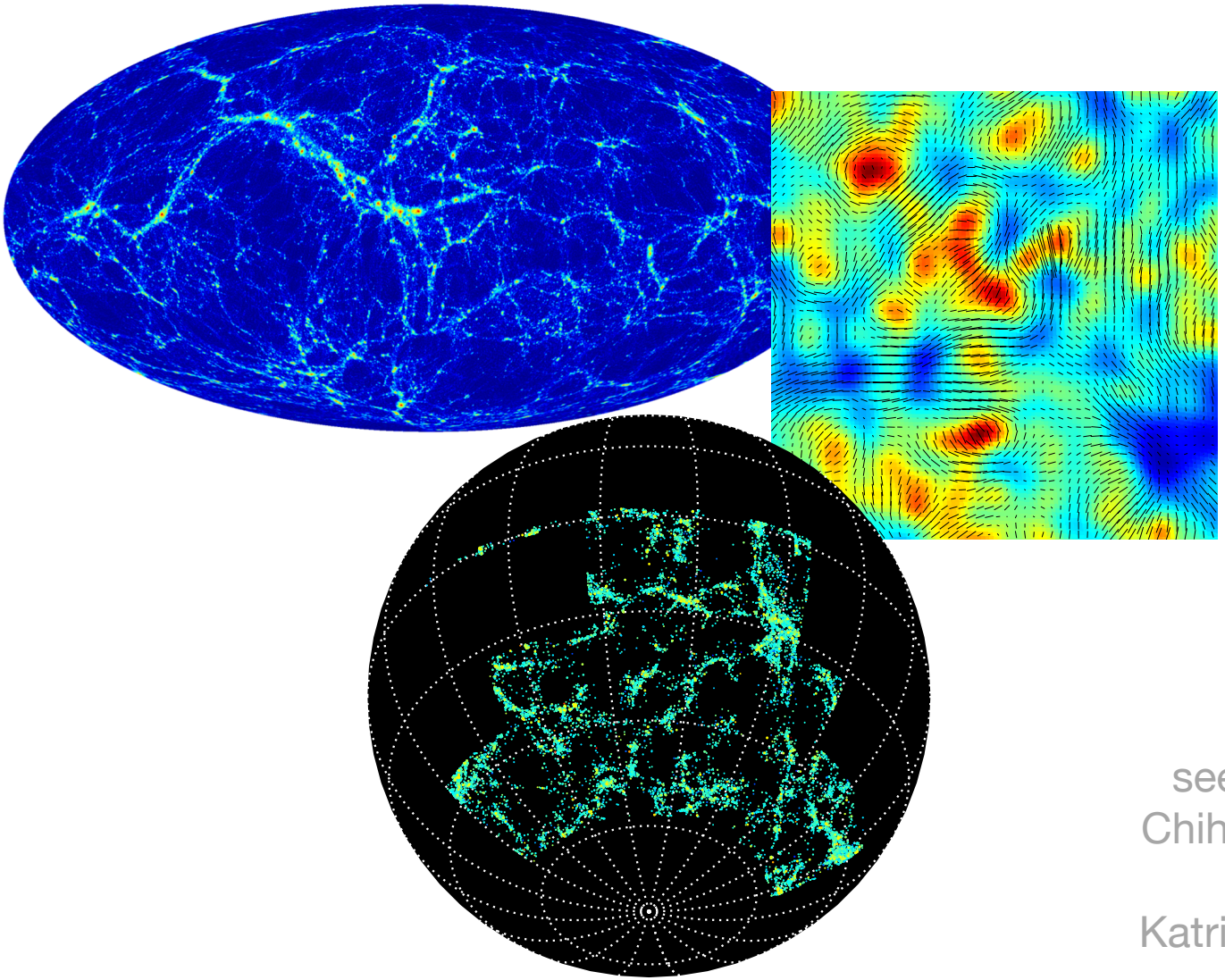
Simulated Galaxy Catalogs for DES and LSST



