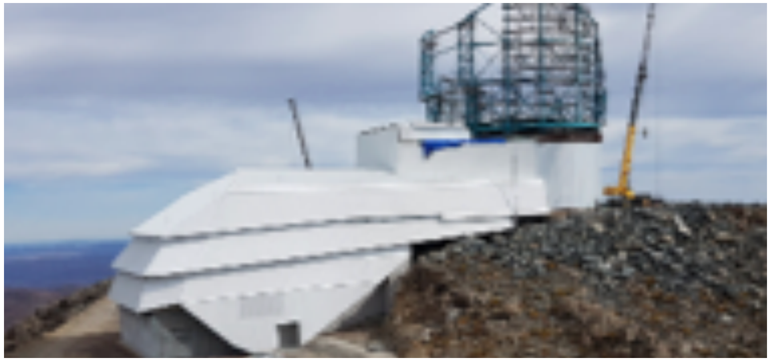
CosmoDC2: A Synthetic Sky Catalog for LSST

Eve Kovacs

In collaboration with:
Hal Finkel, Nick Frontiere, Salman Habib, Andrew Hearin, Katrin Heitmann, Joe Hollowed, Dan Korytov, Patricia Larsen, Nan Li, Adrian Pope, Steve Rangel and Andrew Benson and Yao-Yuan Mao and the LSST-DESC DC2 validation team

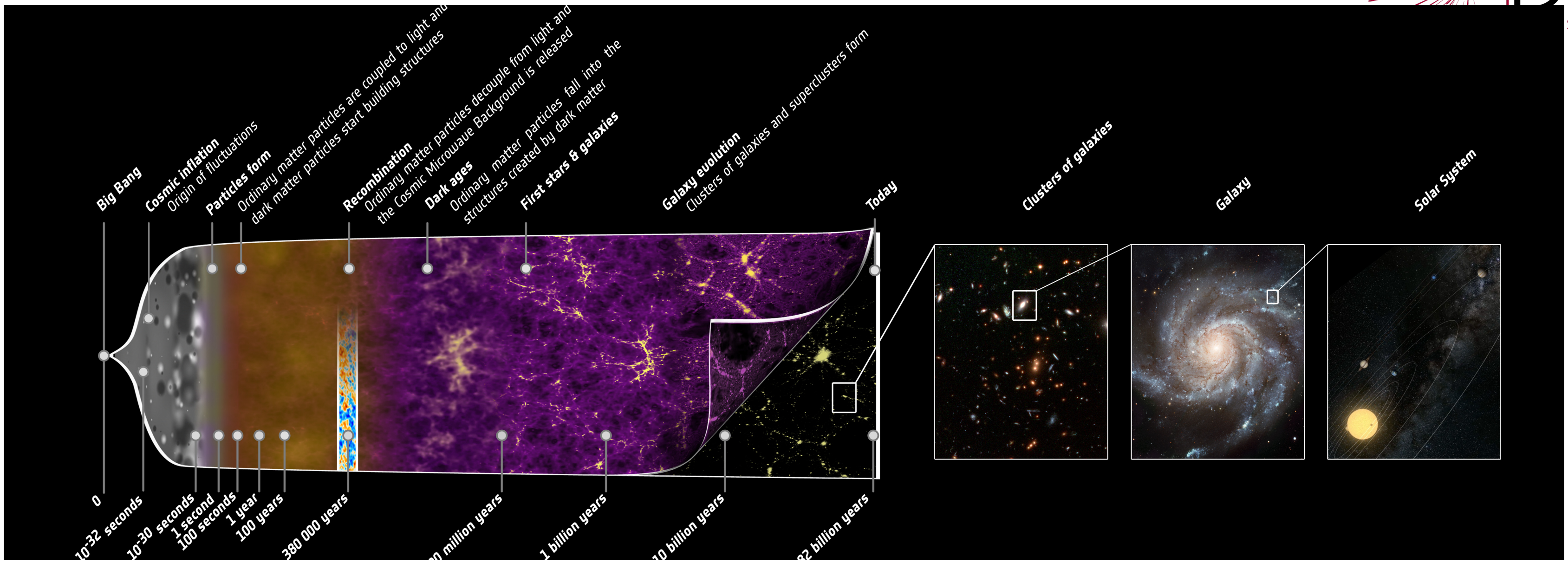
Outline

- Introduction
  - Modeling the Expanding Universe
  - The Era of Precision Cosmology
  - Challenges for Simulated Mock Skies
- Building Synthetic Skies
  - N-body Simulations
  - Modeling the Galaxy-Halo Connection
  - Building CosmoDC2 for LSST-DESC
  - Weak-Lensing Quantities
- Catalog Validation
  - Developing Validation Tests
  - DESCQA Validation Framework
  - CosmoDC2 Validation Tests



Large Synoptic Survey Telescope (LSST)

Forward Modeling the Expanding Universe



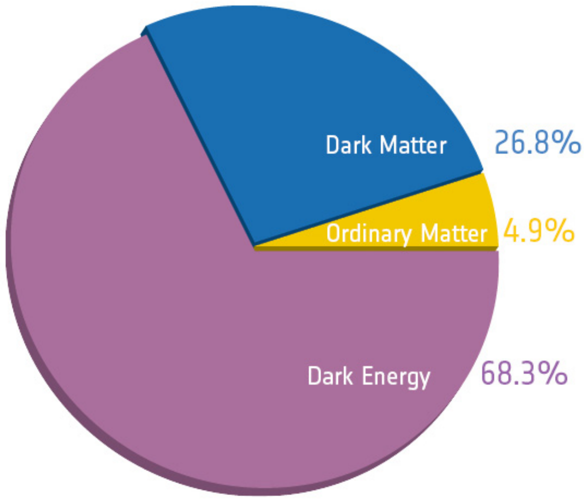
credit: www.nasa.gov

a=0, z=∞ ——————————————————————————> ***a=1, z=0***

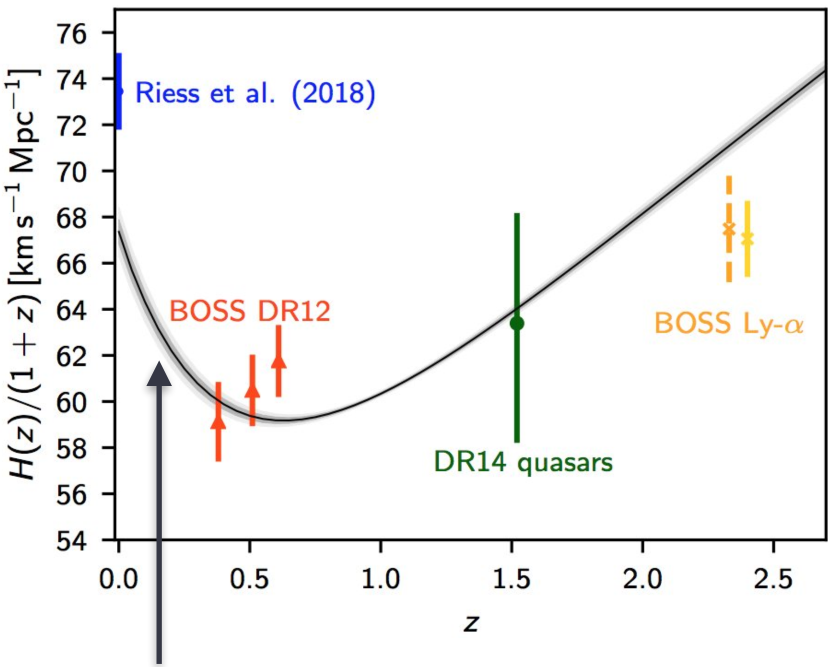
Observe galaxies and transient objects to infer model parameters

- Expansion history ⟶ $H = \dot{a}/a$
- Growth of structure ⟶ $\rho_i(a)$
- Simulations are crucial for understanding nonlinear growth, observational biases, effects of cosmic variance, …

Components of the Model



credit: planck.cf.ac.uk



credit: https://arxiv.org/abs/1807.06209

accelerated expansion for z < 0.5

ΛCDM Dark Energy (*k=0*)

$$w = P/\rho \text{ where } \rho \sim a^{-3(1+w)}$$

$$H^2(z) = (\dot{a}/a)^2 = H_0^2(\Omega_m a^{-3} + \Omega_\Lambda a^{-3(1+w)})$$

Is w constant and = -1?

The Era of Precision Cosmology

Stage 3 surveys are completing now

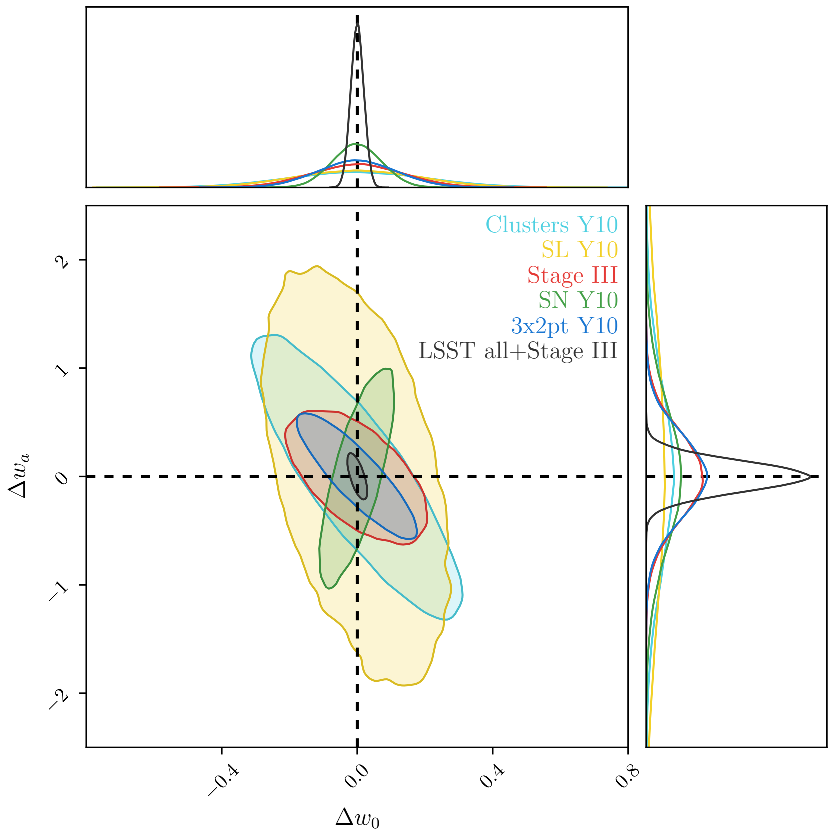
Upcoming stage 4 surveys are even bigger and have more demanding science goals

LSST will image

- 18,000 sq. deg. (1/2 sky)
- large depth: z ~ 3, *r* < *27*
- billions of galaxies
- tiny statistical errors
- FoM will improve from ~130 (Stage 3) to > 600
- **systematic errors will dominate**
- **need large volume, high-fidelity simulations to produce realistic mock skies**

LSST DESC Dark Energy Science Collaboration

$$w = w_0 + w_a(1 - a)$$



Credit: LSST DESC Science Requirements Document v1

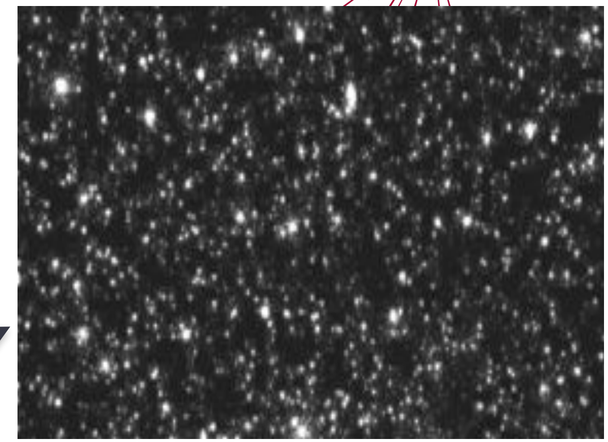
Challenges for Simulated Mock Skies

LSST DESC Dark Energy Science Collaboration

- Catalogs will be used in diverse ways
- Need to supply a wide range of realistically distributed properties

