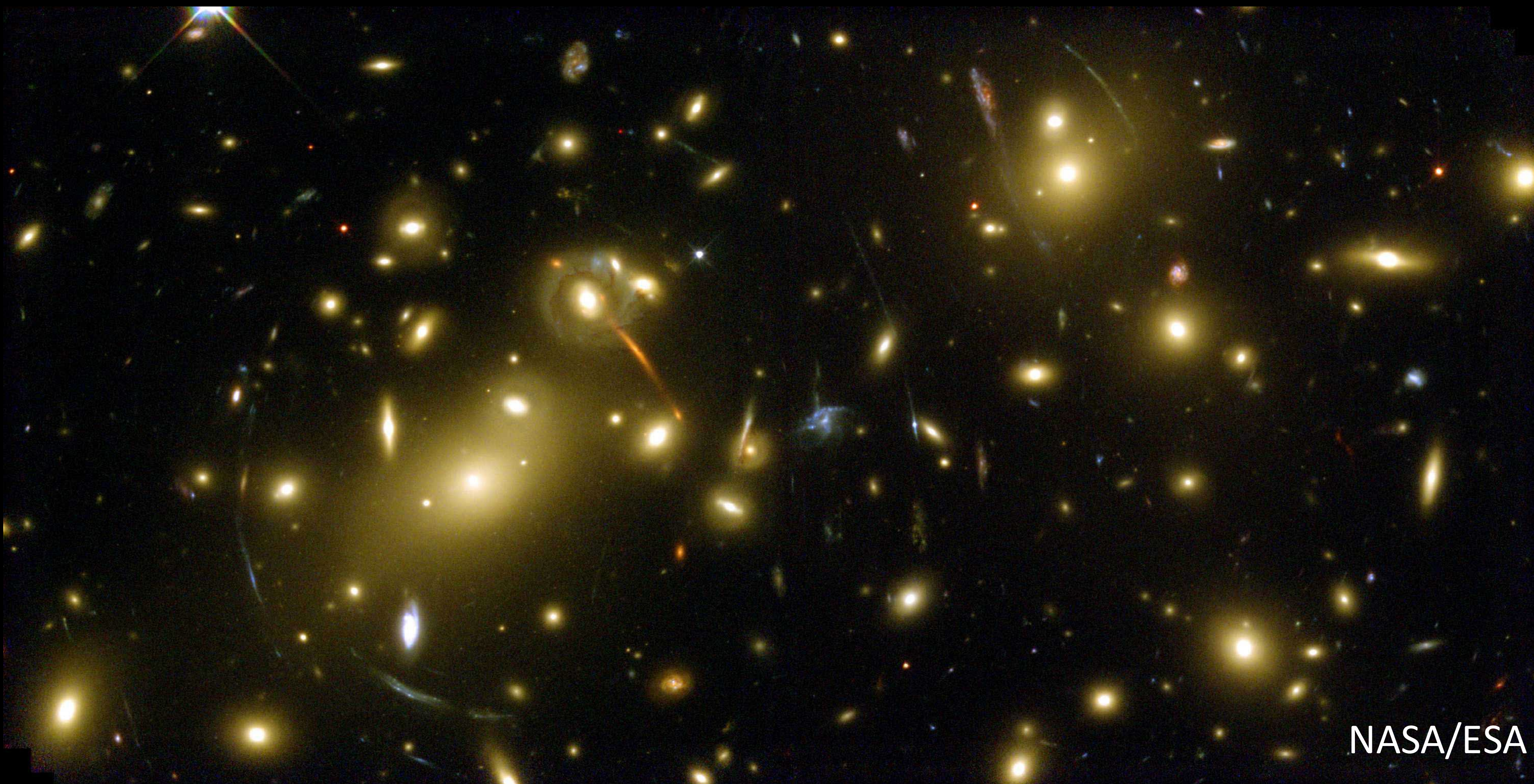
Key issues for cluster cosmology

$$\frac{\mathrm{d}N(\mathcal{O},z|\mathbf{p})}{\mathrm{d}\mathcal{O}\mathrm{d}z} = \Theta(\mathcal{O},z)\int \mathrm{d}M \left[P(\mathcal{O}|M,z,\mathbf{p})\frac{\mathrm{d}n(M,z|\mathbf{p})}{\mathrm{d}M\mathrm{d}z}V(z,\mathbf{p})\right]$$

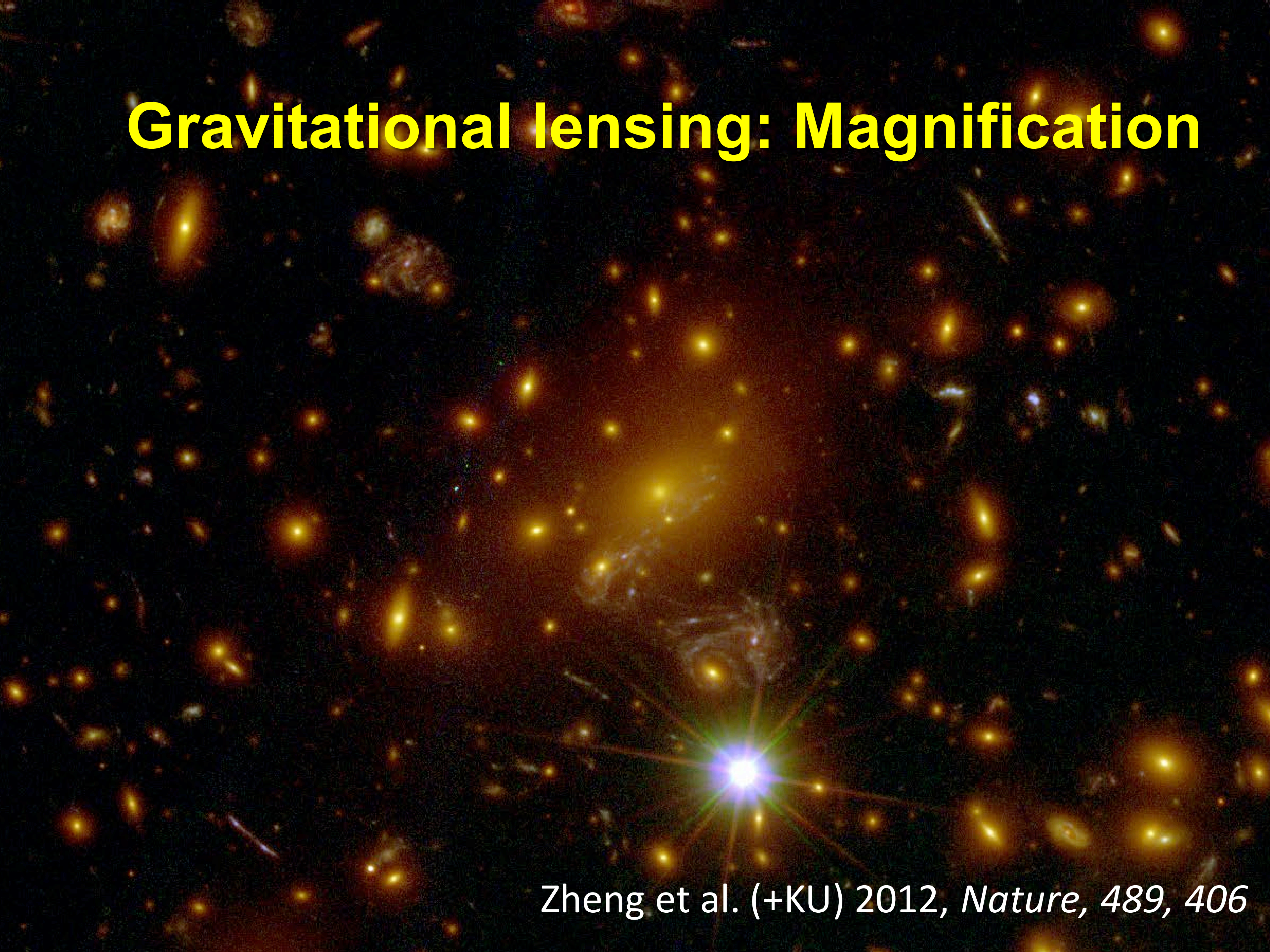
selection — observable–mass relation — mass function

- Accurate determination of the “mass scale” of an observable-selected sample is key to use clusters as a cosmological probe
  - Need to account for all relevant physical and observational effects in cosmological inference and mass calibration (< a few % for all-sky surveys)
- It is also essential to have accurate characterization of observable-selected samples in terms of:
  - Survey selection function across the redshift range
  - Observable-mass scaling relations and their intrinsic scatter (covariance)

Gravitational lensing: Shear



Gravitational lensing: Magnification



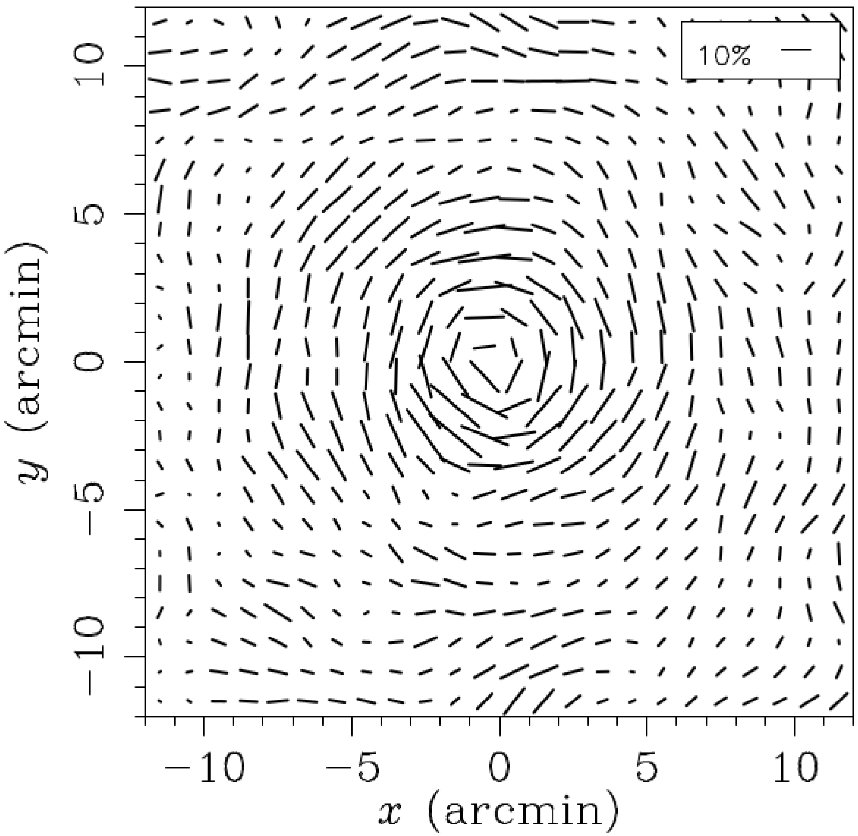
Zheng et al. (+KU) 2012, *Nature, 489, 406*