Risa Wechsler
Stanford/SLAC/KIPAC

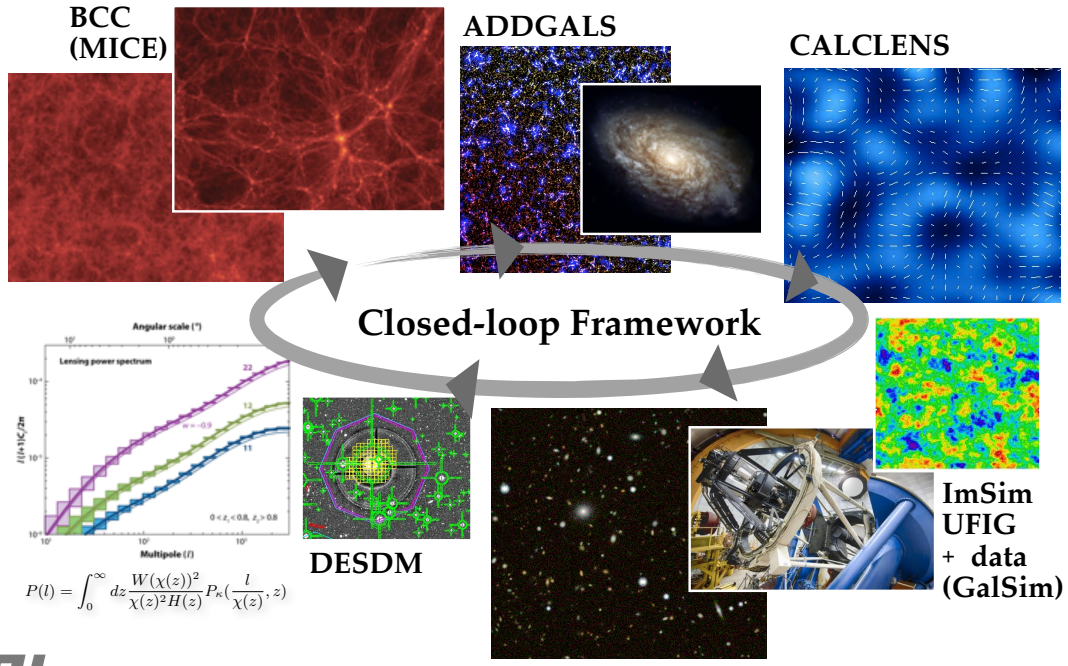
with Matt Becker
Michael Busha
+ DES sim working group
and many others!

see also related talks by
Chihway Chang (yesterday)
and
Katrin Heitmann (tomorrow)

Motivation for Simulated Sky Surveys

- Engaging with realistic, survey-size data helps with e.g.
  - data management and access
  - code parallelization
  - understanding systematics for precision cosmology
  - developing new analysis ideas in a realistic context
  - testing cross correlations and joint analyses
  - assessing followup plans (spectroscopic, multi-wavelength)
  - etc...
- Cosmological probes with DES and LSST are systematics -limited
  - theoretical: mass function, bias function, moments of non-linear structure formation
  - observational: photo-z's, projection effects, star-galaxy separation, impact of mask: **can be quantified with simulated sky surveys**
  - many need full volume & have cosmology dependence
  - need to understand in gory detail **how to go from cosmo parameters --> observables so that we can get from observables in the real data --> cosmo parameters**

End-to-end Simulation @ DES



ETH zürich

- Under development for more than a decade (starting with modeling SDSS)
- From cosmological simulations to catalogs and images.
- Early versions of our galaxy catalogs --> images --> input into data management while survey and survey software was under development.
- Several generations of catalogs have been provided to science collaboration and are in active use in understanding systematics, developing analysis tools, and extracting science.
- Have been critical for many years in many aspects of the project!

Simulated Sky Surveys

Want simulations that allow a realistic cosmology analysis for the main dark energy probes

- cluster abundance and clustering
- galaxy clustering & baryon acoustic oscillations
- lensing: shear-shear correlations; galaxy-galaxy lensing; cluster mass calibration
- cross-correlation between galaxies and the CMB
- etc

Want to produce a realistic simulated sky

- observed properties of galaxies
- large-scale structure of galaxies
- realistic impact of lensing shear on galaxies
- as many relevant observational systematics as possible

Want to produce *many* full area and depth sky surveys; need lightweight simulations

- many cosmological models
- a variety of galaxy models for a given cosmology
- multiple skies for covariance

Strategies for galaxy catalogs

- high resolution: associate all galaxies with resolved halos and subhalos.
  - assign luminosities using abundance matching + galaxy properties based on environment
  - active work on color models, which are not as mature
  - extensive testing against data from SDSS at low z, including correlation functions, group statistics, galaxy-galaxy lensing, etc.
  - need very high resolution, e.g ~ kpc force resolution and 1e8 mass resolution to resolve Mr = -19 galaxies.
  - currently have/creating catalogs based on various boxes with ~ 150-600 Mpc
  - SAM models on the same merger trees using model of Yu Lu, further development informed by empirical results... these are coming along but in my opinion no existing SAMs are there yet.
- medium resolution: minimum needed, in order to produce multiple sky surveys in many cosmologies
  - associate all galaxies with dark matter overdensities + central galaxies where halos are well resolved
  - iterated based on lessons from
  - well developed pipeline
  -