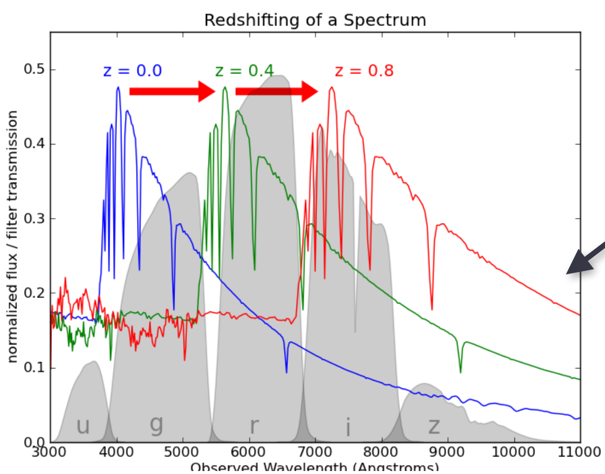
Pipeline Testing

Mock Galaxy Catalog



credit: LSST-DESC SSim Working Group

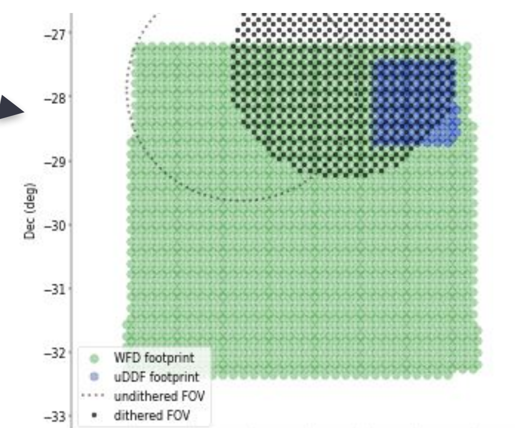
Image Simulations



credit: orgisel.github.io

Testing of Photometric Redshift Codes

Analysis Design and Systematics Calibration



credit:Humna Awan

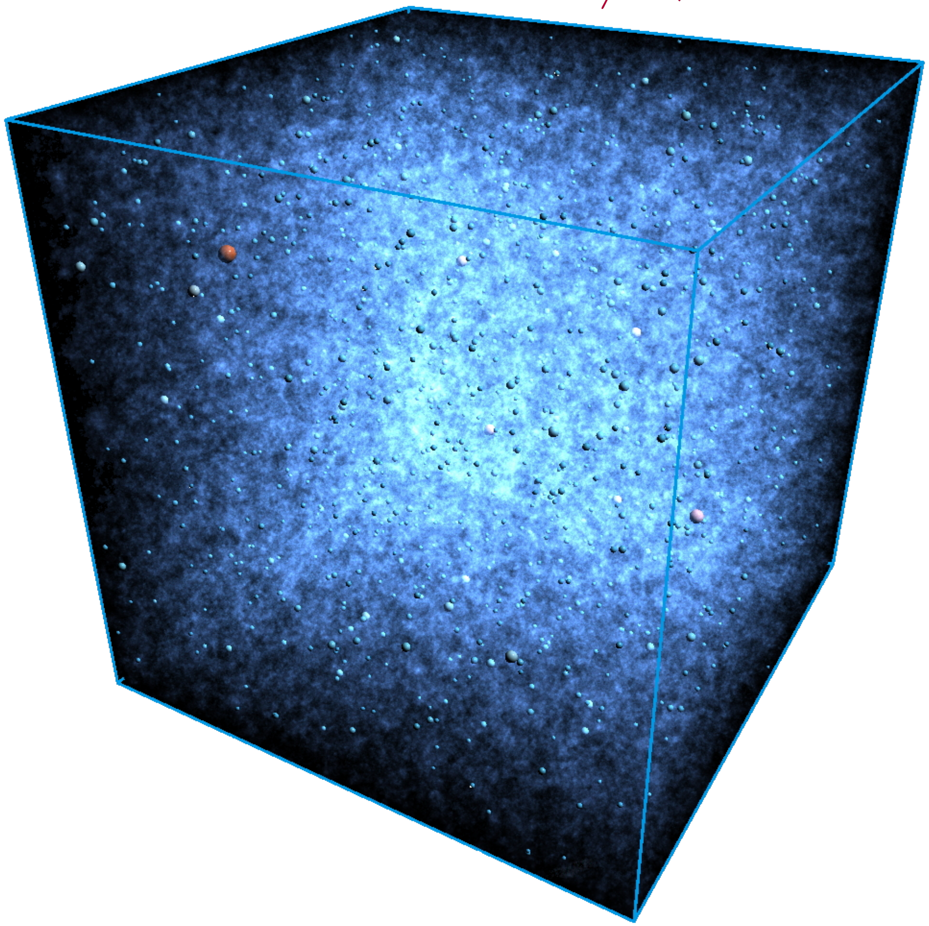
Survey Design

Building Synthetic Skies

CosmoDC2 catalog is based on the Argonne (ANL)

Outer-Rim gravity-only simulation

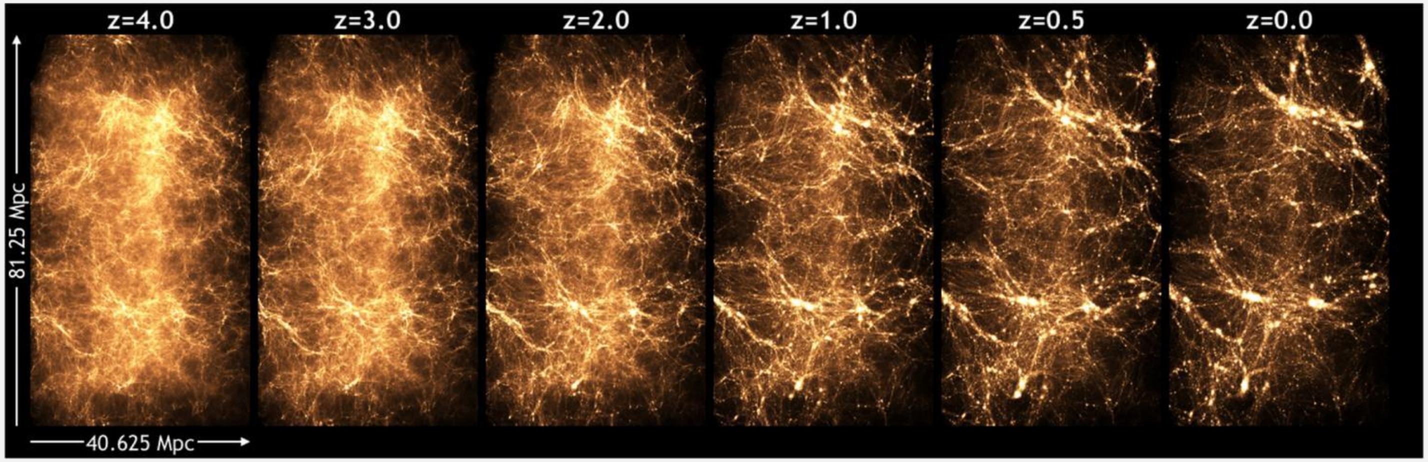
- 4.225 Gpc box
- 1 trillion tracer particles
- 2.6 x 10^9 $M_\odot$ particle mass
- WMAP-7 cosmology
- 5 PB data
- Hydro-dynamical simulations, which include gas physics, are too expensive to run at these large scales



Halos in the Outer Rim Simulation at z=0; spheres show halos with M > $1.8.10^{15}$ $M_\odot$ (Heitmann et al. (2019) image: Joe Insley and Silvio Rizzi)

N-body Simulations

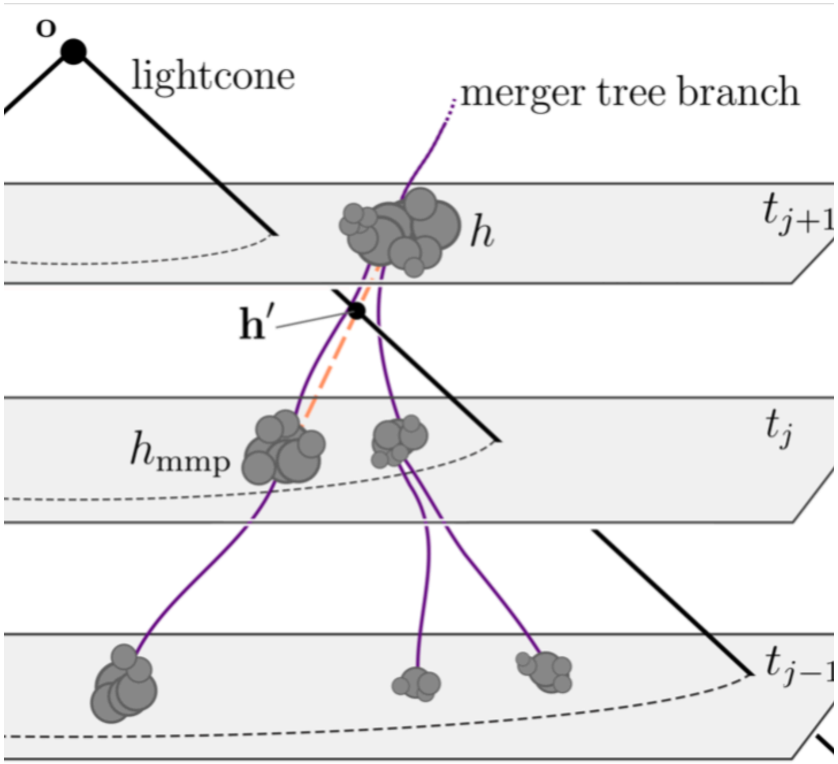
Q-Continuum
Heitmann et al. 2014



- 101 snapshots from z=10 to z=0, evenly spaced in log(*a*)
- Full particle outputs and 1% downsampled outputs are saved
- Halo catalog generated from FoF halo finder with $b = 0.168$, *> 20 particles*
- Halo merger trees generated with new particle-membership based algorithm which uses backwards-in-time processing to remove the effect of temporary mergers (resolves halo splits)
- Smaller companion simulation AlphaQ (360 Mpc^3) used for proto-typing and running the semi-analytic galaxy-formation model (SAM)

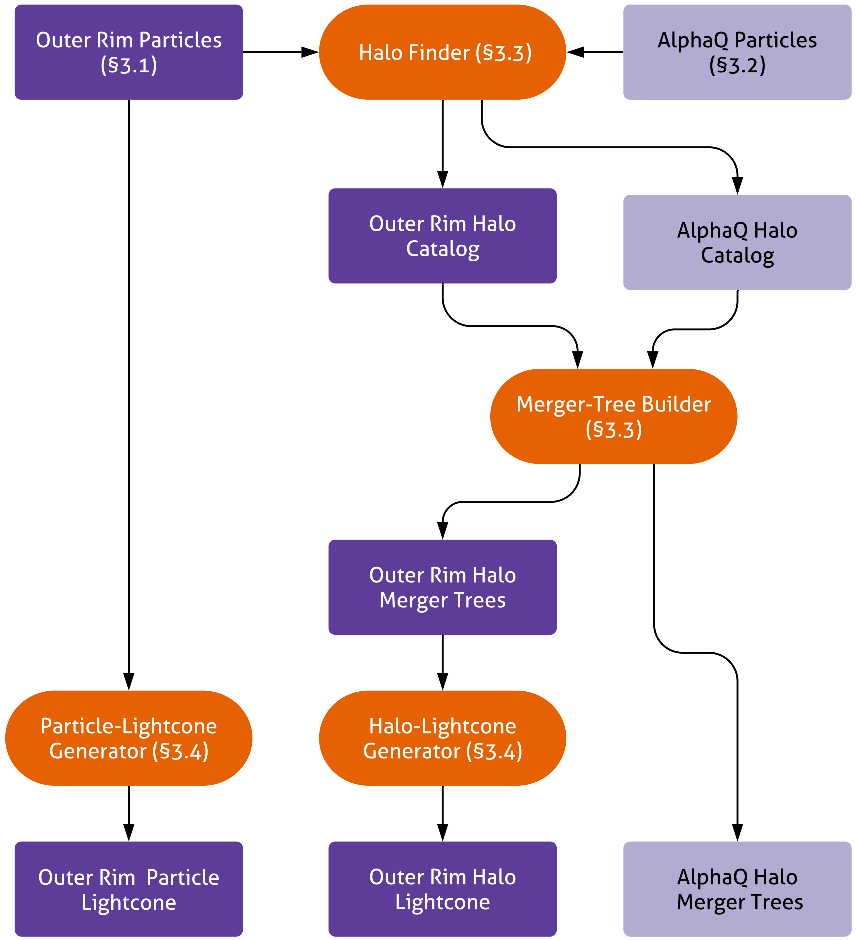
Lightcone Generation

- Build particle and halo light cones from snapshot particle and halo data
- Particle light cones are the basis for computing weak-lensing quantities
- Halo light cones are populated with galaxies to build the mock sky
- Linear interpolation between snapshots is sufficiently accurate
- z~3 corresponds to ~6 Gpc, so box replication with reflection prevents repeated structure along the los
- Merger trees are critical for building halo light cones - properties of h are used for h′



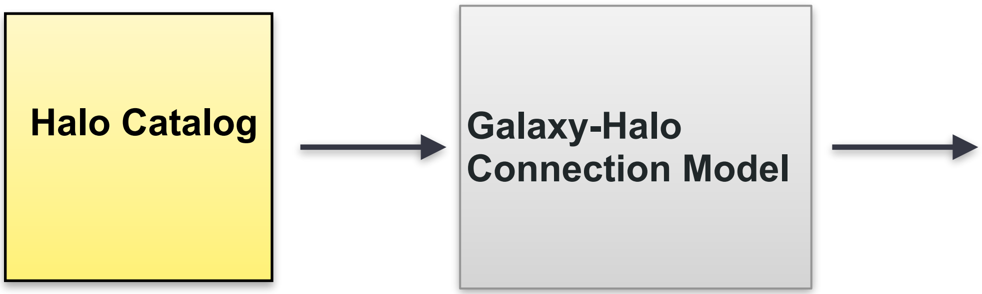
Summary of Simulation Product Workflow

- Outer Rim simulation products
  - Particle Lightcones
  - Halo Lightcones
- AlphaQ simulation products
  - AlphaQ halo merger trees