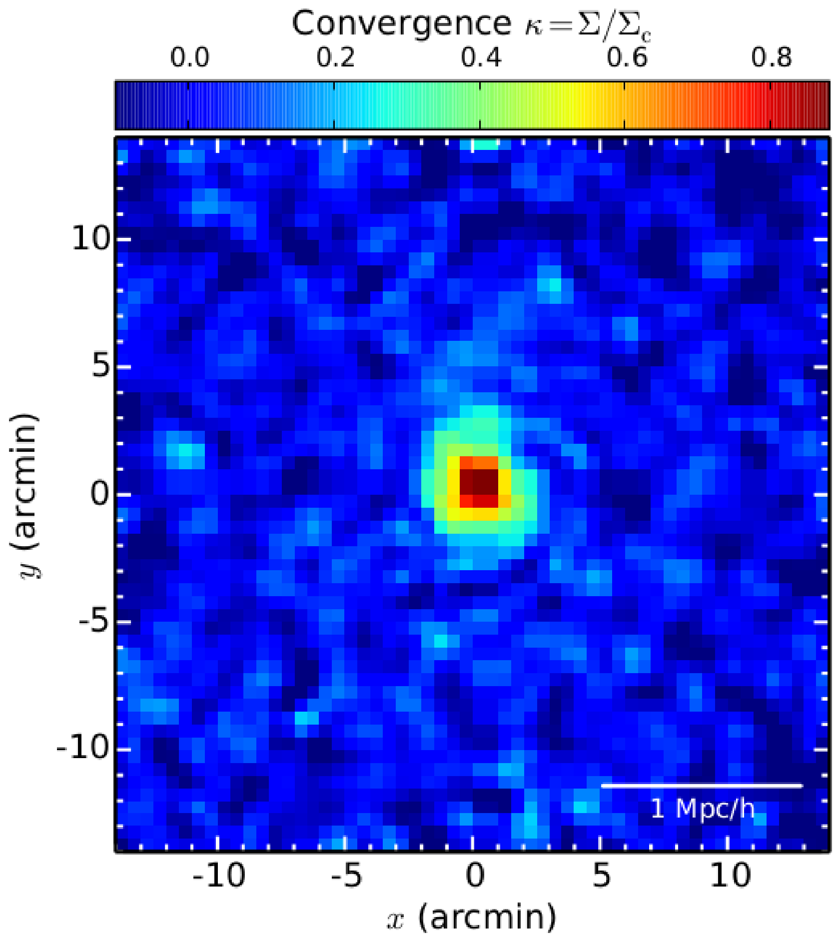
Shear and Convergence

$$\Gamma_{\alpha\beta} = (\partial_\alpha \partial_\beta - \delta_{\alpha\beta}\Delta/2)\psi$$

$$\kappa = \Delta^{-1}(\partial^\alpha \partial^\beta \Gamma_{\alpha\beta})$$

Shear field: Abell 1689 (Umetsu+15)



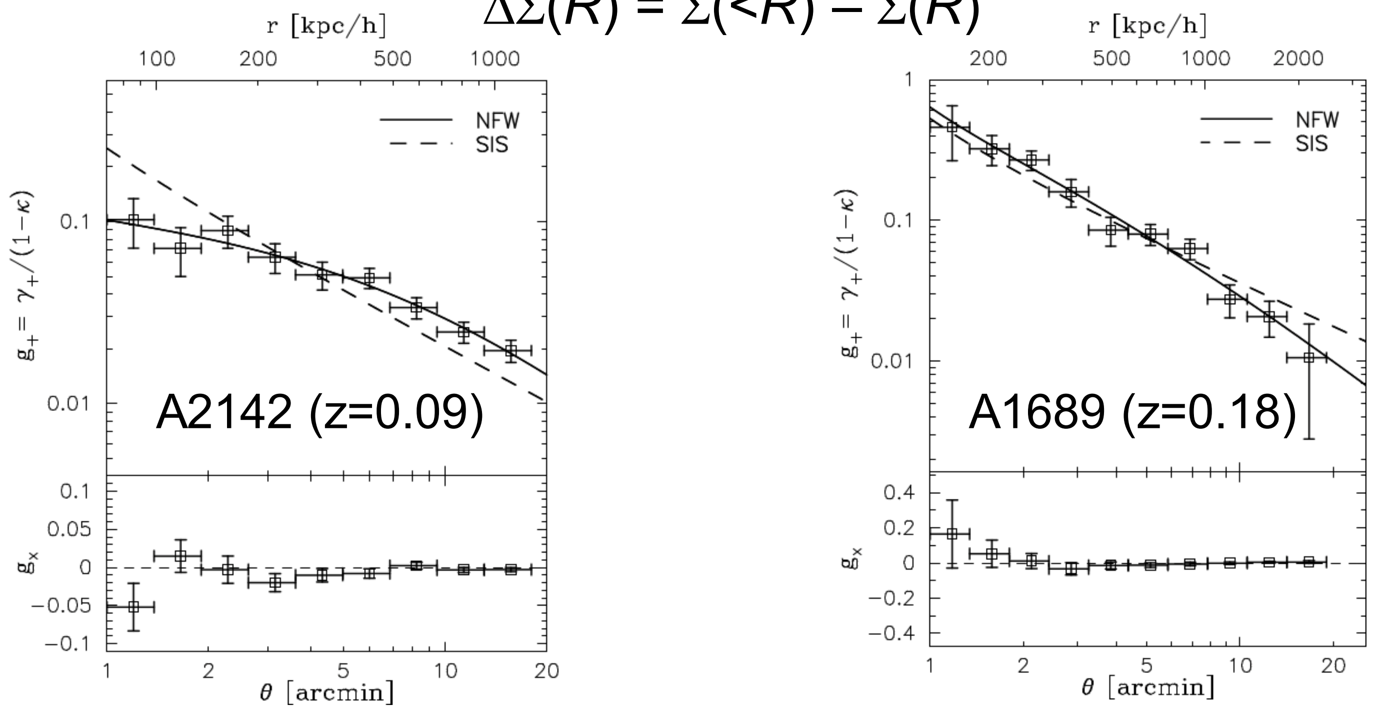


Review: Umetsu 2020, arXiv:2007.00506

Weak-lensing (WL) mass measurements

Tangential shear fitting of individual clusters with an NFW halo description (or its variants: e.g., Einasto, DK14, halo model. ..)

$$\Delta\Sigma(R) = \Sigma(<R) - \Sigma(R)$$



Covariance matrix of $g_+(R)$ at fixed halo mass (Hoekstra 03; Gruen+15; Umetsu+16)

$$C = C^{\rm shape} + C^{\rm lss} + C^{\rm int}$$

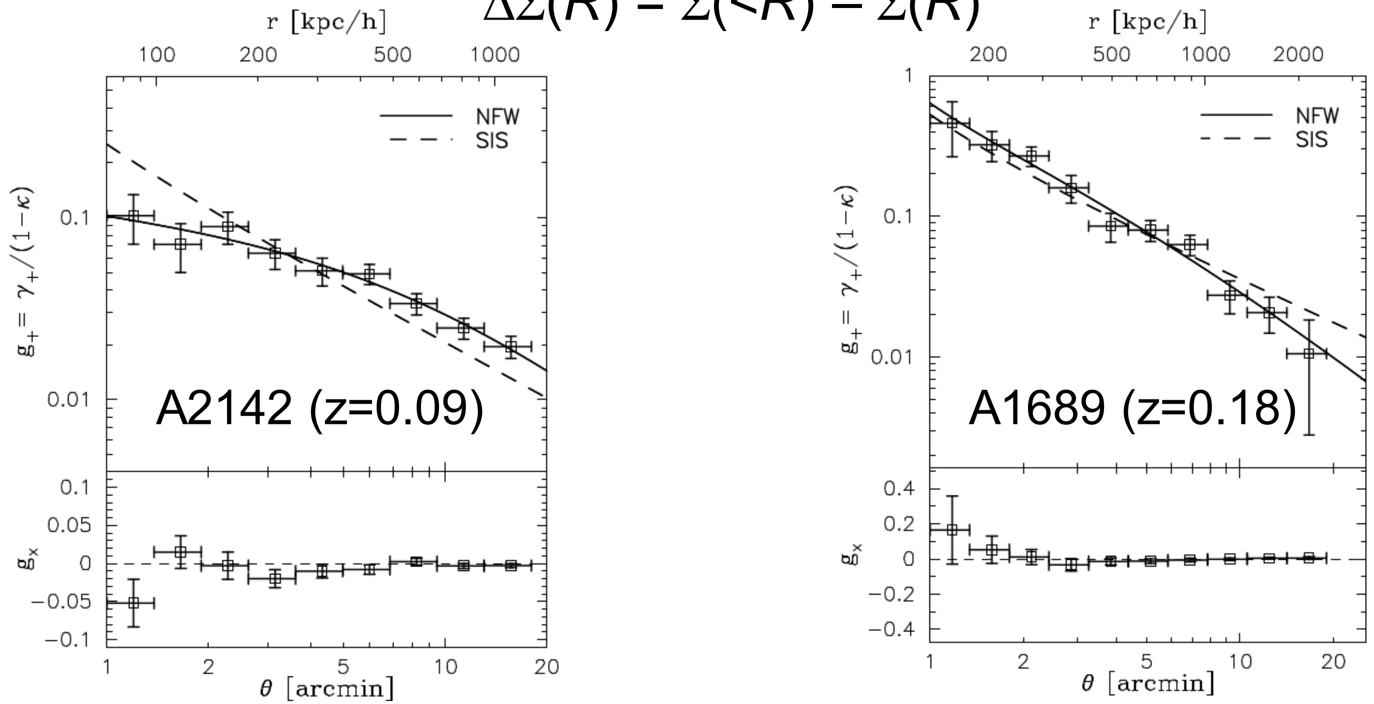
Other features: miscentering (see Tim's talk), baryonic feedback, splashback feature, 2-halo term

For a review, Umetsu 2020, **A&ARv**,28, 7

Weak-lensing (WL) mass measurements

Tangential shear fitting of individual clusters with an NFW halo description (or its variants: e.g., Einasto, DK14, halo model. ..)

$$\Delta\Sigma(R) = \Sigma(<R) - \Sigma(R)$$



Covariance matrix of $g_+(R)$ at fixed halo mass (Hoekstra 03; Gruen+15; Umetsu+16)

$$C = C^{\rm shape} + C^{\rm lss} + C^{\rm int}$$

Other features: miscentering (see Tim's talk), baryonic feedback, splashback feature, 2-halo term

For a review, Umetsu 2020, **A&ARv**,28, 7

Ensemble mass modeling procedures

1. Conventional **single-mass-bin fit** to the stacked lensing profile <g_+>(R)
2. **Population modeling of the stacked lensing profile**, accounting for the mass function and the selection function (e.g., Miyatake+19: ACTPol-HSC)
3. **Hierarchical Bayesian population modeling of all individual clusters** in
   - $M_{\rm WL}$ domain (Umetsu+20: XXL-HSC)
   - $g_+(R)$ domain (Chiu+22: eFEDS-HSC)
4. **Likelihood-free inference using forward simulations** (Tam, Umetsu, & Amara+22)

$d\ln(\gamma_+)/d\ln(M_{200}) \sim 1$ (Oguri & Takada 11)

More flexible: properly accounting for selection/statistical effects, mass modeling bias, etc.