- simulation lightcone
- galaxy luminosities
- SEDs for galaxies
- shear at every galaxy position (current version, 6.2” resolution)
- galaxies lensed / sheared & magnified
- photometry in many bands
- photometric errors & photometric redshifts
- integration with UFIG and preliminary integration with LSST phosim

Large area “Blind Cosmology Challenge” simulations (“Aardvark/Buzzard-v1.0”)

- available simulations:
  - LCDM cosmology; N-body lightcones to z~2 (based on 3 sim boxes with 2048^3 particles)
  - +additional cosmologies and volume (blind parameters for DES Blind Cosmology Challenge)
  - halo finding from rockstar, includes multiple mass def., concentrations, etc.
  - ~ 1 billion galaxies added using ADDGALS, over 1/4 sky (10313 sq. degrees), complete to i ~ 25
  - photometry in many bands, including LSST bands and DES, SDSS (DR8+S82), VISTA (VHS +VIKING), CFHTLS, NDWFS, DEEP, WISE, IRAC
- shear on the full quarter of sky using CALCLENS; currently with 6.2” resolution
- extensive development and testing with SDSS data and other higher redshift data, including early DES data; designed to go to full DES depth
- + stars and quasars
- + simulated spectra (SPOKES) and simulated images (UFIG)
- should contain all of the galaxies in the LSST “gold sample”
- allows science analysis related to clusters, weak lensing, LSS, photometric redshifts, spectroscopic followup design, etc.

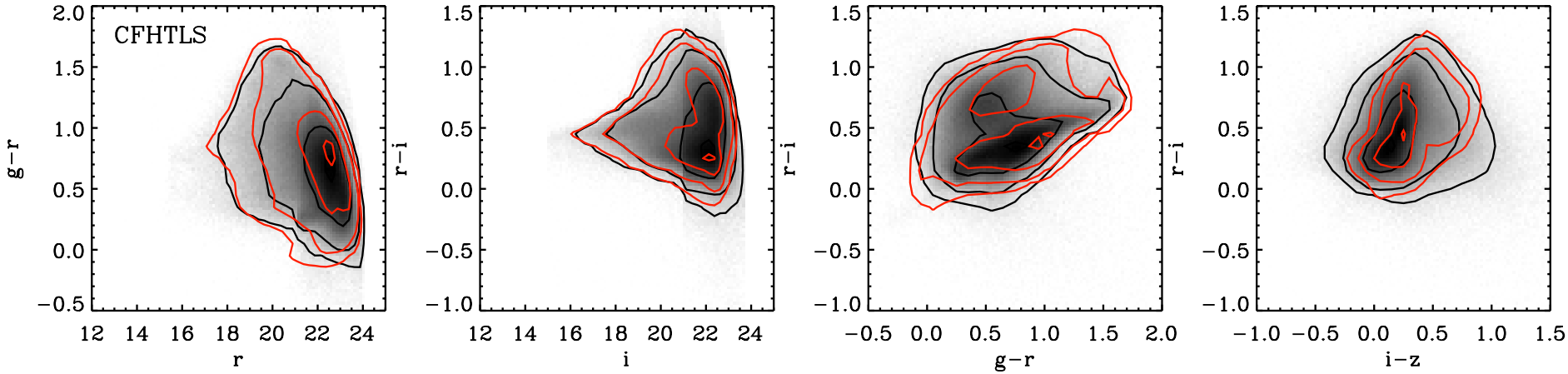
Small-area, high resolution catalogs in progress

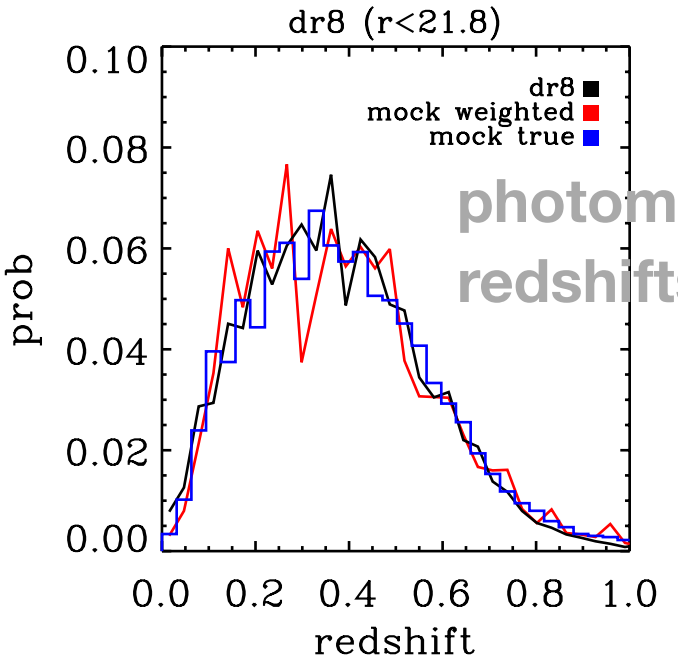
- lightcone based on populating subhalos with galaxies using empirical methods
- currently ~ 100 sq. degrees
- constructing to LSST depth
- includes lensing with CALCLENS