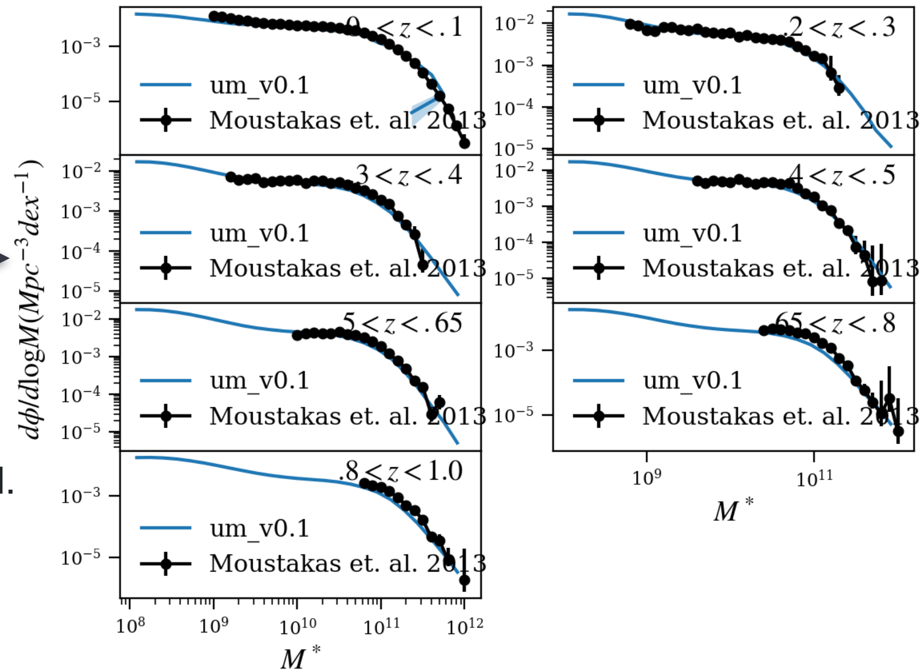


Modeling the Galaxy-Halo Connection

Empirical Models



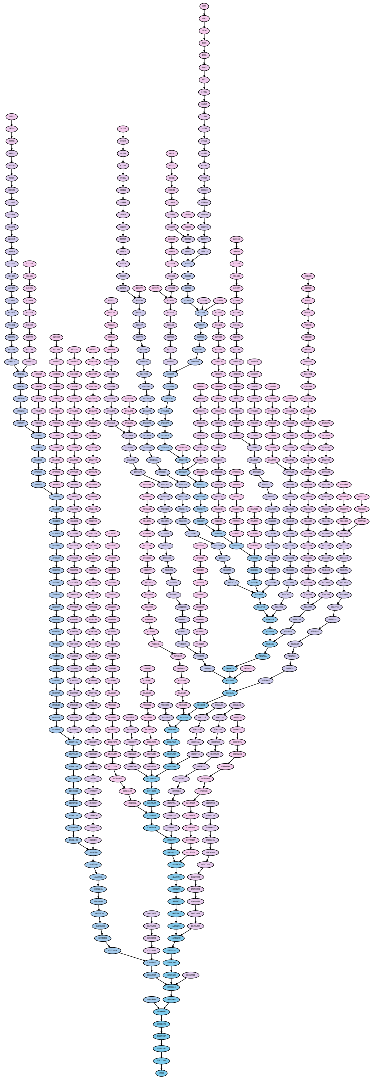
Models vary in complexity. See
 “Universe Machine”, Behroozhi et al.
https://arxiv.org/abs/1806.07893



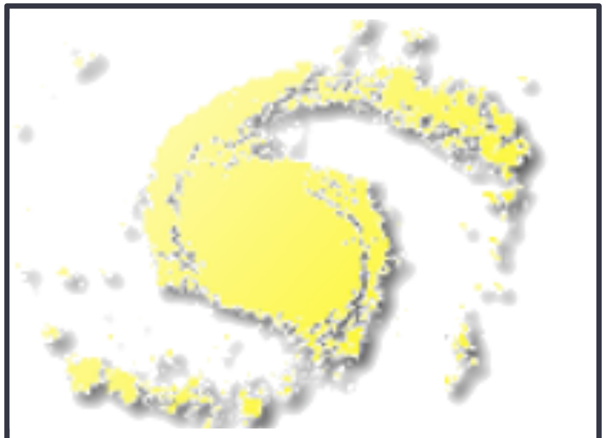
- Model the galaxy-halo connection and fit to empirical data to determine the model parameters. Models vary in complexity (HOD, SHAM, UM)
- Galaxy properties will be only *statistically* correct
- Limited set of properties can be determined, correlations between properties may be lost

Modeling the Galaxy-Halo Connection

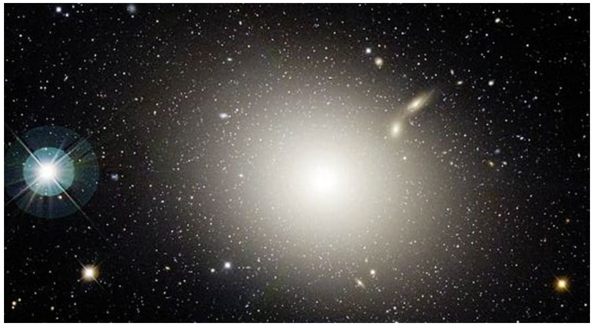
Semi-Analytic Models



Merger Tree
showing mass-assembly history of a selected halo



GALACTICUS
https://users.obs.carnegiescience.edu/abenson/



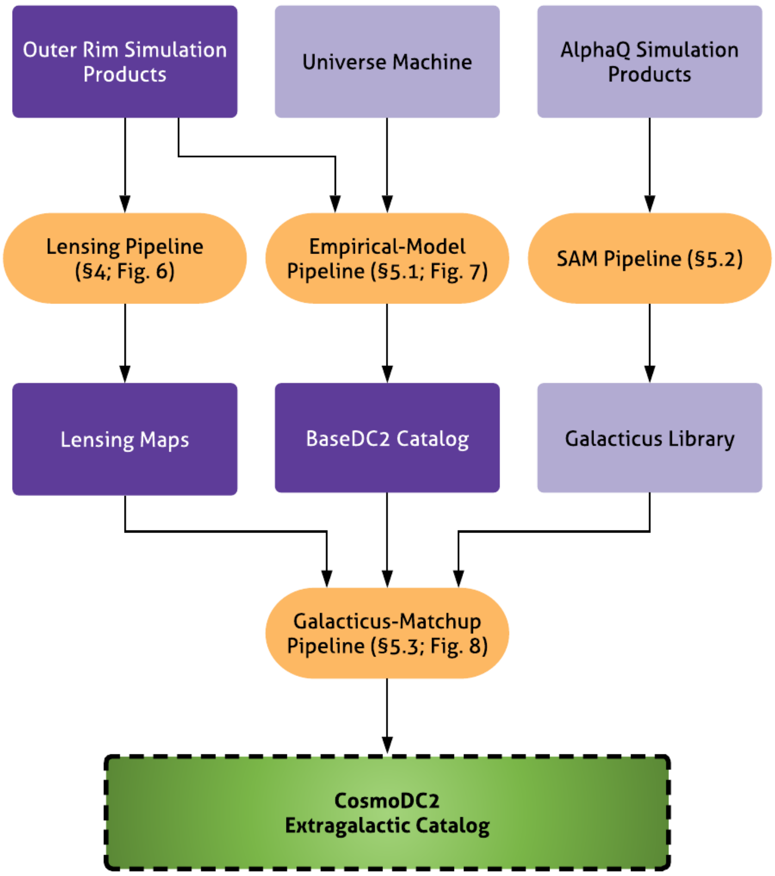
credit: apod.nasa.gov

credit: astronomy.com

- Solve coupled ODE's that track each galaxy's mass-assembly history and determine properties such as M*, luminosity, …
- Expensive to produce, requires many parameters that are difficult to tune

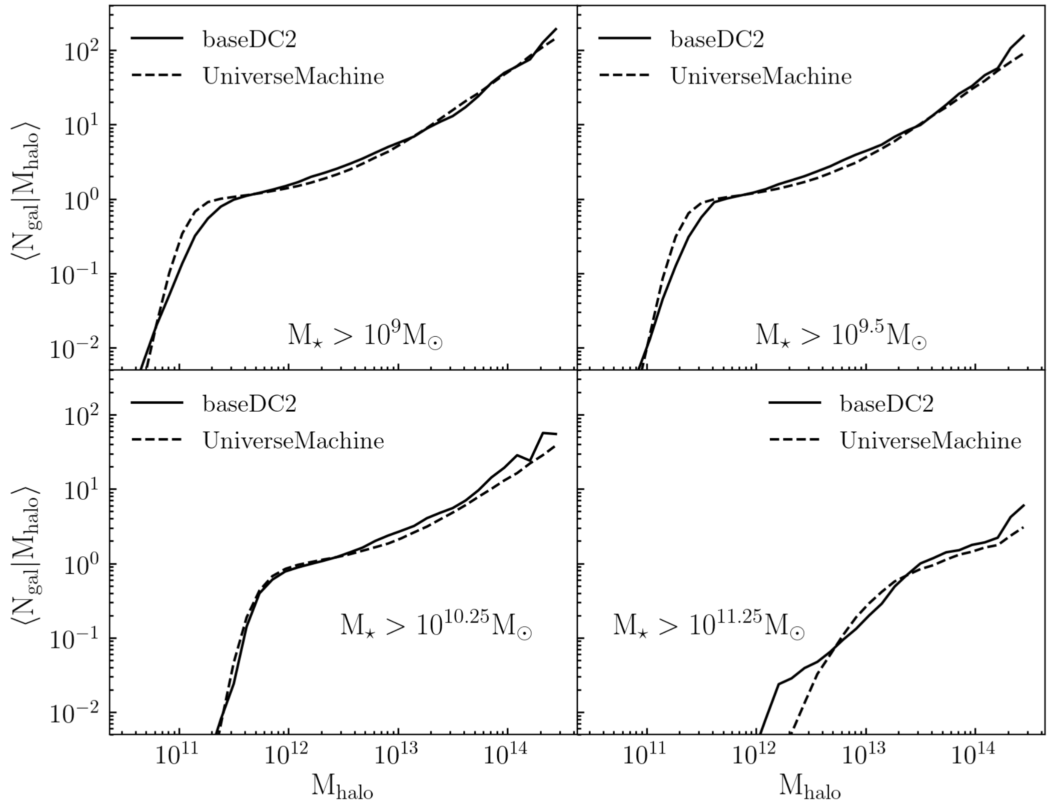
CosmoDC2 Production

- New hybrid method combines empirical and semi-analytic methods and resampling techniques to populate a large simulation with realistically complex galaxies
- Empirical model tuned on smaller simulation establishes the galaxy-halo connection and predicts limited set of properties. Galaxies are resampled into halos in the large simulation.
- Galaxy properties obtained from a small simulation using the Galacticus SAM are used as a library
- Expensive SAM modeling is run once without special tuning
- Galaxies are matched with those in the library to fill in the remainder of the required properties.
- Weak-lensing properties are added in the final amalgamation step



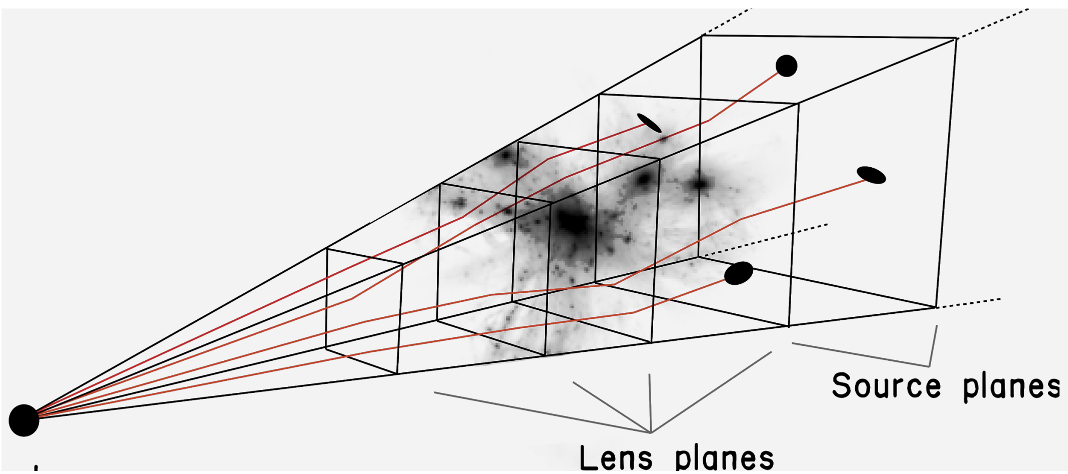
Empirical Model Pipeline

- UM is run on snapshots of the Millennium MultiDark Planck 2
- UM galaxies have realistic distributions of M* and SFR
- Match Outer Rim halos by mass to UM halos
- Resample galaxy content of UM halo into matching OR halo using Galsampler
- Model is augmented by adding flux modeling for rest-frame *gri* LSST filters
- Boost central M*, N_{sat} for missing UM high-mass halos
- Add ultra-faint galaxies to compensate for the mass resolution of the simulation



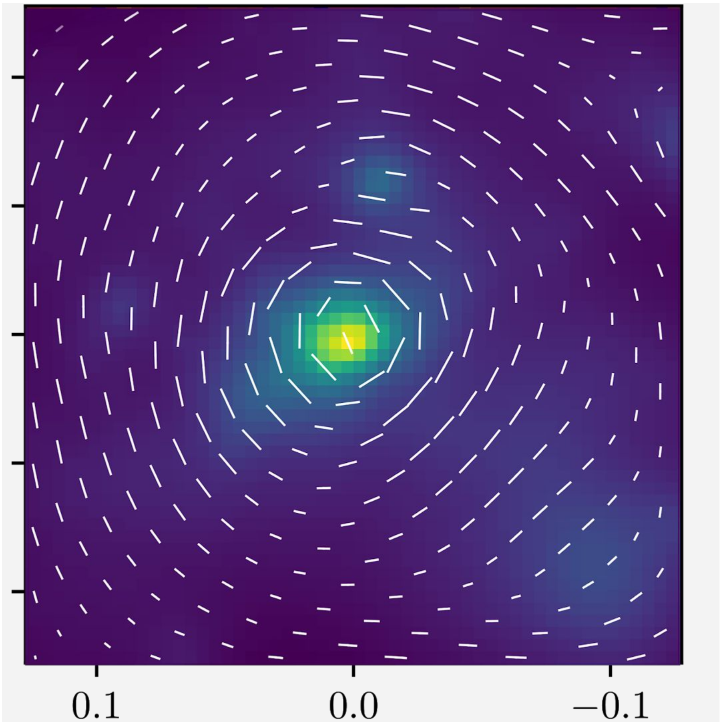
Weak-Lensing and Cosmic Shear

Images of distant galaxies are distorted due to weak gravitational lensing of emitted light by large-scale structure along the line of sight.