Multivariate Gaussian Likelihood Approach

Bayesian population modeling of cluster observables

Multivariate Gaussian likelihood function $P(x,\boldsymbol{y}|\boldsymbol{\theta})$

e.g., Kelly 07 (LINMIX_ERR in IDL), Sereno 16 (LIRA in R), Okabe (HiBRECS in Python; Akino+22 HSC+XXL)

Posterior probability: $P(\boldsymbol{\theta}|x,\boldsymbol{y}) \propto P(x,\boldsymbol{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})$

$$p(x,\boldsymbol{y}|\boldsymbol{\theta}) = \frac{\prod_n^N \int_{-\infty}^{\infty} dX_n \int_{-\infty}^{\infty} d^D\boldsymbol{Y}_n \int_{-\infty}^{\infty} dZ_n\, p(x_n,\boldsymbol{y}_n|X_n,\boldsymbol{Y}_n)p(X_n,\boldsymbol{Y}_n|Z_n,\boldsymbol{\theta})p(Z_n|\boldsymbol{\theta})}{\prod_n^N \int_{y_{\mathrm{th},0,n}}^{\infty} dy_{0,n} \int_{-\infty}^{\infty} dx_n \int_{-\infty}^{\infty} d^{D-1}\boldsymbol{y}_n \int_{-\infty}^{\infty} dX_n \int_{-\infty}^{\infty} d^D\boldsymbol{Y}_n \int_{-\infty}^{\infty} dZ_n\, p(x_n,\boldsymbol{y}_n|X_n,\boldsymbol{Y}_n)p(X_n,\boldsymbol{Y}_n|Z_n,\boldsymbol{\theta})p(Z_n|\boldsymbol{\theta})}$$

$$Z = \ln M_{\Delta,\mathrm{true}}$$
$$X = \ln M_{\Delta,\mathrm{WL}}$$
$$Y = \ln O_{\Delta}$$

Observable vs. true-mass relations:

$$X = \alpha_{X|Z} + \beta_{X|Z}\, Z \pm \sigma_{X|Z} \quad (\alpha_{X|Z} \sim 0, \beta_{X|Z} \sim 1, \sigma_{X|Z} \sim 0.2)$$
$$Y = \alpha_{Y|Z} + \beta_{Y|Z}\, Z \pm \sigma_{Y|Z}$$

M_{WL} as a scattered, weakly biased proxy for M_{true}

O: Baryonic cluster property (e.g., L_X, T_X, M_{gas}, Y_{SZ})

Key ingredients for Bayesian population modeling

- Probabilistic model for the parent population: $P(Z)$
  - Intrinsic parent distribution in "true" halo mass ($Z = \ln M_{\mathrm{true}}$)
    - Modeled by the halo mass function $n(M,z)$ and survey selection function
    - Or approximated by a time-evolving Gaussian mixture model (Sereno 16)
- Conditional probability for mass proxies: $P(X,Y|Z)$

WL mass: $X = \alpha_{X|Z} + \beta_{X|Z} Z \pm \sigma_{X|Z} \quad (\alpha_{X|Z} \sim 0, \beta_{X|Z} \sim 1, \sigma_{X|Z} \sim 0.2)$

Baryonic properties $Y = \alpha_{Y|Z} + \beta_{Y|Z} Z + \gamma_{Y|Z} \ln[E(z)/E(z_{\mathrm{pivot}})] \pm \sigma_{Y|Z}$

→ Selection/statistical effects (e.g., Malmquist bias, Eddington bias) and mass modeling bias can be statistically accounted for

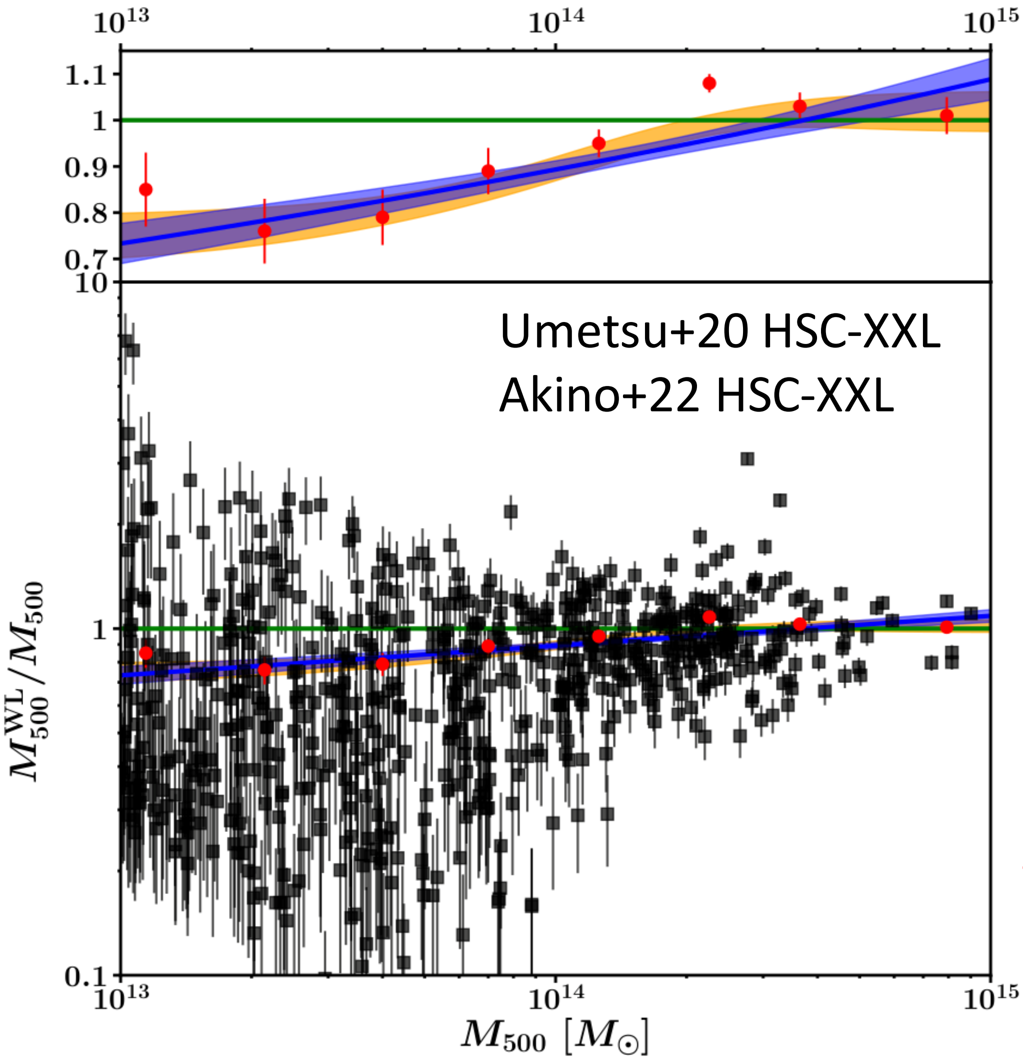
→ Joint multivariate modeling of intrinsic scatter covariance of (X,$\boldsymbol{Y}$)

- Conditional probability for observed proxies: $P(x,y|X,Y)$

→ Accounting for the impact of large measurement scatter (e.g., regression dilution effect on the slope parameter)

Mass modeling bias as a function of $M_{\rm true}$

BAHAMS DM-only simulation



True M_{500} of N-body CDM halos

WL analysis of BAHAMAS mock datasets (Umetsu+20):

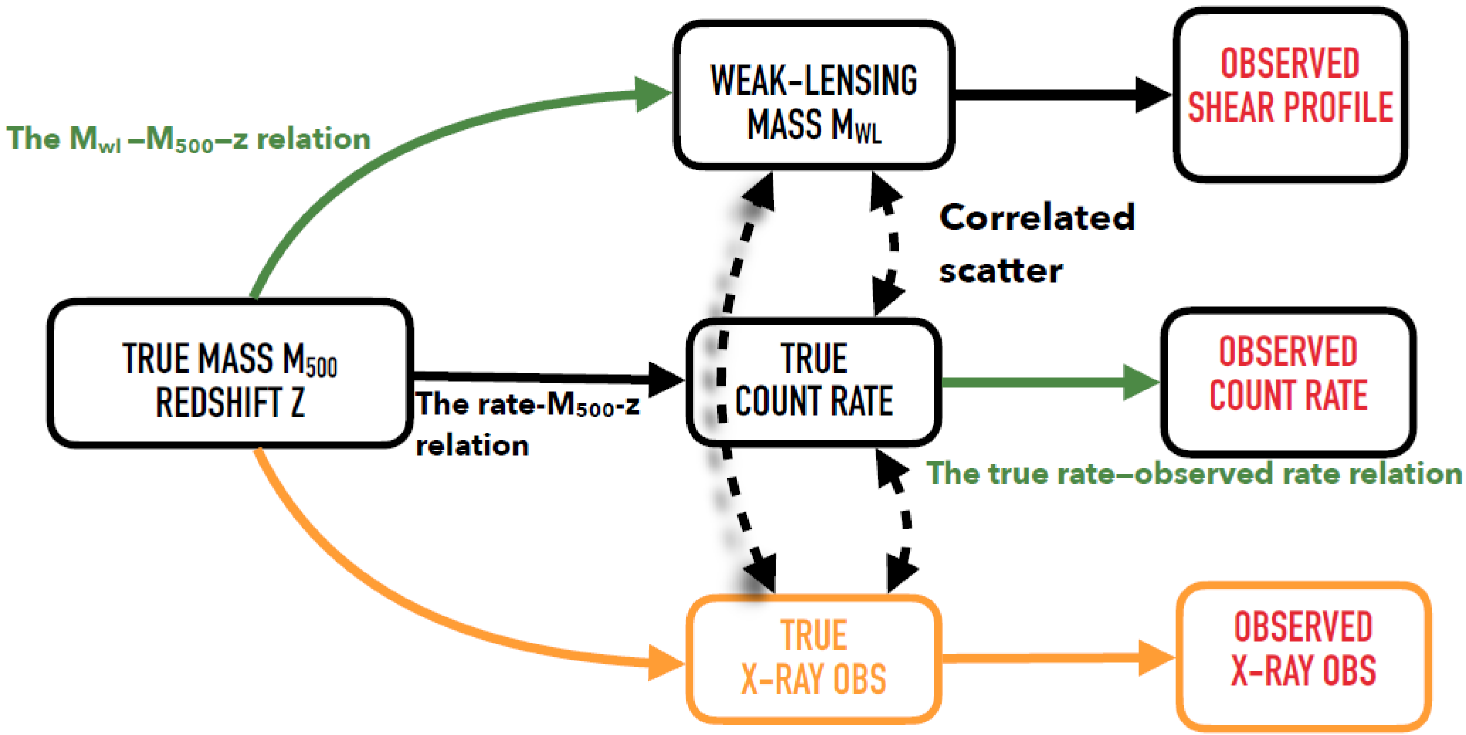
$d\ln(\gamma_+)/d\ln(M_{200}) \sim 1$
(Oguri & Takada 11)

- For $M_{\rm true} > 2e14 M_{\rm sun}/h$, bias is essentially zero.
- At group scales, $M_{\rm WL}$ at fixed $M_{\rm true}$ is underestimated.