Post-catalog production

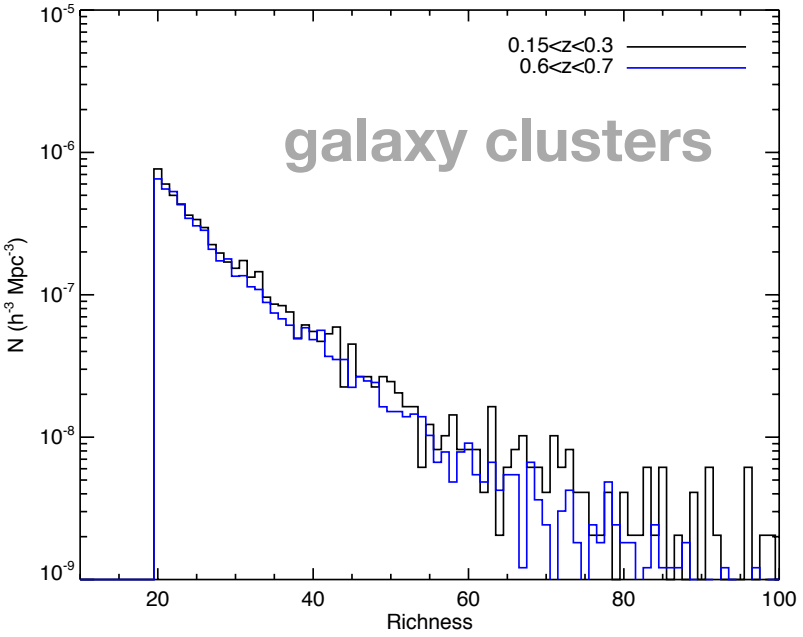
- catalogs + UFIG (with Chihway Chang)
  - produces images (fast), runs sextractor to produce new catalogs
- catalogs + ImSim tools (with Debbie Bard)
  - positions and lensing is now well integrated
  - further work to integrate with LSST (or more general) SED model.