credit: Joe Hollowed

Ray-tracing from source to observer

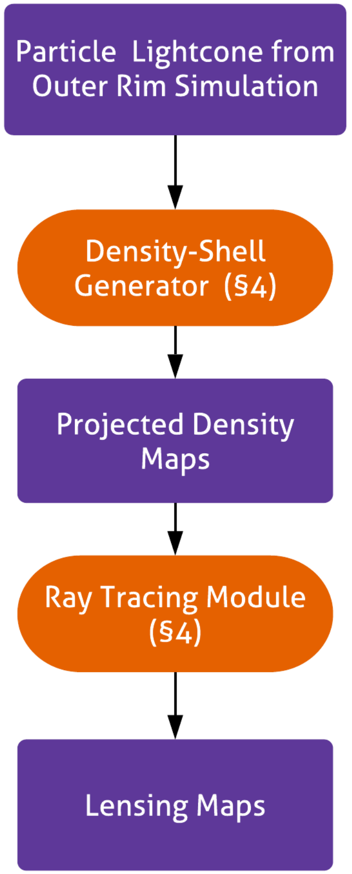


credit: Joe Hollowed

Map of **magnification** and **tangential shear** around a galaxy cluster

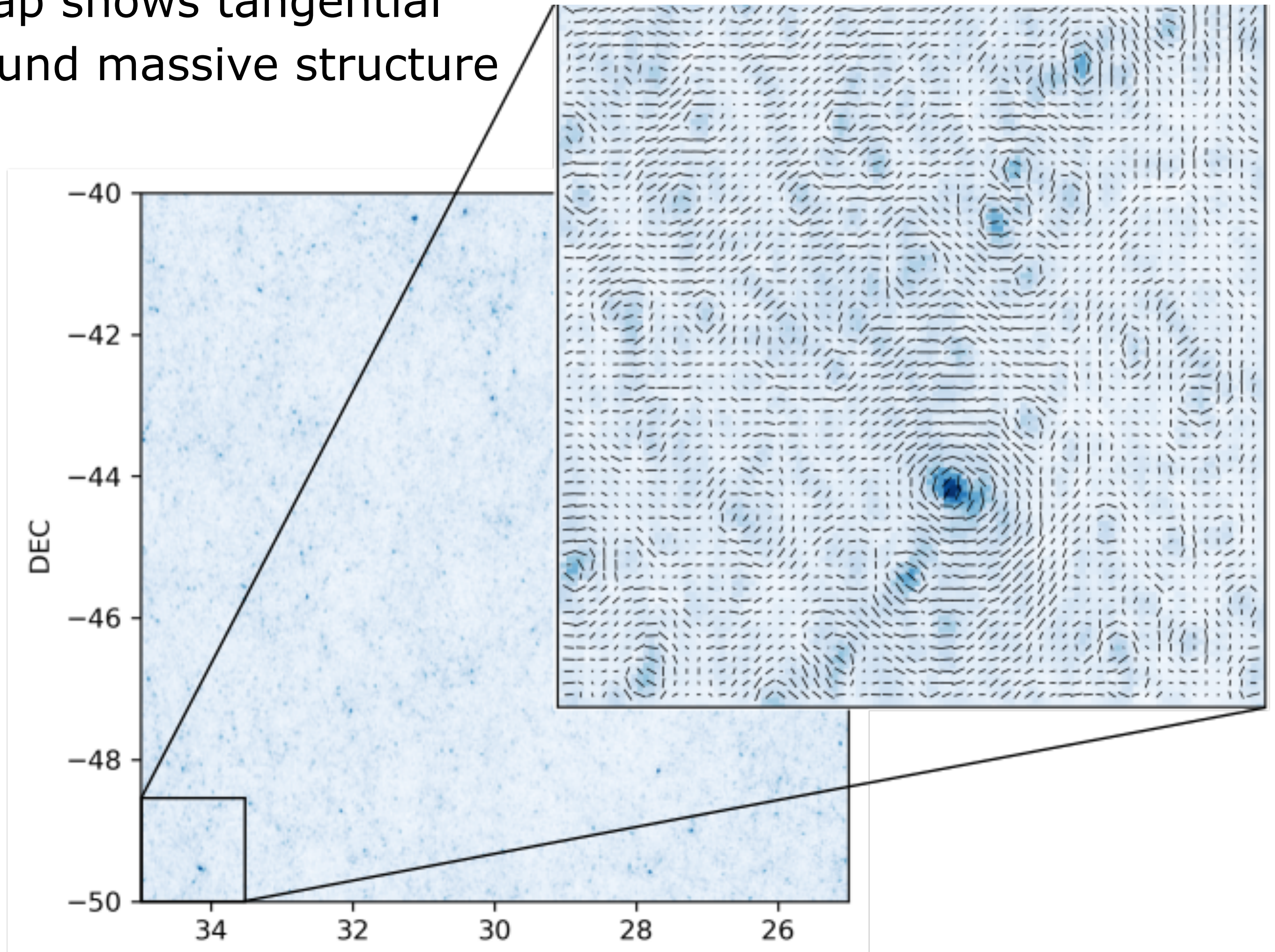
Weak-Lensing Pipeline

- Divide down-sampled particle lightcones into shells ~ 114 Mpc wide
- Estimate surface mass densities on a HEALPix grid with Nside=4096 (~arcmin)
- Conduct ray-tracing to compute deflections and produce lensing maps
- Galaxies are then shifted to their observed positions and assigned lensing quantities by interpolation of the resulting source shell.



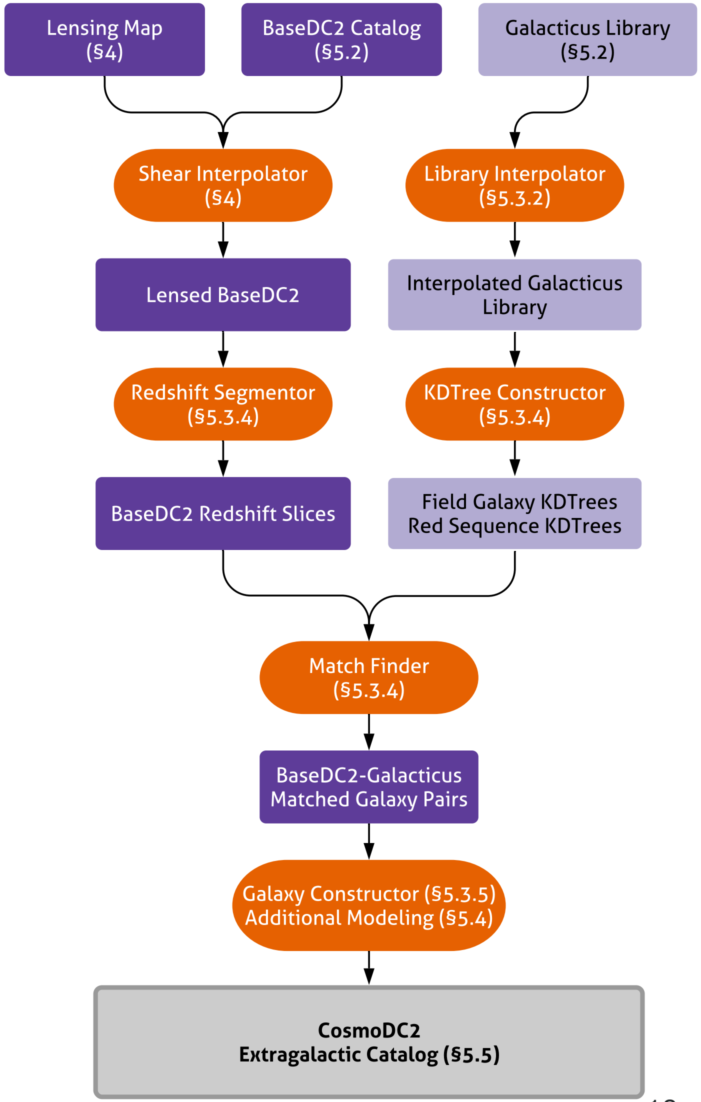
Weak Lensing Shear Map

- Density map shows 100 deg.^{2} patch, z~1
- Overlaid shear map shows tangential shear aligned around massive structure



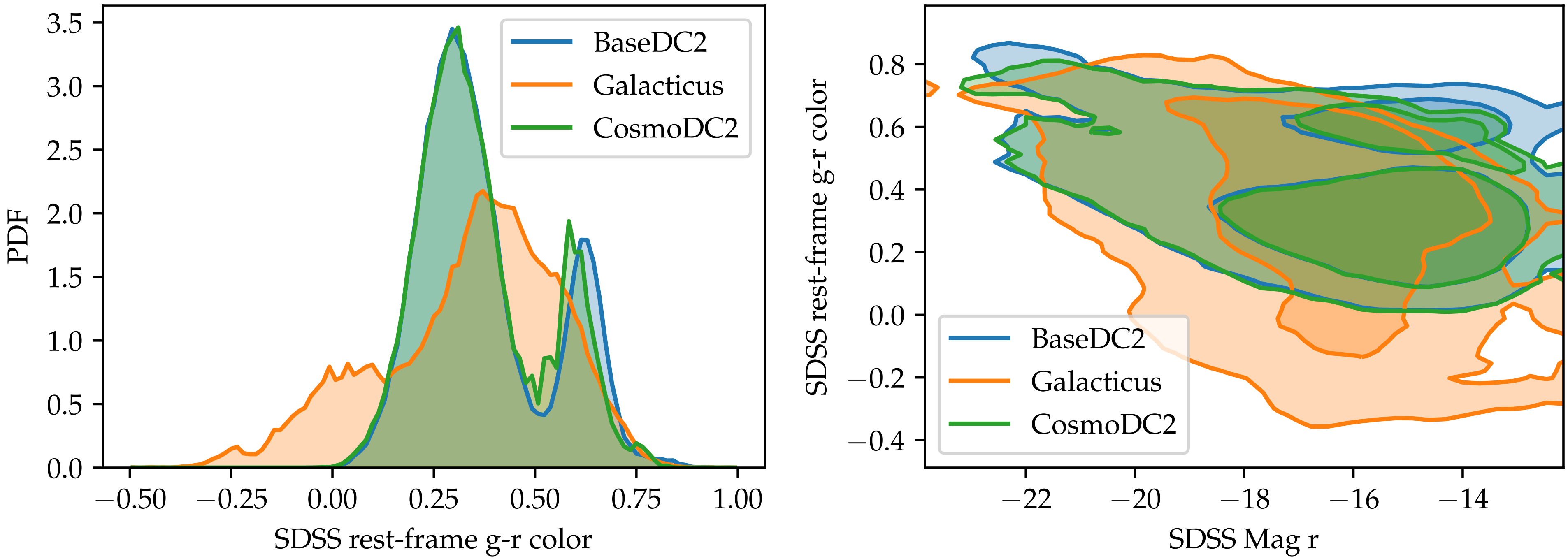
Galacticus Match-up Pipeline

- Incorporate weak-lensing information
- Preserve the careful tuning of the empirical model, but augment the properties with additional information from the SAM library
- Match the empirical *r*-band magnitudes and *g-r* and *r-i* colors with those of Galacticus library galaxies
- Properties of library galaxies must be interpolated between snapshots to minimize discreteness effects
- Library matches to cluster members on the red-sequence must also match the observed DES red sequence
- Additional empirical modeling based on more complex properties is included for galaxy shapes and sizes



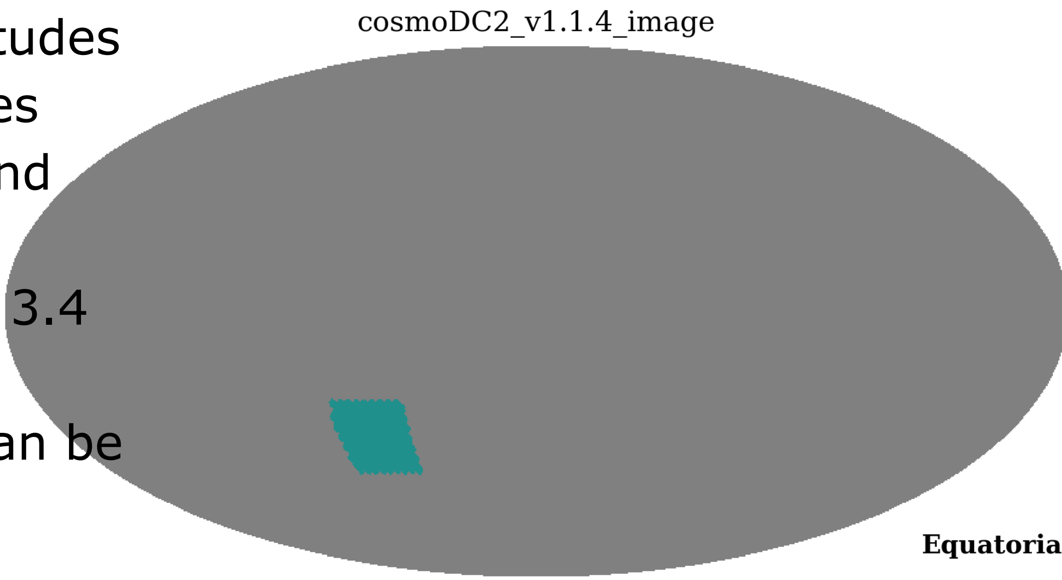
Matchup Distributions

- Matchup procedure compensates for “holes” in the color-magnitude space by down-weighting the luminosity in favor of color



CosmoDC2 Catalog for LSST DESC Data Challenge 2 (DC2)