Mass modeling bias can be statistically corrected for in a forward-modeling manner

See Giocoli+23 (arXiv:2302.00687) for mass bias expected in Euclid WL

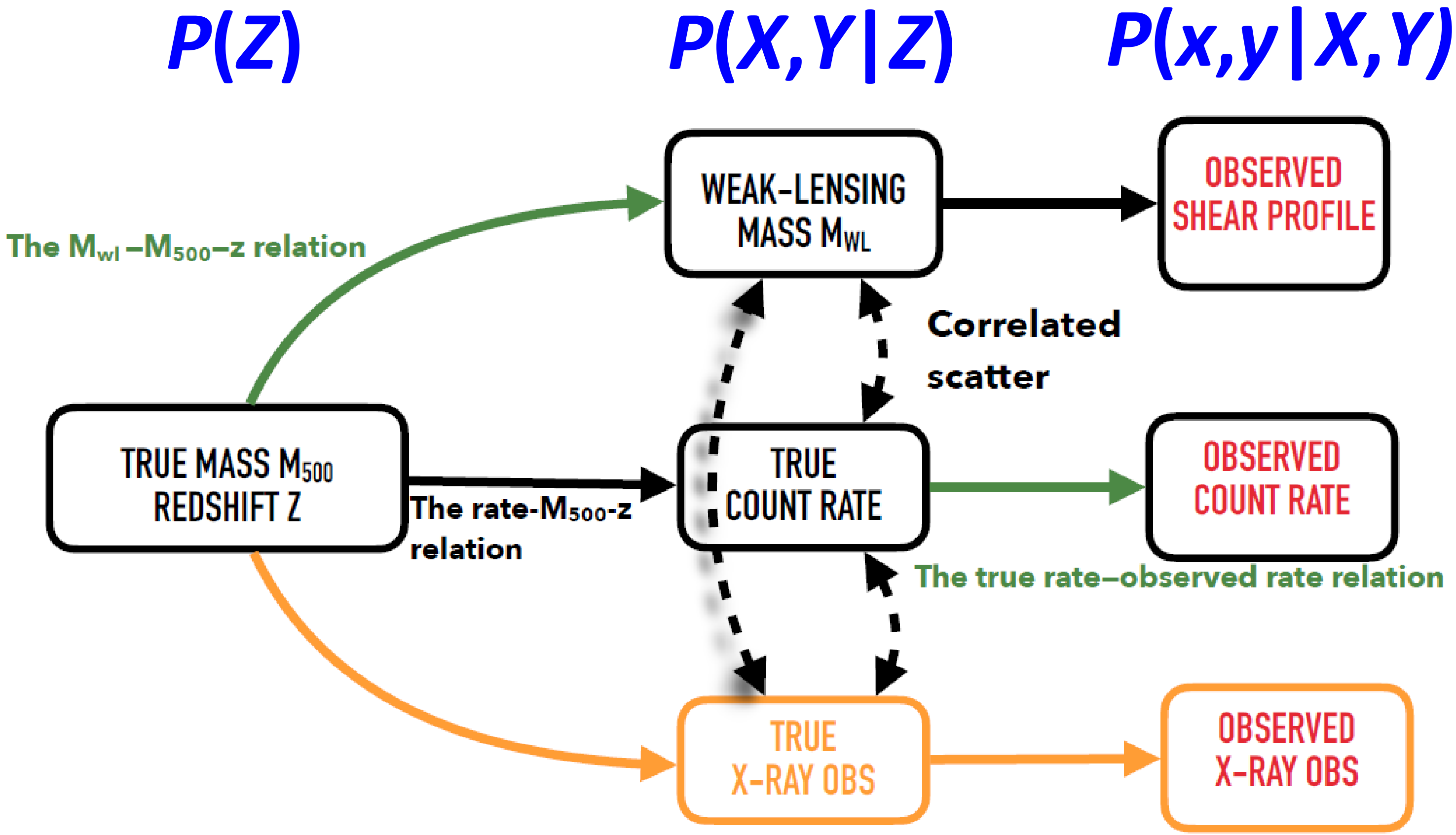
One-step Bayesian population modeling



$$\langle \ln \mathscr{X} | M_{500} \rangle = \ln A_{\mathscr{X}} + \left[B_{\mathscr{X}} + \delta_{\mathscr{X}} \ln \left(\frac{1+z}{1+z_{\mathrm{piv}}} \right) \right] \times \ln \left(\frac{M_{500}}{M_{\mathrm{piv}}} \right) + C_{\mathrm{SS},\mathscr{X}} \times \ln \left(\frac{E(z)}{E(z_{\mathrm{piv}})} \right) + \gamma_{\mathscr{X}} \times \ln \left(\frac{1+z}{1+z_{\mathrm{piv}}} \right)$$

Chiu+22
eFEDS-HSC

One-step Bayesian population modeling



$$\langle \ln \mathscr{X} | M_{500} \rangle = \ln A_{\mathscr{X}} + \left[B_{\mathscr{X}} + \delta_{\mathscr{X}} \ln\left(\frac{1+z}{1+z_{\rm piv}}\right)\right] \times \ln\left(\frac{M_{500}}{M_{\rm piv}}\right) + C_{{\rm SS},\mathscr{X}} \times \ln\left(\frac{E(z)}{E(z_{\rm piv})}\right) + \gamma_{\mathscr{X}} \times \ln\left(\frac{1+z}{1+z_{\rm piv}}\right)$$

Chiu+22
eFEDS-HSC

Simulation-based Bayesian Inference

The Astrophysical Journal, 925:145 (14pp), 2022 February 1 https://doi.org/10.3847/1538-4357/ac3d33

Likelihood-free Forward Modeling for Cluster Weak Lensing and Cosmology

Sut-Ieng Tam[1], Keiichi Umetsu[1], and Adam Amara[2]
[1] Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan;keiichi@asiaa.sinica.edu.tw
[2] Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Burnaby Road, Portsmouth PO1 3FX, UK
Received 2021 September 20; revised 2021 November 6; accepted 2021 November 23; published 2022 February 2

Demonstration for cluster abundance cosmology combined with weak-lensing mass calibration

Sut-Ieng Tam, Keiichi Umetsu, & Adam Amara 2022, *ApJ*, 925, 145

Simulation-based inference: ABC and DELFI

Bayesian inference requires that **the likelihood of data given a model is known**. In general, **Gaussian likelihoods are assumed**.

This Gaussian assumption may not be valid because of

- Complex data reduction and measurement processes (e.g., filtering, masking, data compression: catalog creation)
- Non-gaussian nature of galaxy clusters (non-gaussian signals)

Likelihood-free methods use forward simulations to **bypass the need for an evaluation of the likelihood**. These methods sample the model prior space and “compare” observational vs. simulated data summaries to derive the posterior distribution.

We explore 2 complementary methods to develop a likelihood-free framework for cluster cosmology and WL mass calibration:

Approximate Bayesian Computation (ABC)

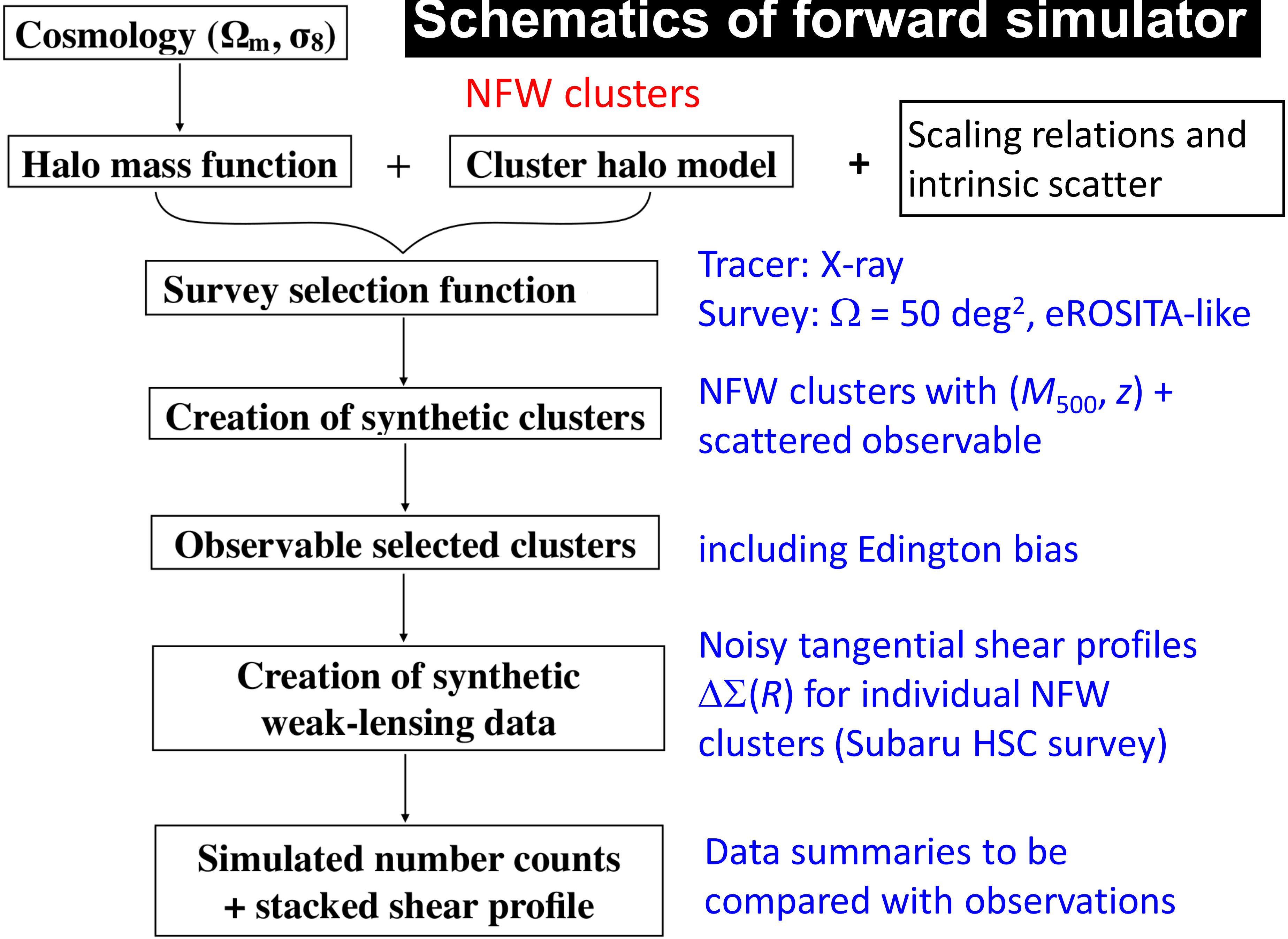
Density-Estimation Likelihood-Free Inference (DELFI)

Tam, KU, Amara 22

Likelihood vs. simulation-based approaches

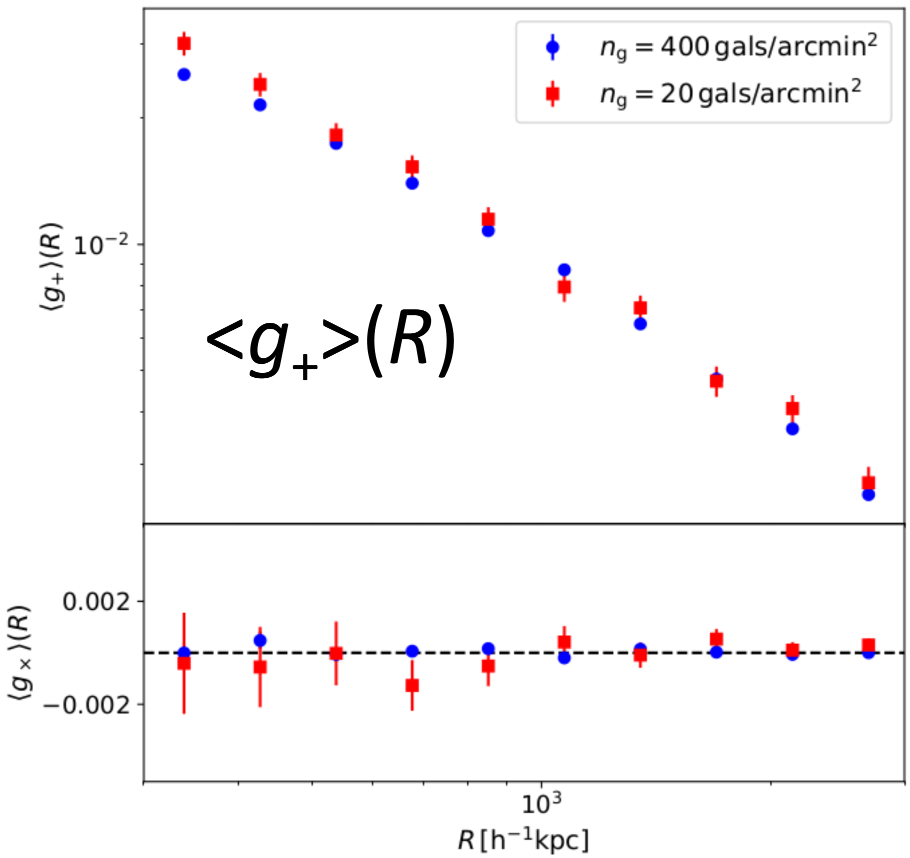
- Gaussian-likelihood inference
  - Gaussian likelihoods assumed a priori
  - Covariance matrices need to be well characterized
  - Exact posterior distribution
  - Fast
- Likelihood-free (simulation based) inference
  - Can incorporate complex physical processes and instrumentation effects in forward simulations
  - Cope with intractable likelihood functions
  - Approximate posterior distribution
  - Computationally expensive!