example validation:
galaxy colors and luminosities



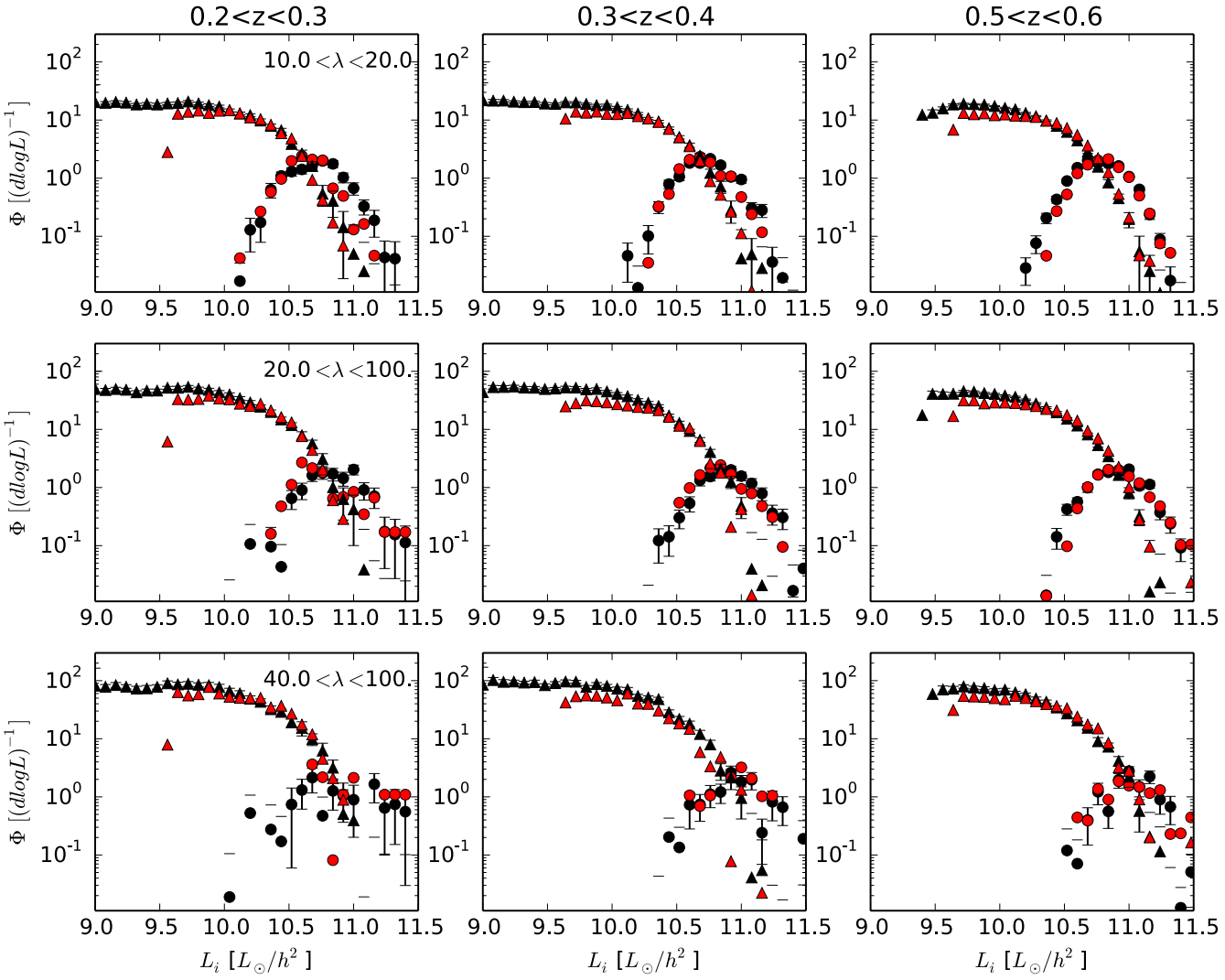


photometric redshifts



galaxy clusters

example validation:
conditional luminosity function in clusters (S82 vs sims)



R. Reddick, w. RW, Rykoff, Rozo

The DES Blind Cosmology Challenge

- Would like to assess the ability of the main DE probes to recover cosmological parameters in realistic sky surveys, including realistic systematic errors
- "VCC" Visible Cosmology Challenge
  - simulated sky with a known cosmology
  - allows code testing with known results
  - this simulation is updated as galaxy model, knowledge of galaxy population, and data model improves
- "BCC" Blind Cosmology Challenge
  - many simulated skies with cosmological parameters that are unknown to collaboration
  - coordinated analysis among LSS, lensing, cluster working groups, which determines the cosmological parameters for this suite of simulated skies.
  - have done a few rounds now
    - more than 10 groups participating (lensing, clusters, clustering, various combos)
    - so far: no groups with a "mature" analysis (e.g. using only observables, finalized through to correct parameters). hope to be there soon!
  - planning challenges targeted to specific DES data releases (SVA1, Y1A1, ...)

Making catalog production modular: where I would like to be.

key caveat: some of these steps depend critically on simulation geometry and resolution.

simulations	SLAC	MICE	Argonne	VIRGO
galaxies	ADDGALS	SAM-ABCD	HOD+colors	SHAM+colors
properties	size model A SED model B	size model Q SED model Z		
IA	linear align.	halo model	Blazek	
stars	addstars	galfast		
lensing	CALCLENS	Hilbert		
images & catalogs	photo error	UFIG	PHOSIM	DESDM
photo-z	ANNz	DES-NN	LePhare	ArborZ
validation	SDSS	COSMOS	DES	CFHTLS

(steps and options listed here meant to be representative, not comprehensive!)

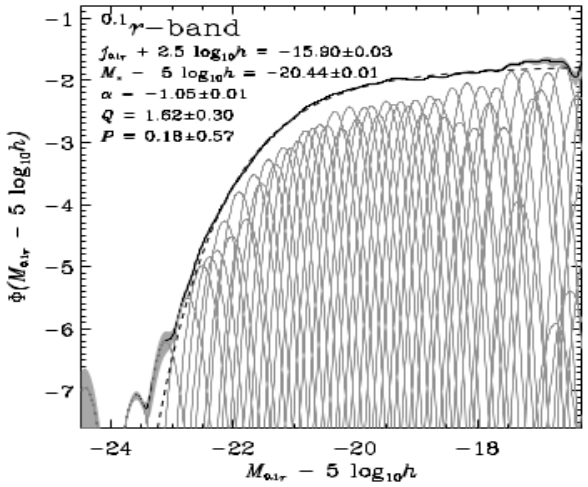
not there yet! but lots of pieces in place.