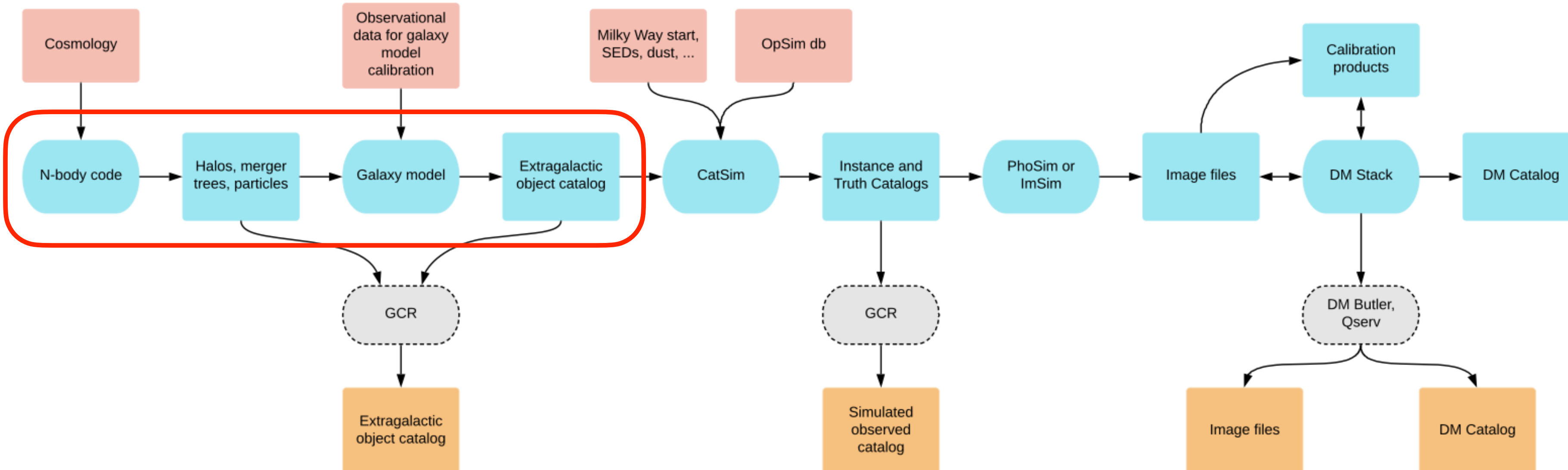
- 440 sq. deg
- 0 < z <3
- Complete to r-band magnitude 28
- Wide range of galaxy properties
  - Stellar mass
  - LSST *ugrizY* rest and observer-frame magnitudes
  - Galaxy shapes and sizes
  - Weak-lensing shears and magnifications
- Delivered in HEALPix format ~ 3.4 sq. deg. each
- Full description of properties can be found in the [CosmoDC2 paper]



Catalog Validation

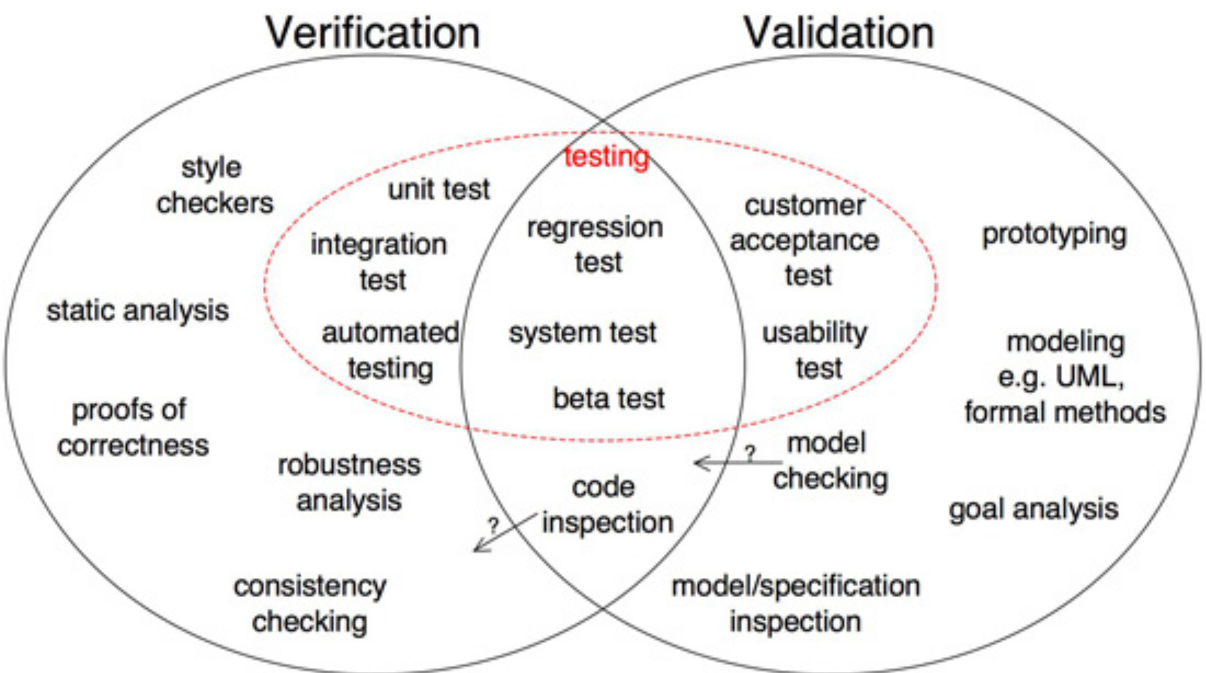
- CosmoDC2 production was accompanied by an extensive validation effort that ran alongside the production pipeline
- Aims of the validation process
- Designing validation and verification tests
- Curation of Validation Data
- Overview of Validation framework: DESCQA
- CosmoDC2 validation and verification tests

Why Validate?



- The DC2 end-to-end pipeline requires a huge investment of:
  - computer time (~200 M hours)
  - human resources (30+ people, + 10+ for N-body sims)
- Ensure that we are correctly capturing the physics we want to include
- Ensure that the catalogs are adequate for the intended science use cases
- Ensure that we are providing correct information for each pipeline component

Verification versus Validation



Credit:https://i1.wp.com/www.easterbrook.ca/steve/wp-content/VandVtoolbox.jpg

- Verification: Are we building the product correctly?
- Validation: Are we building the correct product?
- Both kinds of tests were performed on cosmoDC2

Designing Validation Tests

- What are the science use-cases for the catalog?
  - Eg. provide inputs for image simulations
  - Eg. test photo-z algorithms
  - Eg. test de-blending algorithms
  - Eg. study LSS and cluster systematics
- What galaxy properties must be rendered "correctly"?
  - Eg. galaxy number densities (#/sq. deg.) as a function of redshift, magnitude in *ugrizy*, colors, ...
- What does "correct" mean?
  - What level of agreement between the catalog and the observational data is needed to meet the science objectives

Suite of validation tests give quantitative measures that allow us to:

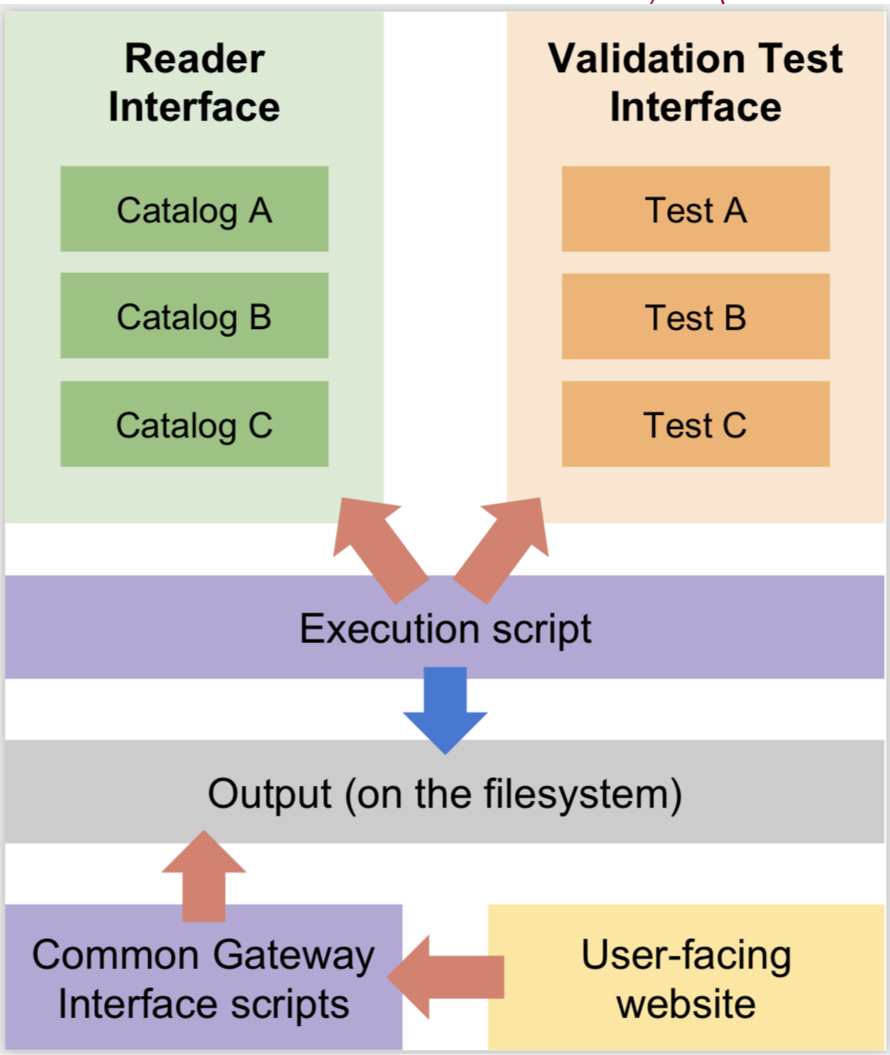
- Assess the level of realism provided by the catalog
- Determine how the catalog can be used

Curating the Validation Data

- What data sets are relevant?
  - Take stock of available data
- What is the best data set to use? Can data sets be combined?
  - Eg. number density tests need number counts from wide and deep surveys (deep for faint objects; wide to reduce effects of cosmic variance)
- What selections have been made on the data?
  - Eg. magnitude/flux selections, redshift ranges, etc.
- Can we replicate the selections in simulated data?
  - Implement the cuts directly, if possible
  - Find a proxy to mock up the selection effect

Overview of Validation Framework DESCQA

- Master script runs selected catalogs against selected tests
- Catalogs are "homogenized" by using the GCR (Generic Catalog Reader)
- Tests are "homogenized" by using the BaseValidationClass in DESCQA
- Tests produce summary statistics and comparison plots
- Web interface displays results "on demand"
- GCRCatlogs repository
- DESCQA repository
- DESCQA paper



DESCQA Summary Results

	cosmoDC2_v1.1.4_small	skysim5000_v1.0.0_small
ApparentMagFuncTest_HSCg	NOT QUITE 0.494	NOT QUITE 0.535
ApparentMagFuncTest_HSCi	PASSED 0.17	PASSED 0.338
ApparentMagFuncTest_HSCr	PASSED 0.342	NOT QUITE 0.411

- Summary page displays matrix of tests, catalogs and scores (if available)
- Each test produces plots of catalog and validation data and delivers a status (pass, ~pass, inspect) and an optional score
- Most tests produce tables of the plotted data
- Some tests produce summary plots comparing all catalogs in the run with the validation data

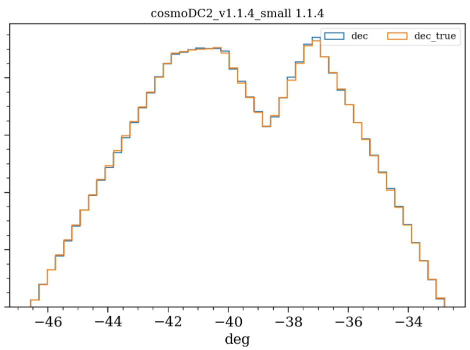
CosmoDC2 Verification Tests

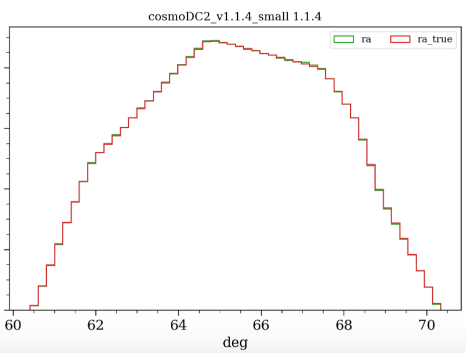
Readiness tests

cosmoDC2_v1.1.4_small VALIDATION_TEST_FAILED

Outputs:

- SUMMARY.html
- p00_deg.png
- p01_deg.png
- p02_redshift.png
- p03_ellipticity.png
- p04_ellipticity_total.png
- p05_size.png
- p06_size_bulge.png
- p07_velocity.png
- p08_shear.png
- p09_position_angle.png
- p10_convergence.png
- p11_magnification.png
- p12_sersic_disk.png
- p13_sersic_bulge.png
- p14_mag.png
- p15_Mag.png
- p16_stellar_mass.png
- p17_SED.png
- p18_Disk SED.png
- p19_Bulge SED.png
- p20_A_v.png
- p21_A_v_disk.png
- p22_A_v_bulge.png
- p23_R_v.png
- p24_R_v_disk.png
- p25_R_v_bulge.png
- p26_bulge_to_total_ratio.png
- traceback.log