Schematics of forward simulator

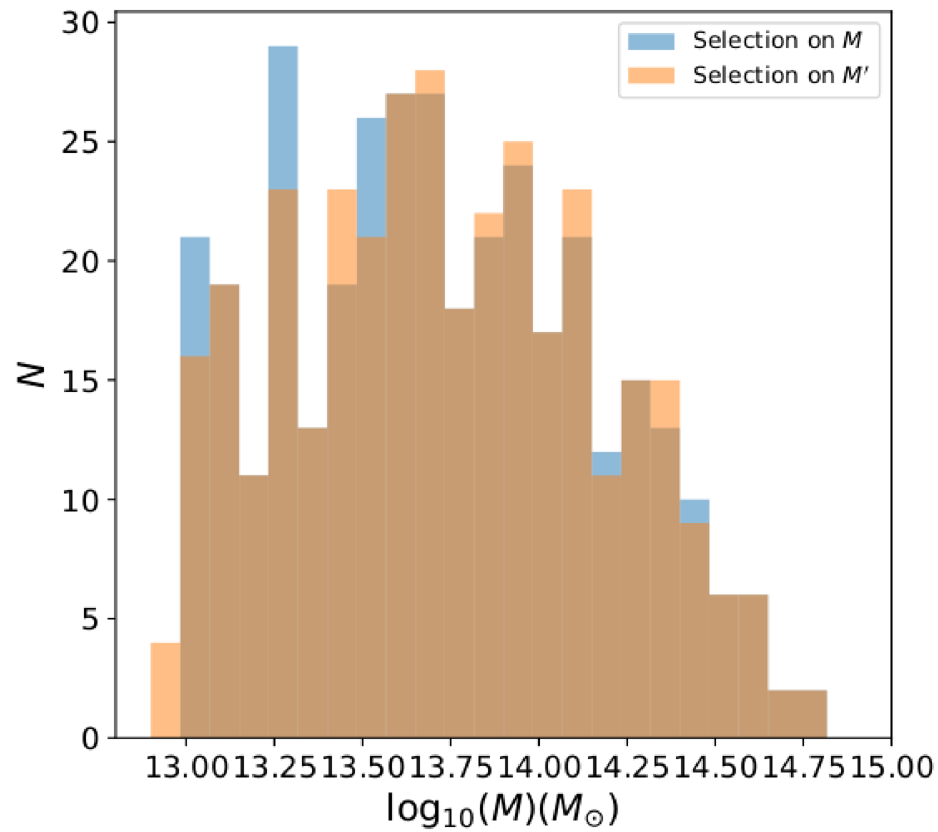


Synthetic survey data created with a fiducial model $\Omega_m = 0.286$, $\sigma_8 = 0.82$

Data summary vector #1



Data summary vector #2

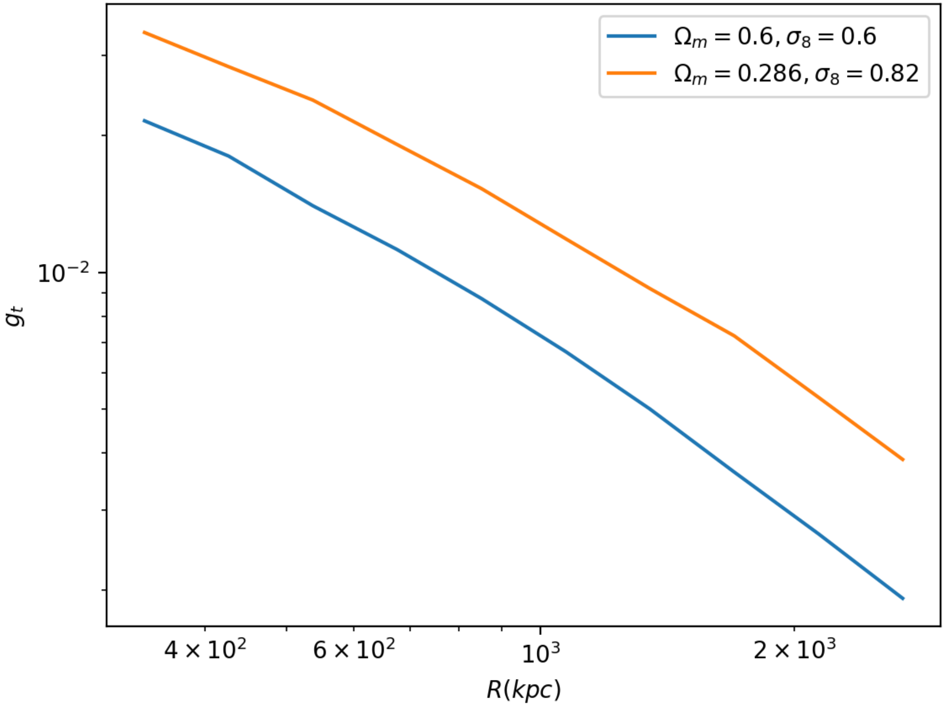


Tam, KU, Amara 22

Predicted data summaries for a given cosmology

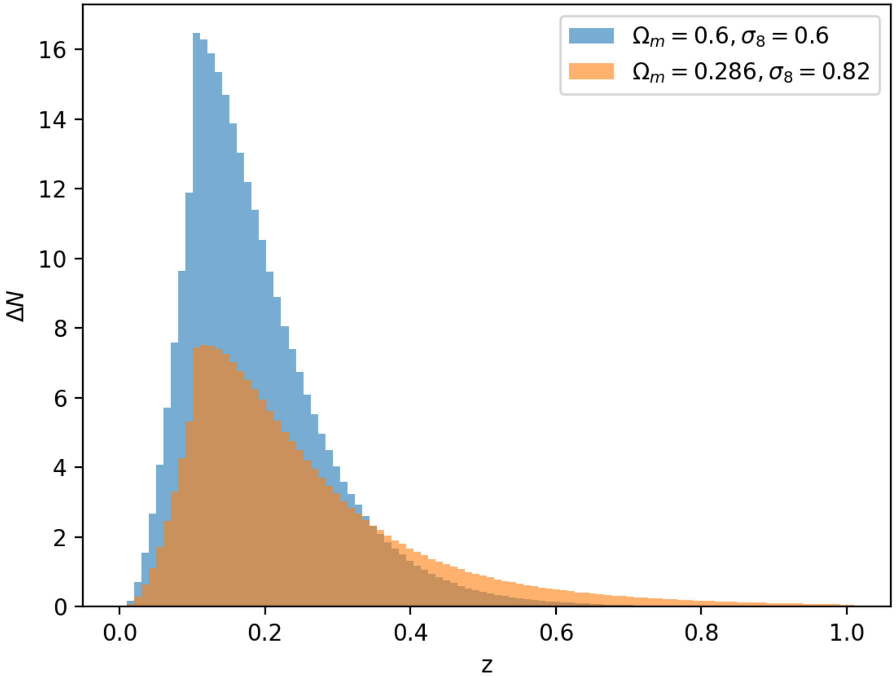
Distance metric #1

$$d_1 = \sum_{i=1}^{N_{\rm bin}} \left[\langle g_+^{\rm obs}\rangle(R_i) - \langle g_{+,i}^{\rm sim}\rangle(R_i|\boldsymbol{p})\right]^2$$



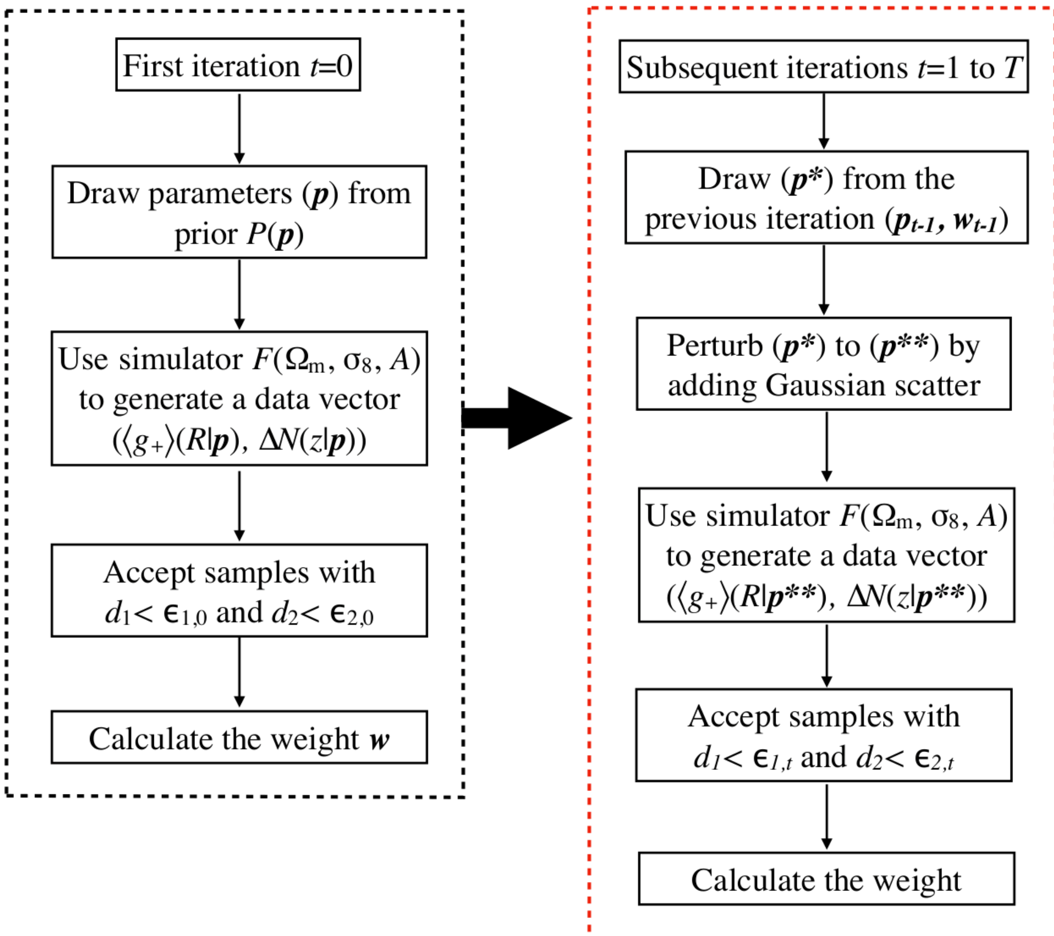
Distance metric #2

$$d_2 = \sum_{k=1}^{N_z} \left(\Delta N_k^{\rm obs} - \Delta N_k^{\rm sim}\right)^2$$