Lessons from DES experience that may be useful for LSST

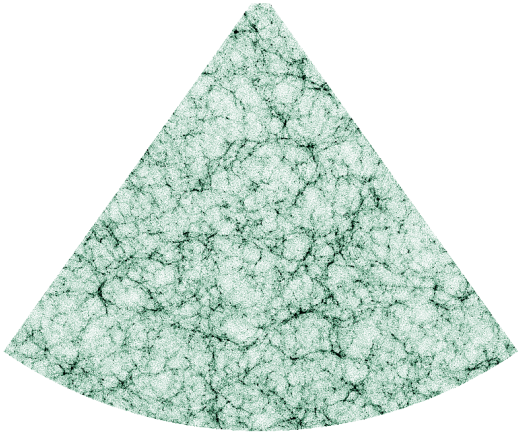
- Many people want simulations but no one knows what they want (until they try something and it doesn't work)
  - Matt Becker: "I have to estimate this thing. I don't know how to estimate this thing. I better find some simulations that allow me to estimate this thing, so that I can proceed with my analysis."
- Simulation requirements are very different for different purposes.
  - examples at different stages: making sure the instrument meets requirements. developing analysis tools. calculating covariance matrices for data analysis.
  - examples with different science goals: quantifying star-galaxy separation. galaxy cluster finding. galaxy-galaxy lensing.
- Very challenging to define simulation requirements for the various needs.
  - This needs the buy-in of the end users (collaboration scientists) -- should not just be the role of the simulators!
  - Needs a lot of communication and iteration. Good to start engaging early!
  - People who have requirements need to define them, and then somewhere in the pipeline these defined requirements need to be validated (often lots of assumptions in all directions)
- Incredibly useful to standardize formats, validation tests, etc, and build modular code.
  - this is hard and somewhat unrewarded work, so it's still in its infancy.
  - the time is right to make things modular and do more coherent comparisons between elements.
- Everyone agrees they need simulations to do the science but the mechanisms to support the work (both necessary hardware and people) are still fuzzy.