- Running readiness test on cosmoDC2_v1.1.4_small 1.1.4
- Found 292781506 entries in this catalog.
- shear_2_treecorr, shear_2 seem be to identical!
- mag_y, mag_y_lsst, mag_Y_lsst seem be to identical!
- mag_y_lsst_no_host_extinction, mag_Y_lsst_no_host_extinction seem be to identical!
- mag_g_lsst, mag_g seem be to identical!
- mag_i_lsst, mag_i seem be to identical!
- mag_r_lsst, mag_r seem be to identical!
- mag_true_Y_lsst, mag_true_y_lsst, mag_true_y seem be to identical!
- mag_true_Y_lsst_no_host_extinction, mag_true_y_lsst_no_host_extinction seem be to identical!
- mag_true_g, mag_true_g_lsst seem be to identical!
- mag_true_i_lsst, mag_true_i seem be to identical!
- mag_true_r, mag_true_r_lsst seem be to identical!
- mag_true_u_lsst, mag_true_u seem be to identical!
- mag_true_z_lsst, mag_true_z seem be to identical!
- mag_u_lsst, mag_u seem be to identical!
- mag_z, mag_z_lsst seem be to identical!
- Mag_true_y_lsst_z0, Mag_true_Y_lsst_z0 seem be to identical!
- Mag_true_y_lsst_z0_no_host_extinction, Mag_true_Y_lsst_z0_no_host_extinction seem be to identical!
- It is true that `galaxy_id < 1e11`
- It is true that `size_minor_bulge_true <= size_bulge_true`
- It is true that `size_minor_disk_true <= size_disk_true`
- It is true that `size_minor_true <= size_true`
- It is true that `(size_bulge_true != 0) | (stellar_mass_bulge == 0)`
- It is true that `stellar_mass_bulge <= stellar_mass`
- It is true that `stellar_mass_disk <= stellar_mass`
- It is true that `1.0 / magnification ~== (1.0 - convergence)**2.0 - shear_1**2.0 - shear_2**2.0`
- It is true that `size_minor_true / size_true ~== (1.0 - ellipticity_true) / (1.0 + ellipticity_true)`
- It is true that `size_minor_disk_true / size_disk_true ~== (1.0 - ellipticity_disk_true) / (1.0 + ellipticity_disk_true)`
- It is true that `size_minor_bulge_true ~== size_bulge_true * (1.0 - ellipticity_bulge_true) / (1.0 + ellipticity_bulge_true)`
- It is true that `ellipticity_1_true ** 2.0 + ellipticity_2_true ** 2.0 ~== ellipticity_true ** 2.0`
- It is true that `ellipticity_1_disk_true ** 2.0 + ellipticity_2_disk_true ** 2.0 ~== ellipticity_disk_true ** 2.0`
- It is true that `ellipticity_1_bulge_true ** 2.0 + ellipticity_2_bulge_true ** 2.0 ~== ellipticity_bulge_true ** 2.0`
- galaxy_id is all unique
- halo_id[is_central] is all unique

Quantity	min	max	median	mean	std	f_inf	f_nan	f_zero	f_outlier
dec	-46.59	-32.81	-39.67	-39.55	3.061	0	0	0	0
dec_true	-46.57	-32.8	-39.68	-39.56	3.063	0	0	0	0
ra	60.41	70.34	65.38	65.36	2.188	0	0	0	0
ra_true	60.41	70.33	65.39	65.37	2.19	0	0	0	0
redshift	0.000889	3.117	1.914	1.871	0.7097	0	0	0	0
redshift_true	0.001107	3.036	1.915	1.871	0.7097	0	0	0	0
ellipticity_1_bulge_true	-0.7824	0.7938	1.528e-06	-3.008e-07	0.04866	0	0	0	0.1157
ellipticity_1_disk_true	-0.9477	0.948	1.186e-05	6.64e-06	0.2268	0	0	0	0.01927
ellipticity_1_true	-0.9464	0.9472	4.055e-06	1.146e-05	0.1571	0	0	0	0.09654
ellipticity_2_bulge_true	-0.7619	0.7706	-1.109e-07	3.485e-06	0.04867	0	0	0	0.1157
ellipticity_2_disk_true	-0.9484	0.9471	-8.692e-07	2.752e-05	0.2268	0	0	0	0.01929
ellipticity_2_true	-0.9471	0.944	-2.965e-07	1.572e-05	0.1571	0	0	0	0.09655
ellipticity_bulge_true	1.625e-07	0.7982	0.01681	0.03898	0.05671	0	0	0	0.07598
ellipticity_disk_true	2.12e-06	0.95	0.2135	0.2527	0.1976	0	0	0	0.0006328
ellipticity_true	1.186e-06	0.9499	0.07701	0.1461	0.1674	0	0	0	0.0404
size_bulge_true [log]	-2.639	1.415	-1.439	-1.412	0.2613	0	0	0	0.01535
size_disk_true [log]	-2.54	1.639	-1.116	-1.098	0.2838	0	0	0	0.008806
size_minor_bulge_true [log]	-2.831	1.312	-1.469	-1.447	0.2587	0	0	0	0.01372
size_minor_disk_true [log]	-3.584	1.382	-1.334	-1.337	0.3499	0	0	0	0.006508
size_minor_true [log]	-3.519	1.382	-1.411	-1.382	0.3184	0	0	0	0.006811
size_true [log]	-2.635	1.494	-1.267	-1.247	0.3297	0	0	0	0.002791
size_bulge_true	0.002296	26.03	0.03643	0.04825	0.05073	0	0	0	0.05961

Assembling Validation Tests and Criteria

LSST DESC Dark Energy Science Collaboration

- LSST-DESC Analysis Working Groups helped to develop validation tests for ensuring the catalog meets the scientific requirements for their analyses
- Validation criteria specify the required level of agreement between the catalog and the observational data
- Strong collaborative effort between Cosmological Simulations group and analysis working groups.

Test	WGs	Implemented	Validation Data	Criteria	"Eyeball" Check by WG	Issue
p(e)	WL	✓ @evevkovacs	✓ COSMOS	✓	✓ (WL @rmjarvis)	#14 #81
p(position angle)	WL	✓ @msimet	✓ uniform	✓	✓ (WL @msimet)	#76 #82
size distribution	WL	✓ @msimet	✓ COSMOS	✓	✓ (WL @msimet)	#77 #80
size-luminosity	WL	SOON @vvinuv	✓ vdW+14, COSMOS	✓	✓ (WL @msimet)	#13 #56
shear 2pt corr.	WL	✓ @patricialarsen	✓ camb	✓	✓ (WL @patricialarsen)	#35 #54
N(z)	LSS	✓ @evevkovacs	✓ DEEP2	✓	✓ (PZ @sschmidt23)	#11 #107
dN/dmag	LSS	✓ @duncandc	✓ HSC	✓	?	#7 #47
red sequence colors	CL	SOON @j-dr	✓ DES Y1	✓	✓ (CL @erykoff)	#41 #101
CLF	CL	SOON @chto	✓ SDSS	?	?	#9 #102
galaxy-galaxy corr	WL, LSS, TJP	✓ @vvinuv @morriscb	✓ SDSS, DEEP2	×	✓ (LSS @slosar)	#10 #38

Example of GitHub issues that were used to track the status of tests

Extragalactic Catalog Validation Targets

- LSST-DESC WGs identified a number of critical galaxy properties that required validation
- There was considerable *overlap* between the requirements of different groups
- Critical Properties:
  - ❖ Galaxy Number Densities
  - ❖ Galaxy 2pt Correlation Functions
    - ➢ Projected and angular gg
    - ➢ gg, $\gamma\gamma$ and gγ (3x2pt)
  - ❖ Emission-Line Galaxies (for LSS)
  - ❖ Galaxy Shapes and Sizes
  - ❖ Galaxies in Cluster Environments
    - ➢ Well-defined red-sequence (RS) galaxies
    - ➢ cluster richness and velocity dispersions
  - ❖ Stellar-Mass Function

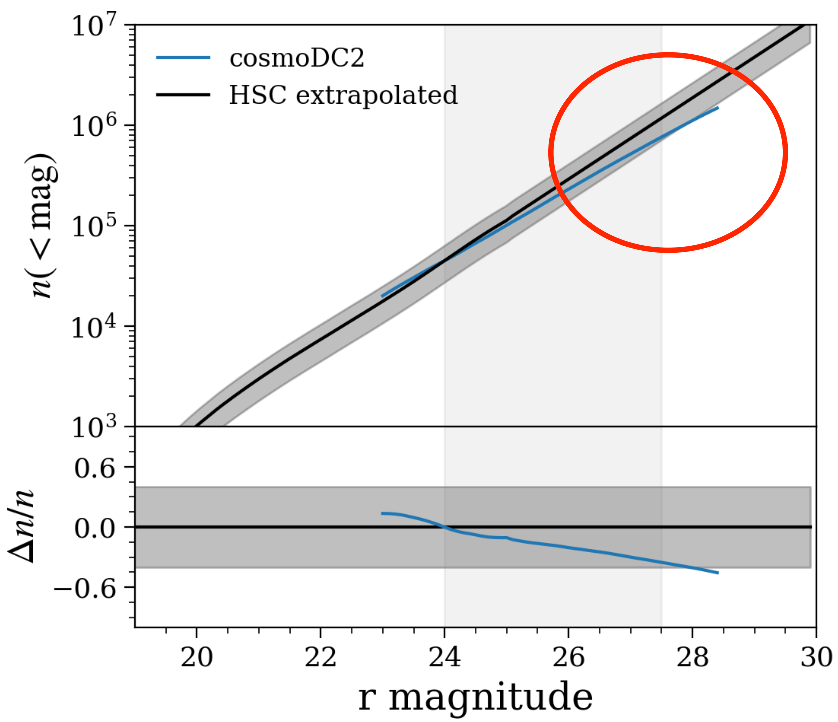
Galaxy Number Densities

- Science drivers:
  - LSS, WL (3x 2pt)
    - noise (~ 1/N) - needs to be correct for systematics estimates
    - N(z) - estimates of effective redshifts for tomographic binning
  - Photo-z's - assess accuracy of photo-z algorithms -> distribution and redshift evolution of SEDs must be accurate
  - Blending - assess accuracy and efficiency of deblending algorithms
    - Simulations crucial for estimating effects of unrecognized blends
- Validate number densities of galaxies as a function of redshift and Flux(λ) (filter band-passes (*ugrizY*)
  - Complex multi-dimensional space not fully explored by available data - need to consider projections of this space and match to available data

Galaxy Number-Density Tests

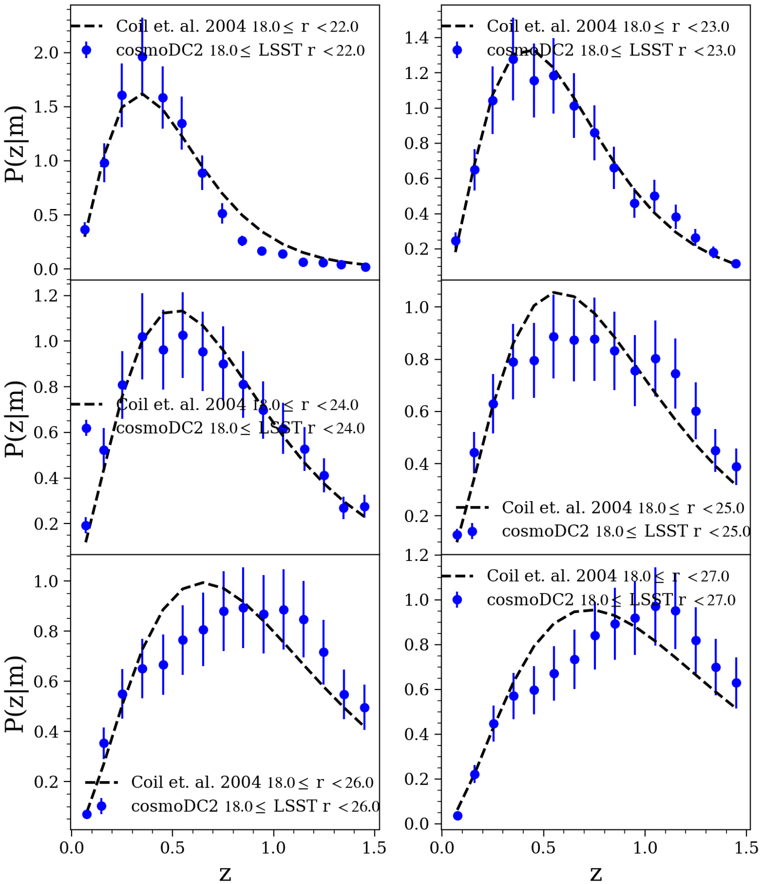
- Cumulative number densities as a function of magnitude in *grizY* (tests absolute normalization)
- N(z) for selected magnitude ranges (tests shape)
- Color distributions for selected magnitude and redshift ranges (tests 1-d projections of galaxy colors)
- Color-color and color-magnitude distributions for selected magnitude and redshift ranges (tests 2-d projections)

dn(<mag)/dmag



- Validation data requires a wide (to reduce cosmic variance) and deep (for faint magnitudes) survey
  - Extrapolate results from the HSC survey (wide)
  - Cross check with numbers from pencil-beam surveys
- Validation region: $24 < r < 27$
- Validation criteria:
  - +/- 40% fractional difference
- Addition of faint synthetic galaxies were critical for passing this test