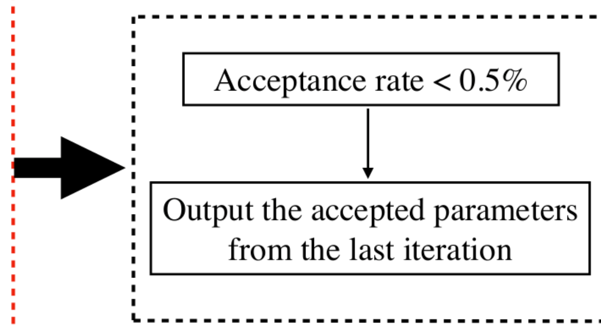
Tam, KU, Amara 22

Rejection based ABC-PMC



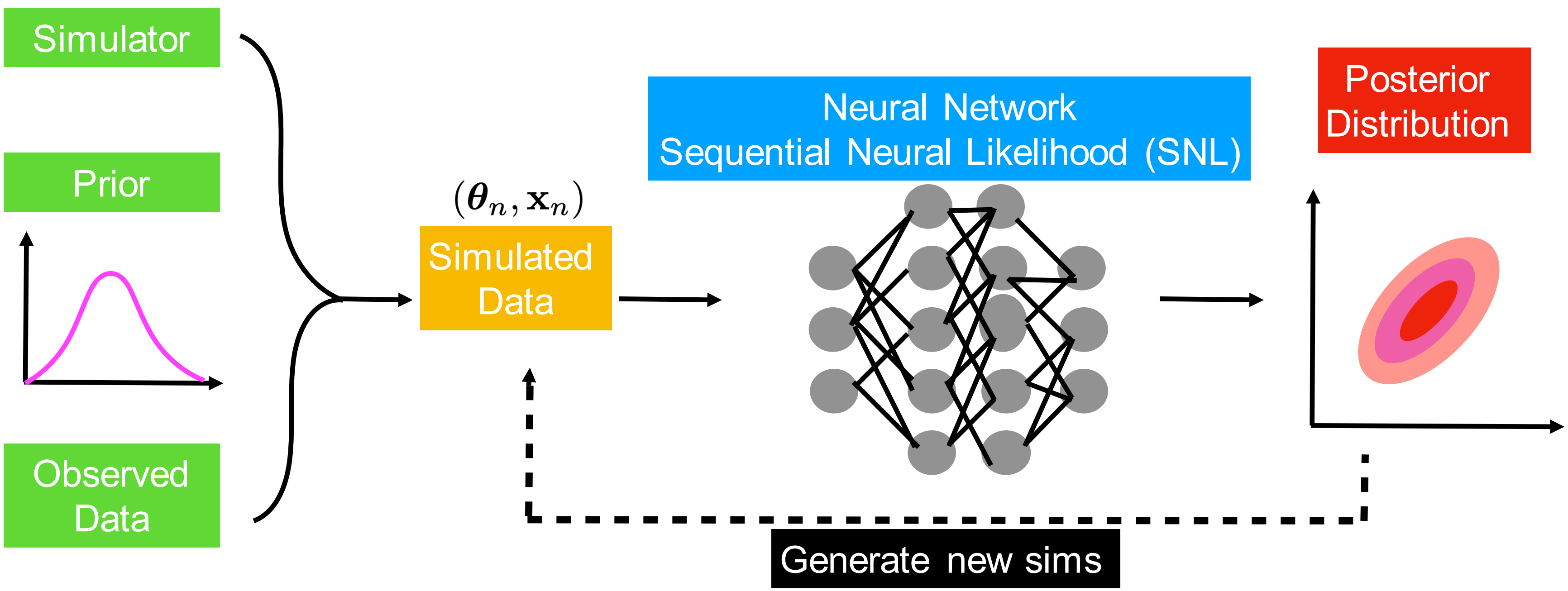
Limitation of ABC:
"Small acceptance rates and low efficiency"
As the iteration proceeds and ε → zero, more realizations are rejected

To have 1000 accepted samples, 1st iteration step with acceptance rate ~ 60%, generate N = 1000/0.6 ~ 1700 realizations; Last iteration step with acceptance rate of ~ 0.5%, generate N ~2 × 10^5 realizations.



Iteration until the stopping criterion is satisfied

PYDELFI
(Density-Estimation Likelihood-Free Inference)



$$p(\boldsymbol{\theta} \mid \mathbf{x}_o) \propto \boxed{p(\mathbf{x}_o \mid \boldsymbol{\theta})}\, p(\boldsymbol{\theta})$$

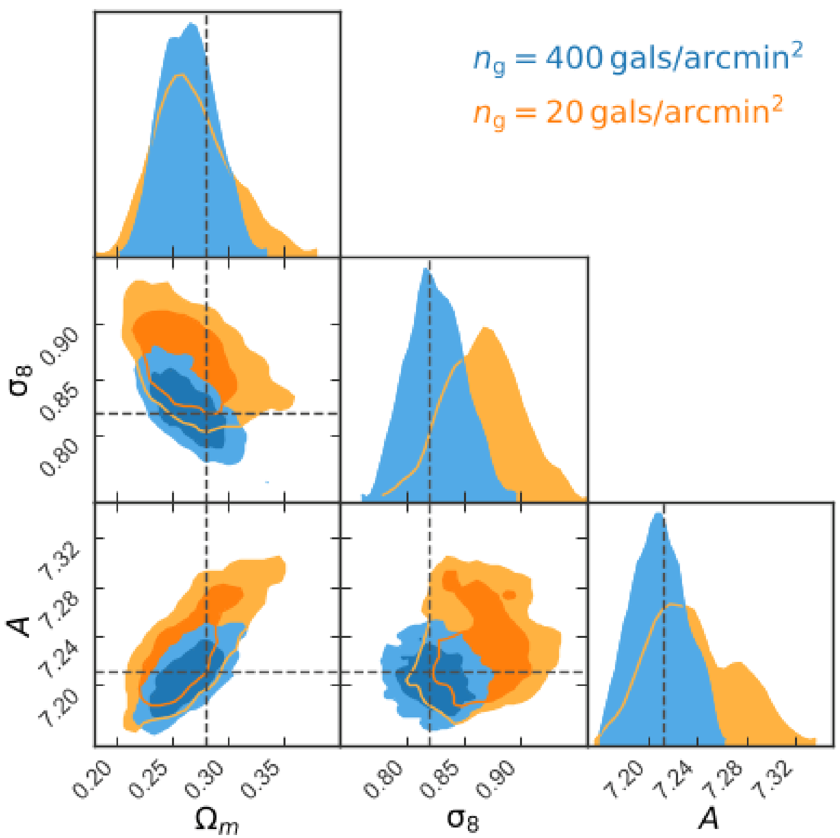
Learn the likelihood by training neural density estimators $q_{\boldsymbol{\phi}}(\mathbf{x} \mid \boldsymbol{\theta})$ on $(\boldsymbol{\theta}_n, \mathbf{x}_n)$

Cosmological inference with ABC-PMC and PYDELFI

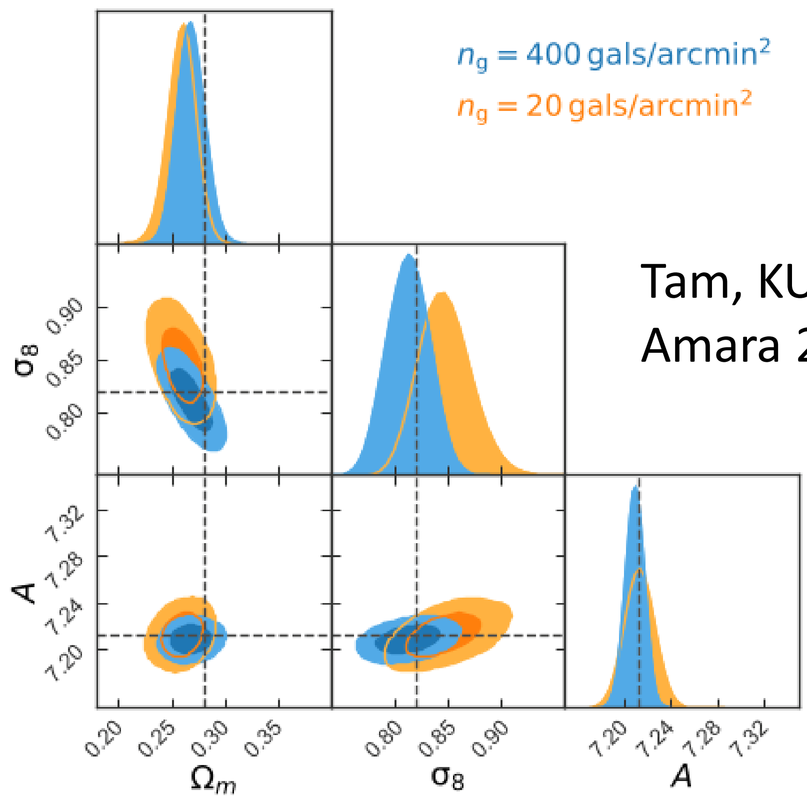
n_g = 400 gals/arcmin^2: ≈ nearly noise-free (cosmic-variance limited)

n_g = 20 gals/arcmin^2: deep ground-based survey

ABC: $O(10^6)$ simulations



DELFI: $O(10^5)$ simulations



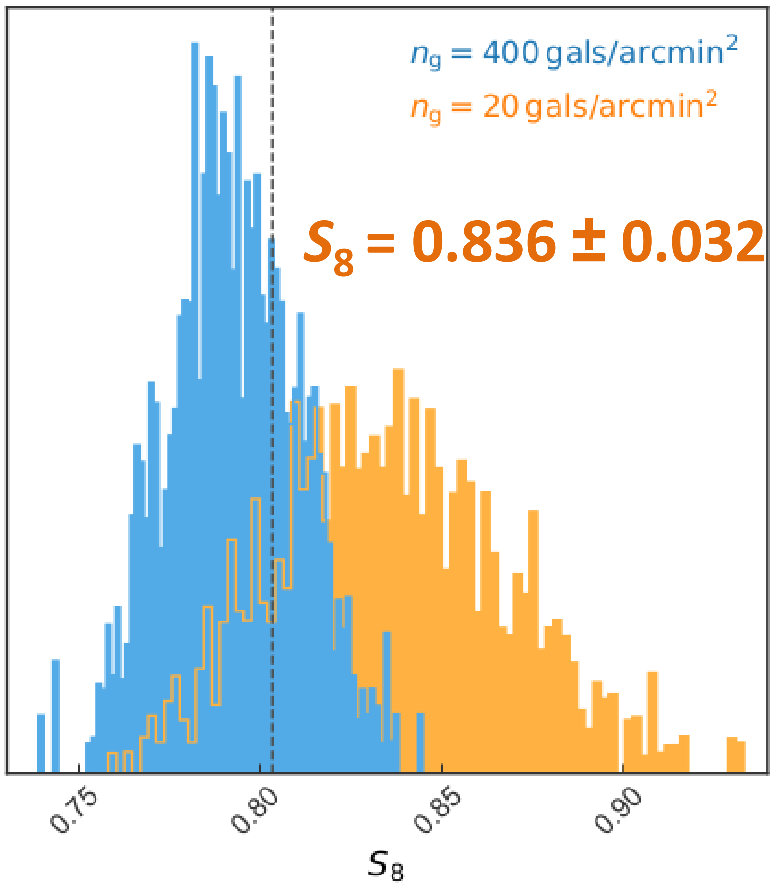
Tam, KU, Amara 22

Posterior constraints on $S_8 = \sigma_8(\Omega_m/0.3)^{0.3}$

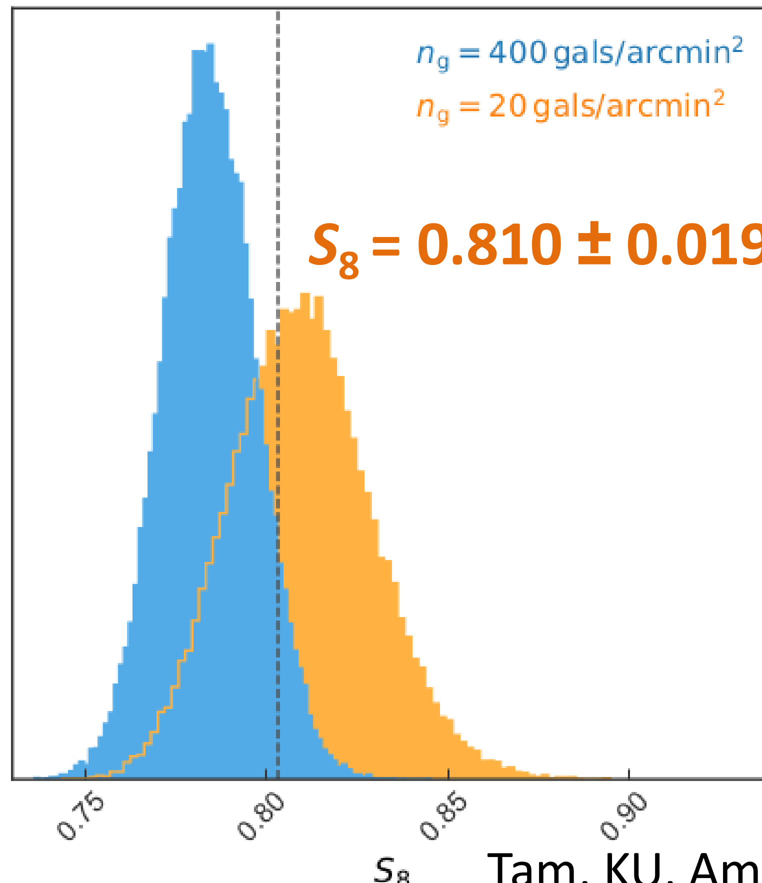
n_g = 400 gals/arcmin^2: ≈ nearly noise-free (cosmic-variance limited)