additional info

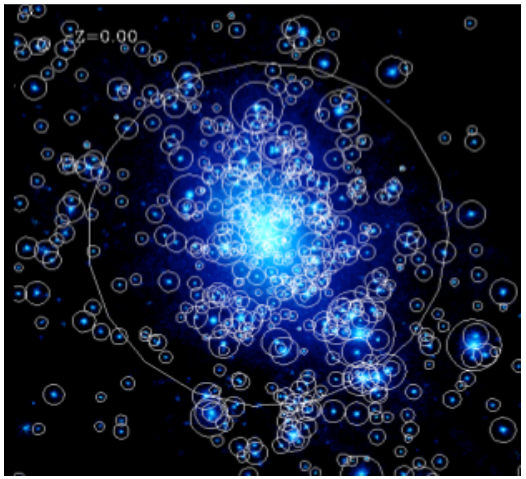
Luminosity Function



Dark Matter Lightcone



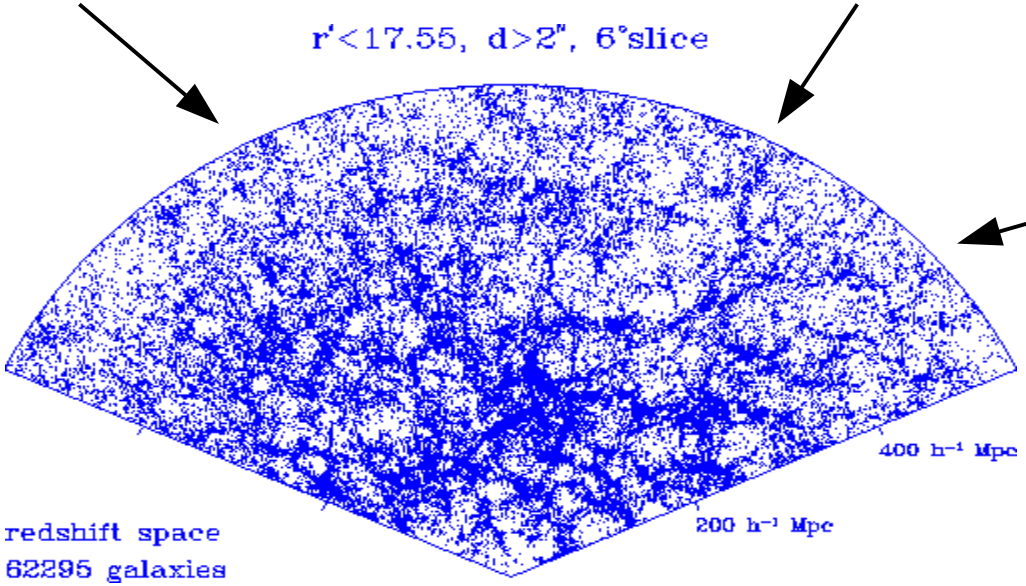
Tuning Simulation



List of galaxies with r-band magnitudes

Dark matter distribution: particles and halos

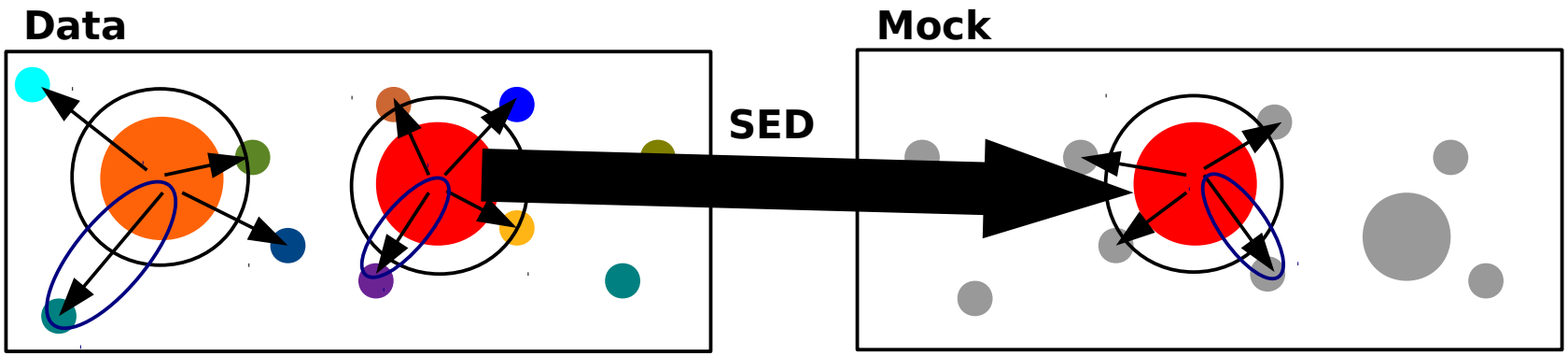
P(δdm|Mr,z): a relation between galaxy magnitudes and environment



ADDGALS
basic algorithm

ADDGALS: Color Assignment

- Once we have an r-band catalog (from any algorithm), we add SEDs using a training set of spectroscopic DR6 galaxies.
- Colors mapped to preserve the color-density relation
- Using training set, we measure P(SED|Mr,Δ5), the probability linking an SED to a r-band magnitude and local density
  - Δ5 is the projected distance to the 5th nearest galaxy
  - Colors are k-corrected
  - A model for the red fraction as a function of z is assumed.



CALCLENS: Curved-sky grAvitational Lensing for Cosmological Light conE simulatioNS

See Becker 2013

CALCLENS is a multiple-plane ray tracing algorithm designed to add weak lensing signals to mock catalogs from N-body light cones.

Features:

- works on the curved sky
- fast, approximate 2D Poisson solver
- works in the Limber approximation
- fully redshift dependent shear
- captures all of the magnification effects (i.e., finds galaxy images correctly)

Other “Features”:

- approximate 2D Poisson solver
- works in the Limber approximation

photometric errors

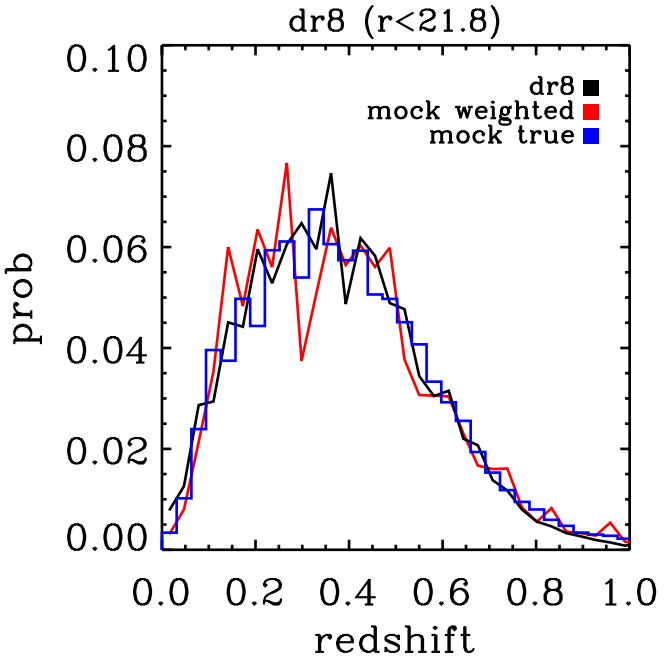
with Eli Rykoff

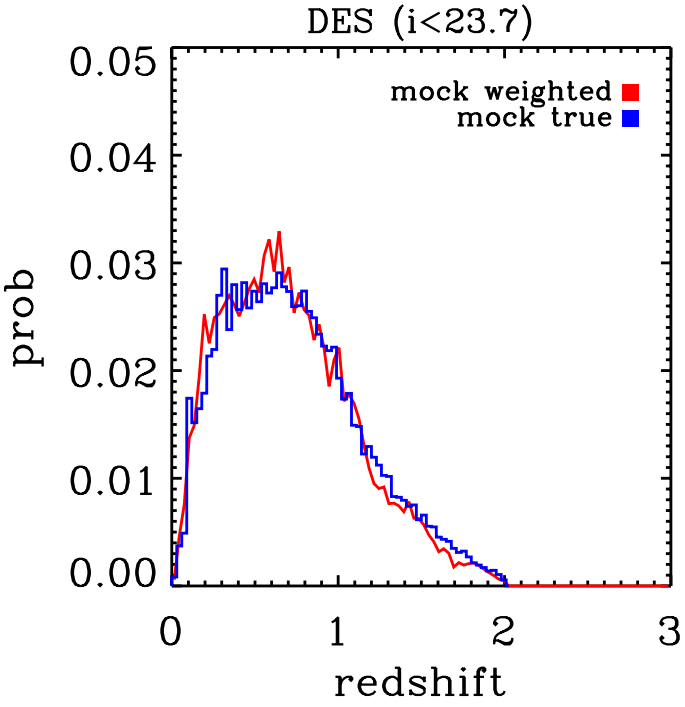
Survey			Limits				Phot. Errors?	Complete?
DECam	u	g	r	i	z	Y		
DES 10-σ		24.9	24.5	23.7	23.2	21.5	yes	5-σ in g r i z Y
DES 10-σ, SV		23.6	23.2	23.0	22.4	20.7	yes	yes
DES 5-σ galaxies		25.5	25.0	24.4	23.9	22.0	yes	yes
DES 5-σ stars		26.5	26.0	25.3	24.7	23.0	no	yes
SDSS	u	g	r	i	z			
DR8[a]	20.4	21.7	21.2	20.8	19.3		yes	yes
Stripe82[a]	22.1	23.4	23.1	22.6	21.2		yes	yes
VISTA	z	Y	J	H	K_s			
VHS[b]			20.1	19.7	19.5		yes	yes
VIKING[b]	21.6	20.9	20.8	20.2	20.2		yes	yes
VIDEO	25.7	24.6	24.5	24.0	23.5		no	no
CFHT	u^*	g'	r'	i'	z'			
CFHTLS Wide[c]	24.3	24.7	24.0	23.7	22.5		yes	yes
RCS2[d]		23.9	23.8	23.1	21.9		yes	yes
	B	R	I					
DEEP2	24.7	24.6	24.1				yes	no?
NOAO	B_w	R	I	J	K_s			
NDWFS[e]	25.1	23.8	23.6				yes	no
FLAMEX[f]				21.2	20.9		yes	no
IRAC	3.5μ	4.6 μ						
IRAC-Shallow	6.4	8.8						
FLAMEX								
SHELA								
WISE	3.4μ	4.6 μ						
WISE[f]	19.4	18.6					yes	yes?

self consistent photometric error model based on existing data from surveys.

BCC photometric redshifts

- DR8 photometric redshifts
  - using the methods of Sheldon, Cunha et al
  - uses similar training set as DR8
  - p(z) for all galaxies with r < 21.8
- DES photometric redshifts
  - optimistic DES training set using 150 0.8 sq. degree patches with galaxies to i < 24
  - “current” training set using only existing data.
- have run this with several different codes (NN, ANNz, LePhare, ArborZ, etc.)





with Michael Busha and
DES photoz folks (Cunha, Abdalla, Gerdes, ...)