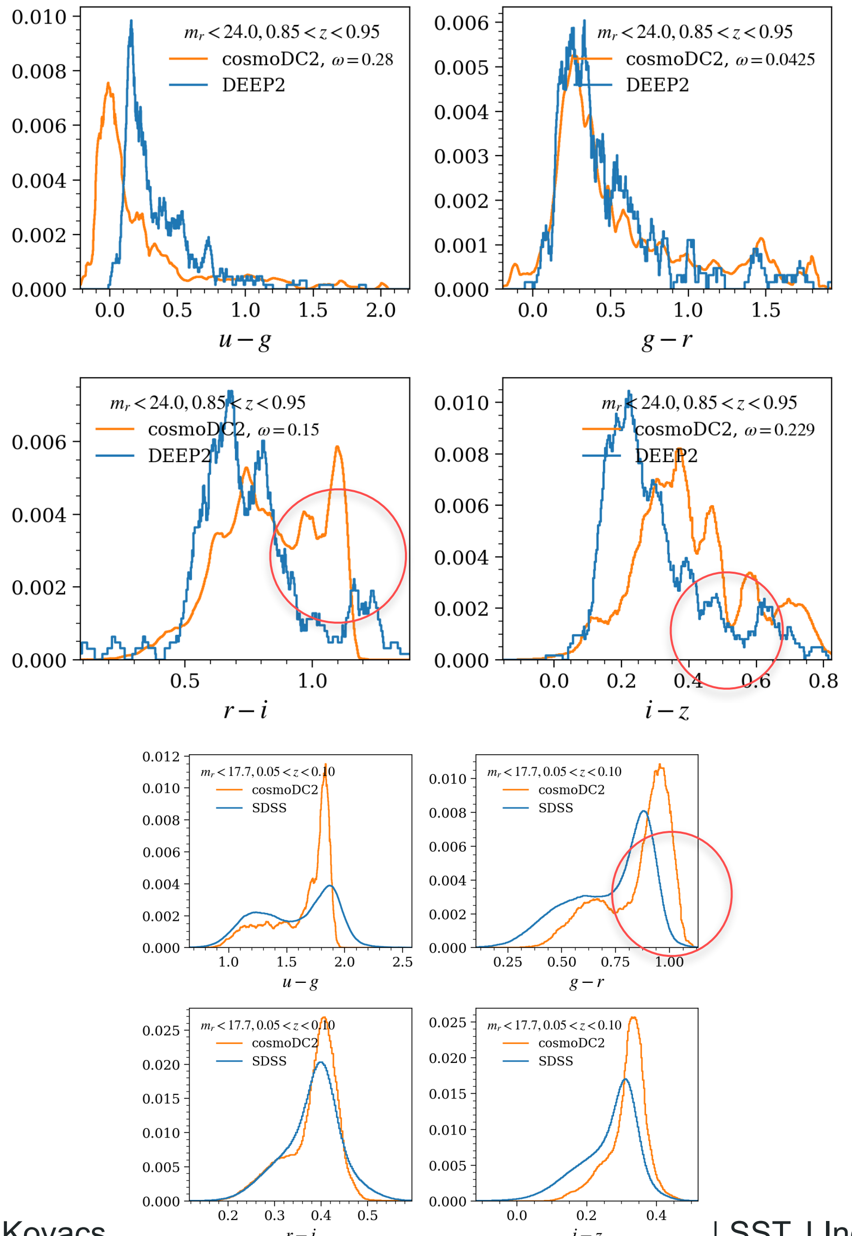
Number Density vs Redshift



Validation Test:

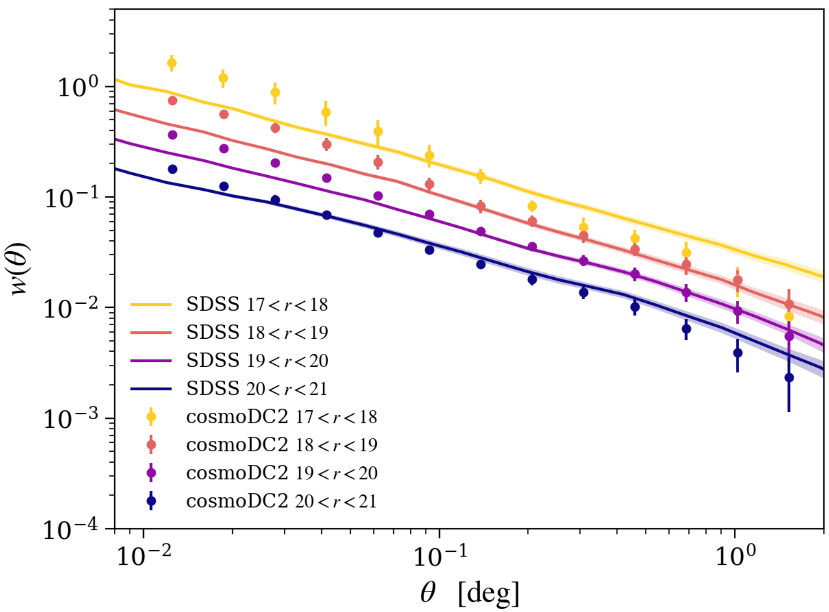
- P(z|m)
- Validation data from DEEP2 for magnitude-limited samples
  - Fits to observed distributions (dotted lines)
  - 0 < z < 1.5
- Jack-knife errors for catalog data to account for cosmic variance

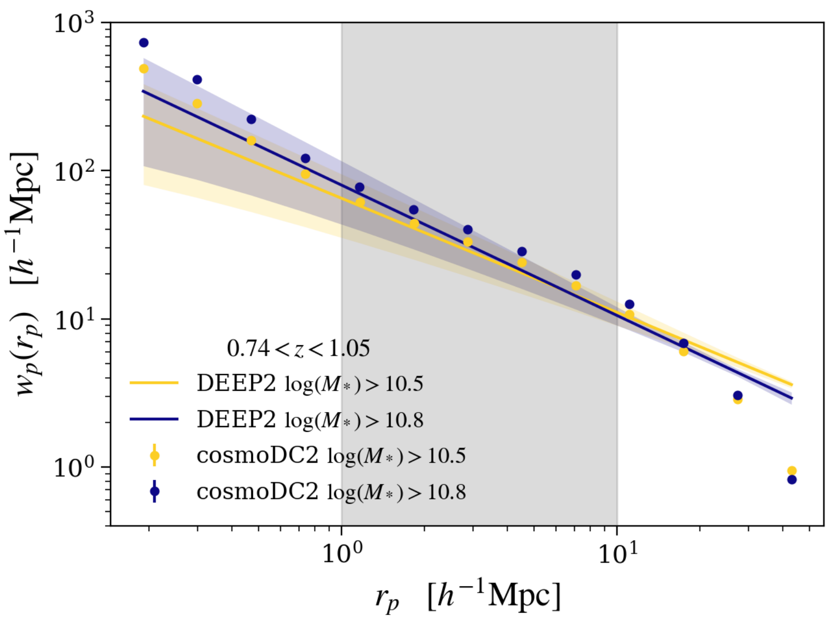
Color Distributions



- Science Drivers:
  - PZ
- Data
  - SDSS *griz* catalog (Zhou et. al.)
  - DEEP2 (CFHT *griz*) catalog (Zhou et. al.)
  - Test uses cosmoDC2 SDSS magnitudes and includes filter transformations for SDSS -> CFHT
- cosmoDC2 contains more red galaxies to ensure that an abundant red sequence is available for redMaPPer

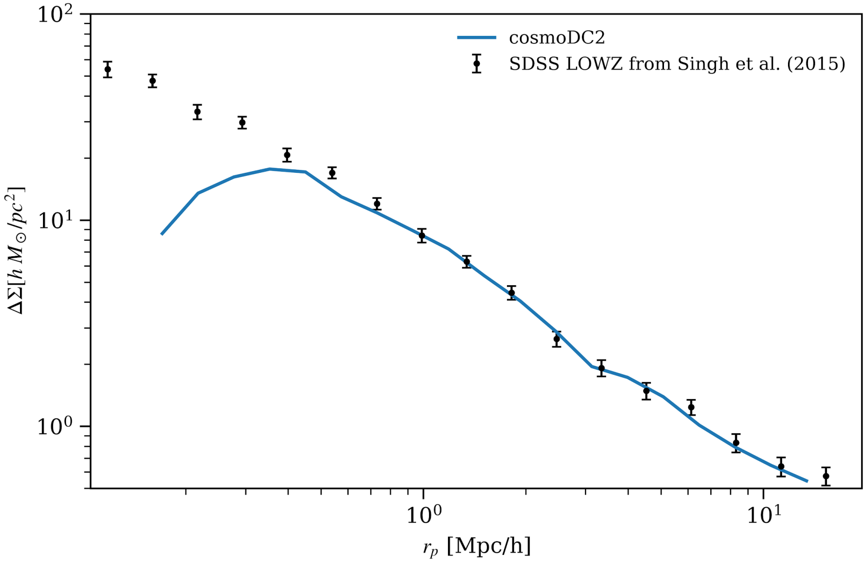
Correlation Functions



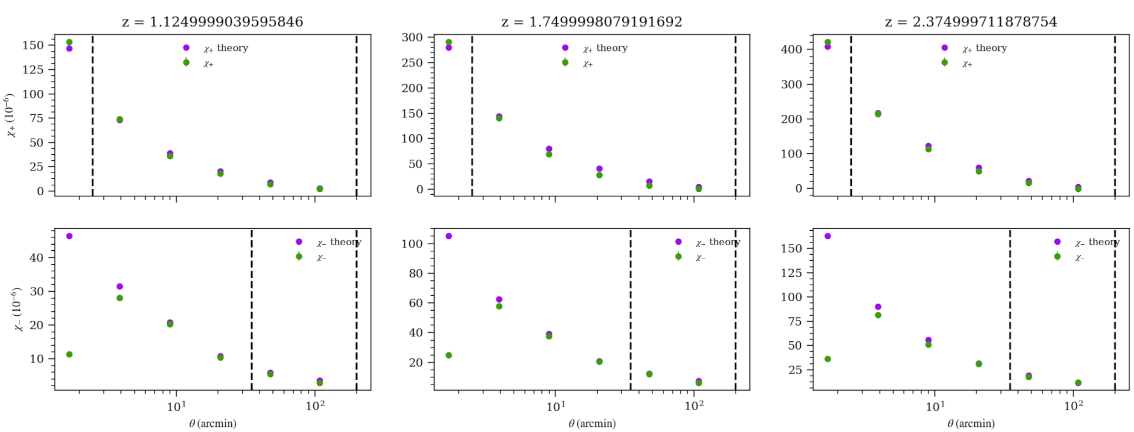


- Science Drivers:
  - LSS
  - WL
  - PZ (calibration)
- Data
  - SDSS (Wang et. al.)
    - $17 < r < 21$
  - DEEP2 (Mostek et. al.)
    - $0.74 < z < 1.05$
    - Fits for 2 stellar-mass cuts
  - SDSS
    - $w_p(r_p)$ for selected color and magnitude cuts

Galaxy-Lensing and Shear-shear correlations

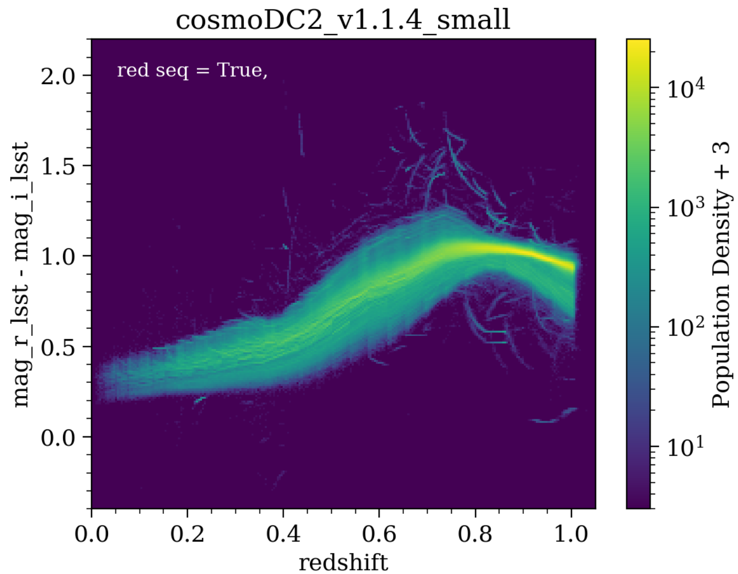


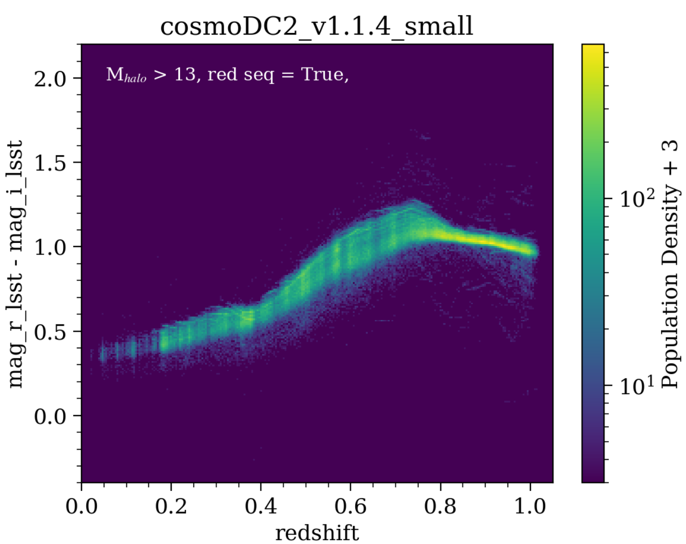
- Science Drivers
  - WL
- Data
  - SDSS LOWZ sample (Singh et al) for the excess surface mass density: $\Delta\Sigma \sim < \gamma_t(r_p) > / \Sigma_{crit}$
  - Theory predictions for ξ^+, ξ^- components of the $\gamma\gamma$ auto-correlation are based on CAMB power spectra; confirms sign conventions for shears



- Galaxy number densities after cuts to mimic the LOWZ sample are in good agreement
- Test confirms that weak-lensing quantities and galaxy positions are correlated as expected