n_g = 20 gals/arcmin^2: deep ground-based survey

ABC: $O(10^6)$ simulations



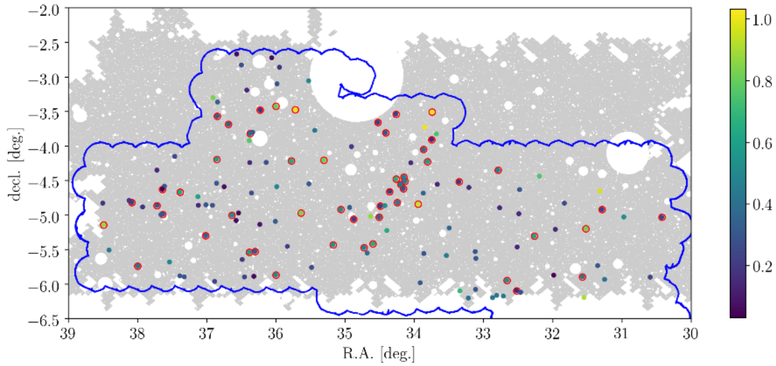
DELFI: $O(10^5)$ simulations

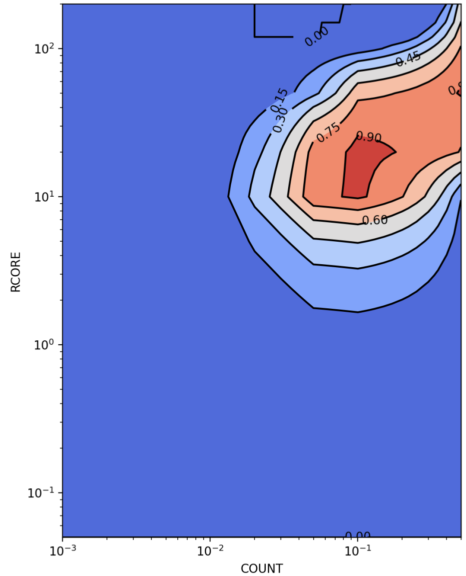


Tam, KU, Amara 22

Simulation-based inference with PYDELFI: Application to HSC-XXL [preliminary]

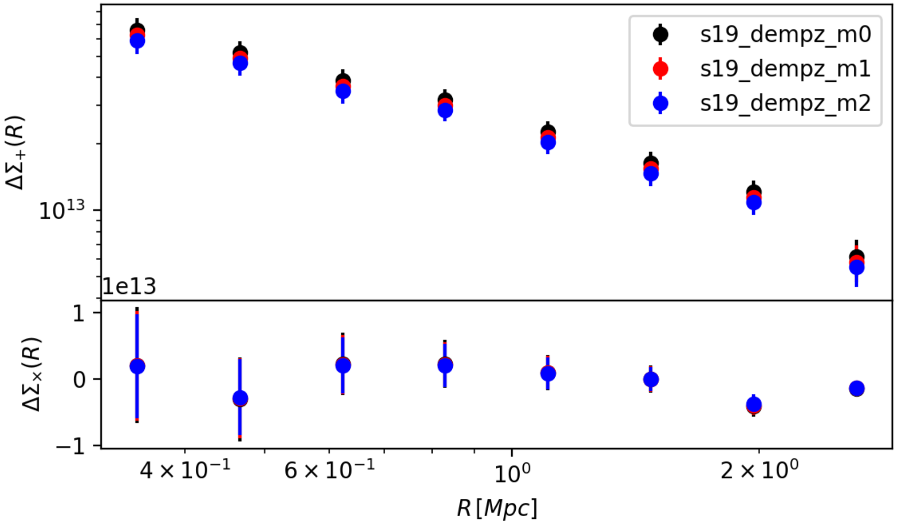
Subaru HSC-XMM field and XXL-N





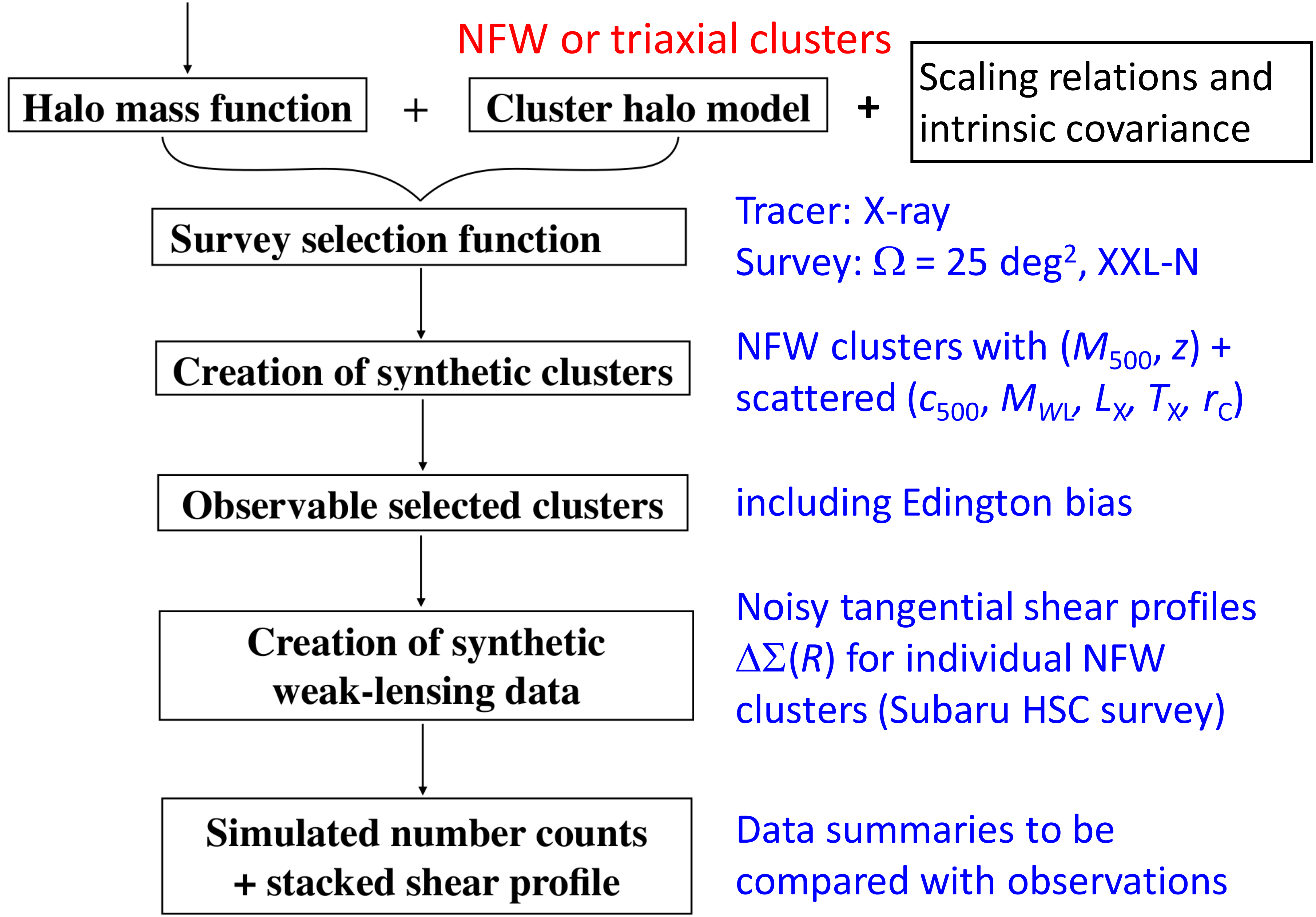
XXL DR2 catalog and DR2 selection function: N = 98 spec-confirmed C1 clusters in XXL-N (25 deg^2) from Garrel+22

Stacked tangential shear profiles $\Delta\Sigma(R)$ using 3 blinded HSC S19A shape catalogs



Tam, KU+ HSC-XXL (in prep.)

Schematics of forward simulator



Data summary vectors

$$\mathbb{D}_1 = \{ \langle g_+ \rangle (R_i) \}_i$$

$$\mathbb{D}_2 = \{ \Delta N (z_i) \}_i$$

$$\mathbb{D}_3 = \{ L_{X,i} \}_i$$

$$\mathbb{D}_4 = \{ T_{X,i} \}_i$$

e.g., 5 quantiles

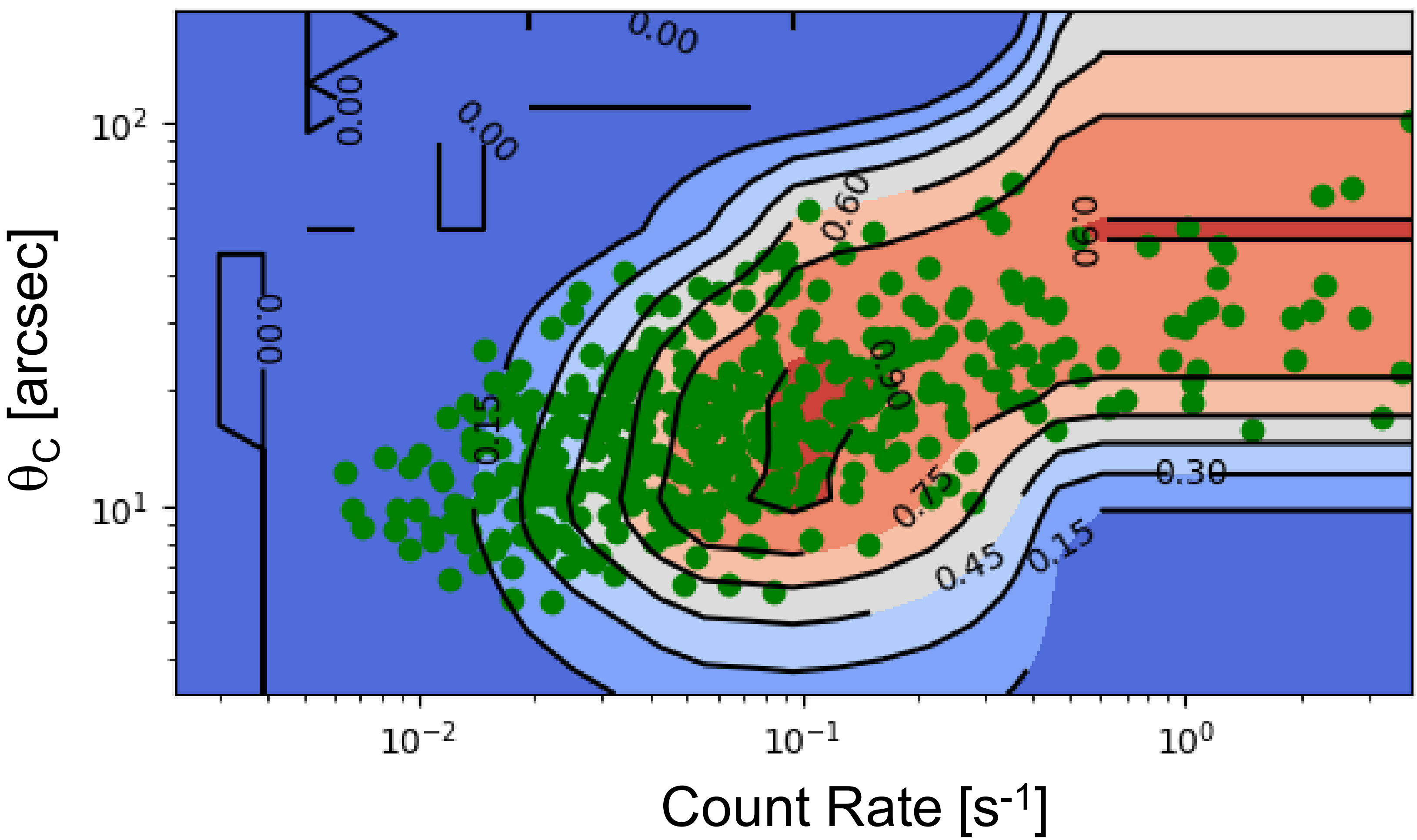
We may include more data summaries, if we are to float other scaling relations