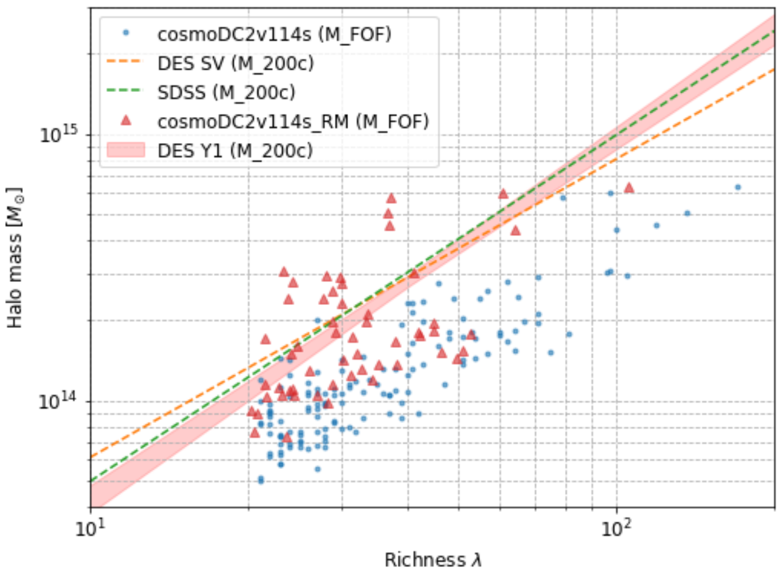
Galaxy Clusters

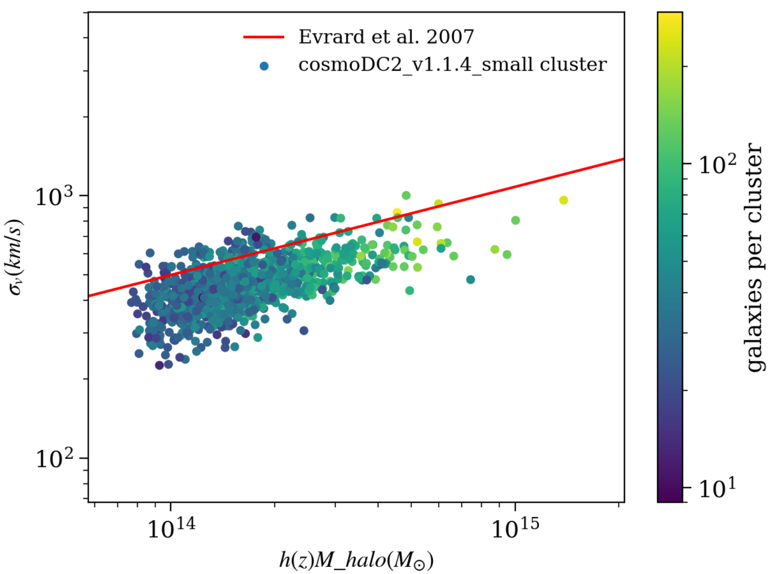




- Science Drivers:
  - CL
- Data
  - DES SV and Y1 data
  - Empirical fits to RS mean and scatter as a function of z
- cosmoDC2 data was tuned on the mean RS colors as observed by DES for galaxies in group- and cluster-sized halos

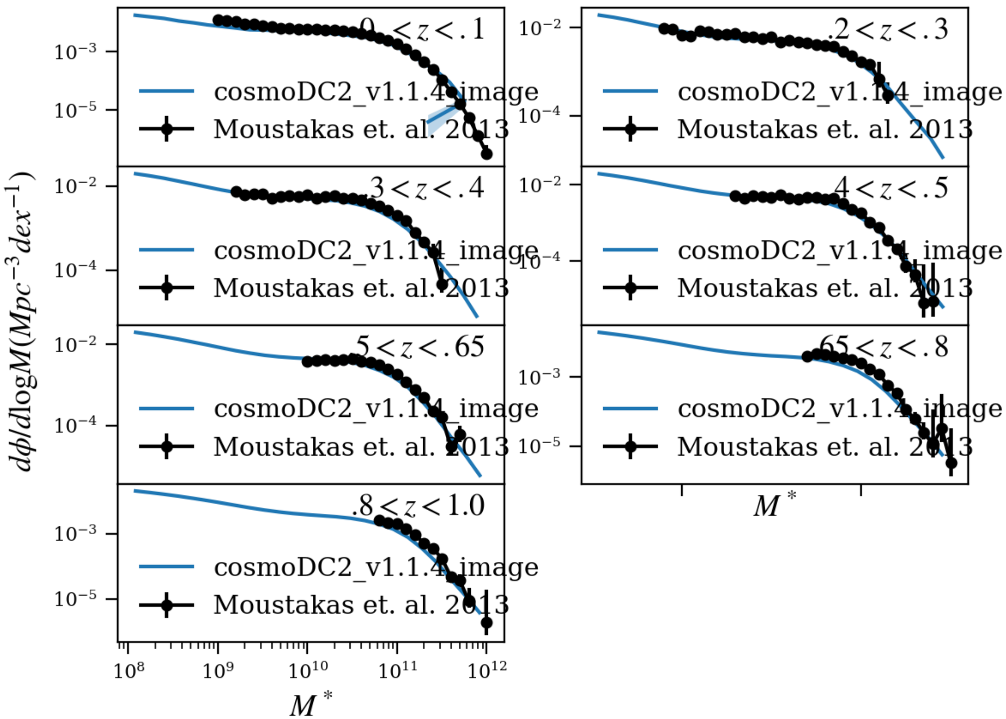
Tests on Galaxy Clusters





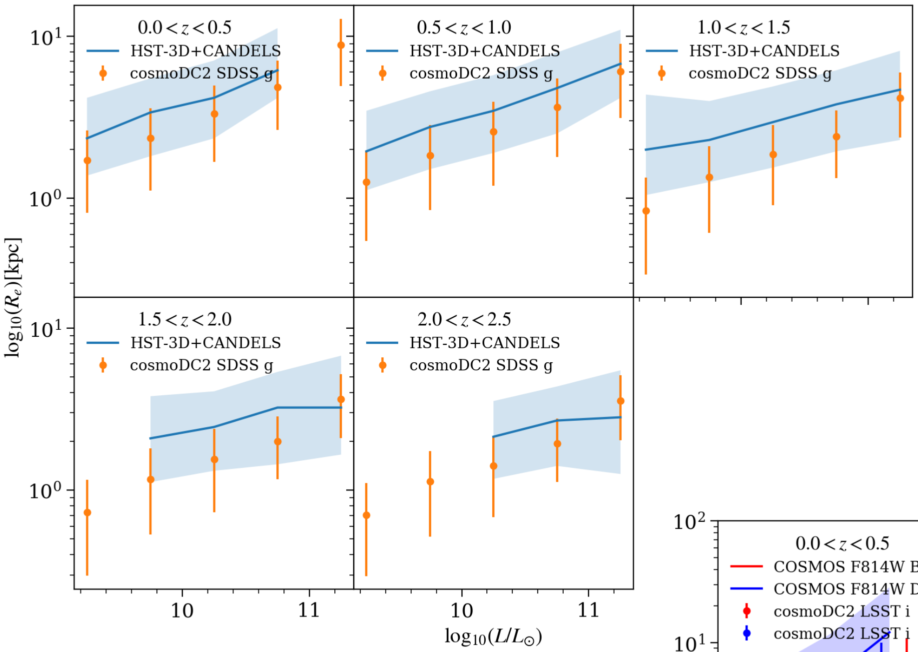
- Science Drivers:
  - CL
- Data
  - DES Y1 (McClintock et al)
    - Parameterized mass-richness relation
    - Mass definitions require conversion from M_{200m} to M_{200c}
  - Velocity dispersions from Evrard et al

Stellar Mass Function

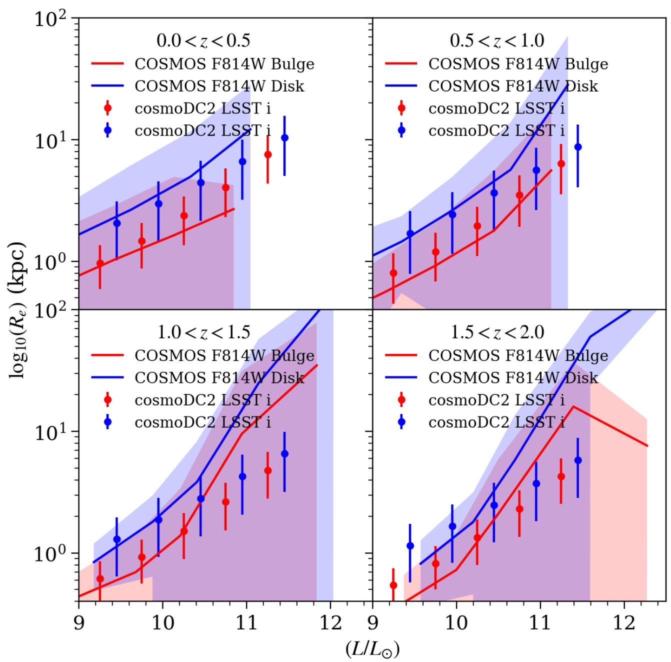


- Science Drivers
  - Fundamental probe of the galaxy-halo connection
  - Relevant for SN science as SN rates and properties are strongly correlated with host stellar mass
- Data
  - PRIMUS (Moustakas et al)
    - redshift slices in 0<z<1
  - CMASS
    - Complex color selection
    - 0.4 < z < 0.7
    - Mean M* and average number densities compare well (93 cf 101)

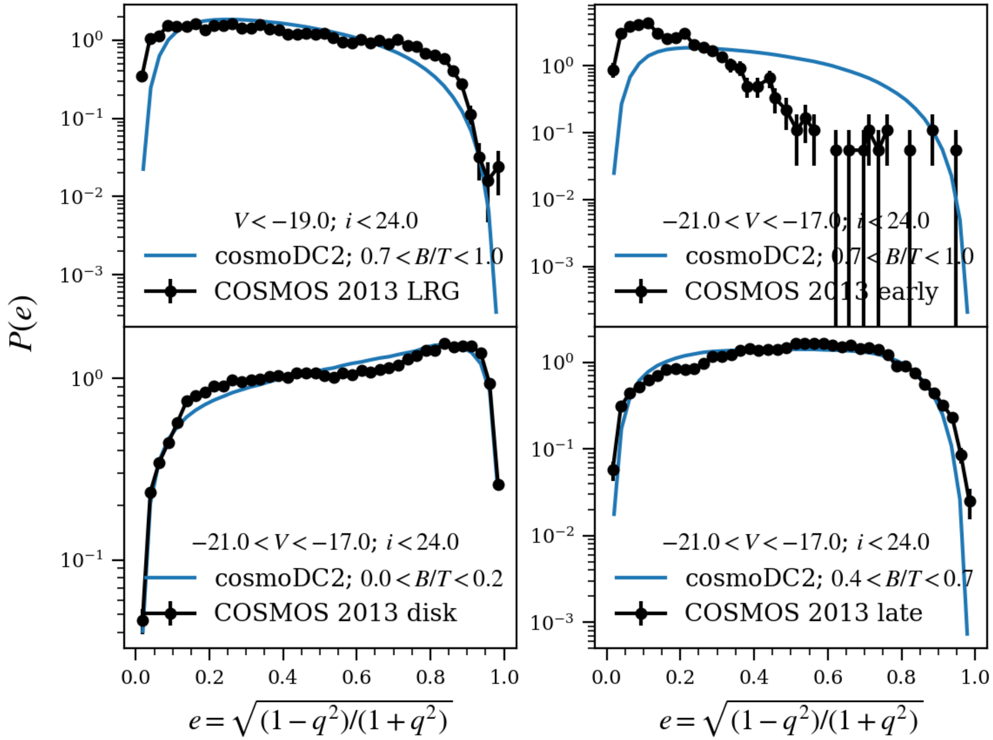
Size Distributions



- Science Drivers
  - WL
  - Blending
- Data
  - HST-3D + CANDELS
  - Mandelbaum et al
- CosmoDC2 was tuned to match these distributions



Ellipticity Distributions



- Science Drivers:
  - WL
  - Blending
- Data
  - COSMOS (Joachimi et al)
  - Complex morphological selection is mimicked by using B/T cuts
  - CosmoDC2 galaxies were tuned to these data

Catalog Production Summary and Outlook

LSST DESC
Dark Energy Science Collaboration

- New hybrid method simulates realistically complex galaxies with an extended range of properties
- Leverages smaller simulations by resampling galaxies into larger volumes
- Catalogs can be produced in “embarrassingly parallel” mode in just a few days, once the halo catalogs, tuned empirical model and SAM library are available.
- Method retains flexibility to allow for adjustments to the empirical model if required
- Could replace the SAM library by resampling galaxies from a smaller hydro-simulation

Catalog Validation Summary and Outlook

- Validation testing is required to determine how a catalog can be used effectively
- LSST DESC test suite is applicable to ground-based imaging surveys targeting dark-energy science
- Curated set of validation data is relevant for imaging surveys using optical bands, but could be extended to other bands
- Test suite can be adapted for other surveys by augmenting the validation data and adjusting the validation criteria, depending on the science goals of the survey.
- Paper is in preparation

Coming Soon: SkySim5000

- 5000 sq. deg. catalog, z~3
- Improved mitigation of discreteness effects
- Tri-axial NFW satellite-galaxy distributions based on Outer Rim halo shapes

Backup Slides

N-body Simulation Equations

$$\ddot{\mathbf{x}} + 2\frac{\dot{a}}{a}\dot{\mathbf{x}} = -\frac{\nabla\Phi}{a^2}$$ Equation of motion for tracer particles in expanding Universe

$$\frac{\dot{a}}{a} = H = \frac{H_0}{a^{3/2}}\sqrt{\Omega_{tot} + a^3\Omega_\Lambda}$$ CDM + baryons + DE

$$\nabla^2\Phi(\mathbf{x}) = 4\pi G a^2[\rho(\mathbf{x},t) - \rho_b(t)]$$ Poisson equation

Designing Verification Tests

- Checks to ensure code produces the expected results
- Tests can run during production or in post processing
    - assert statements for conditions that must be true
    - Check that quantities reproduce known results in known cases
    - Check that indices have correct ranges
    - Check that ids that should be unique are unique
    - Check that quantities that should match do match