e.g., $$\mathbb{D}_5 = \{ r_{c,i} \}_i$$

$$\vdots$$

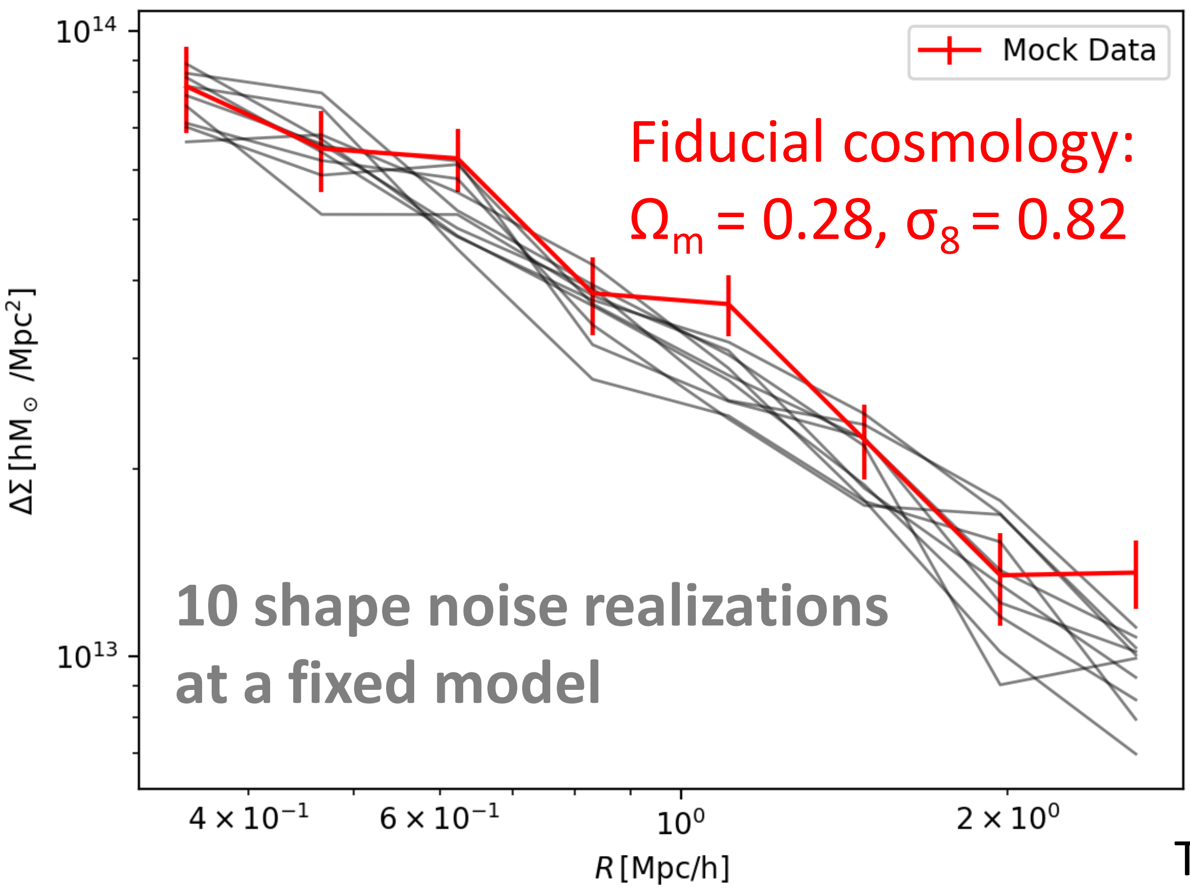
XXL selection function and simulated NFW clusters

● : realization of NFW clusters in CR-θ_C space before X-ray selection



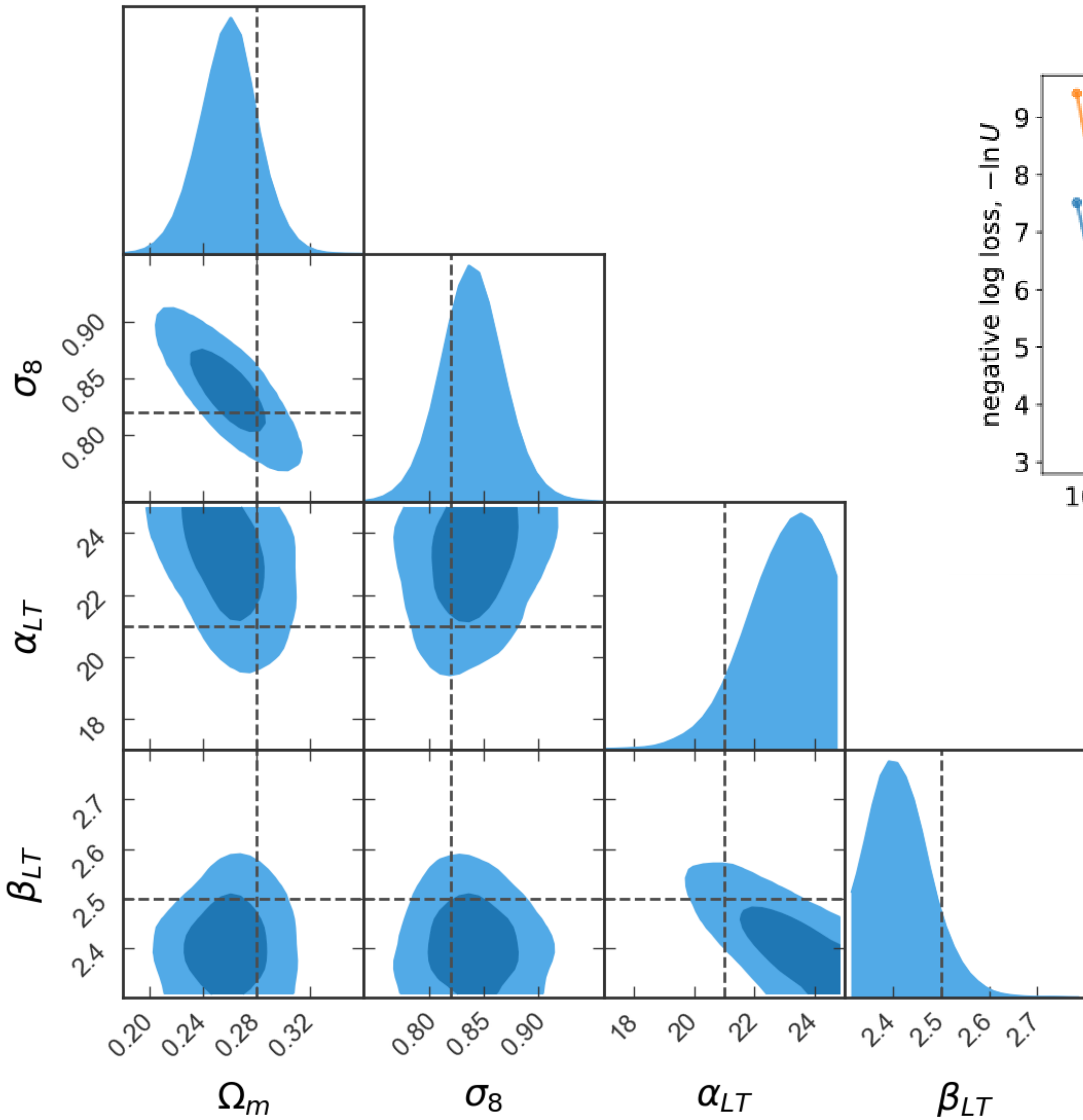
Stacked $\Delta\Sigma(R)$ profile from synthetic Subaru-HSC survey data

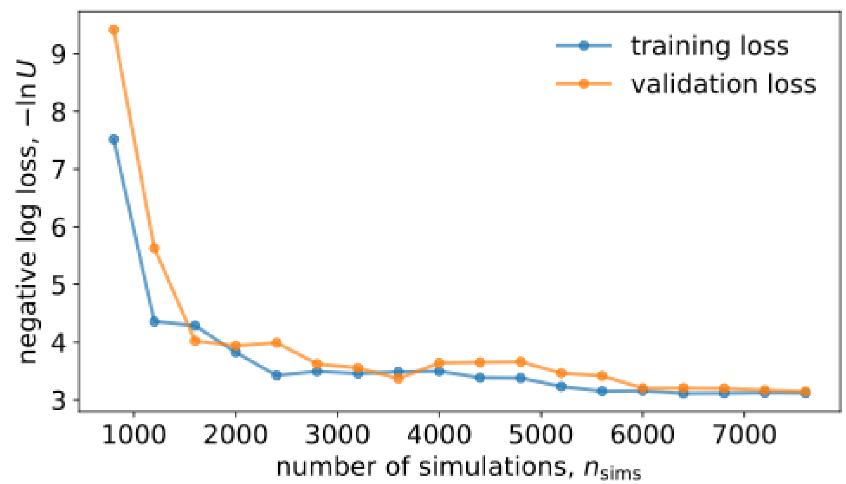
NFW clusters are assigned with $\ln(M_{WL}) = \ln(M_{true}) \pm 20\%$
In addition, shape noise is added to individual $\Delta\Sigma(R)$ profiles



Tam, Umetsu+ (in prep)

Simulation-based cosmological inference from synthetic XXL-HSC observations





Fiducial cosmology:
Ω_m=0.28, σ_8=0.82

Tam, Umetsu+ (in prep)

Summary

- Bayesian population modeling of multi-wavelength cluster surveys provides a flexible approach for cluster cosmology.
  - Two-step WL mass calibration in M_{WL} domain (Umetsu+20: XXL-HSC)
  - One-step WL mass calibration in $g_+(\mathrm{R})$ domain (Chiu+21: eFEDS-HSC)

- Recent results from joint X-ray + WL surveys show multi-variate mass scaling relations on group/cluster scales that are consistent with self-similar predictions (Umetsu+20; Sereno+20; Chiu+22).

- Likelihood-free cosmological inference using forward simulations (ABC, DELFI) will allow for even more flexible and accurate modeling of all physical, observational, and statistical effects (Tam, Umetsu, & Amara+22).