

Building mock catalogues for large surveys

Carlton Baugh

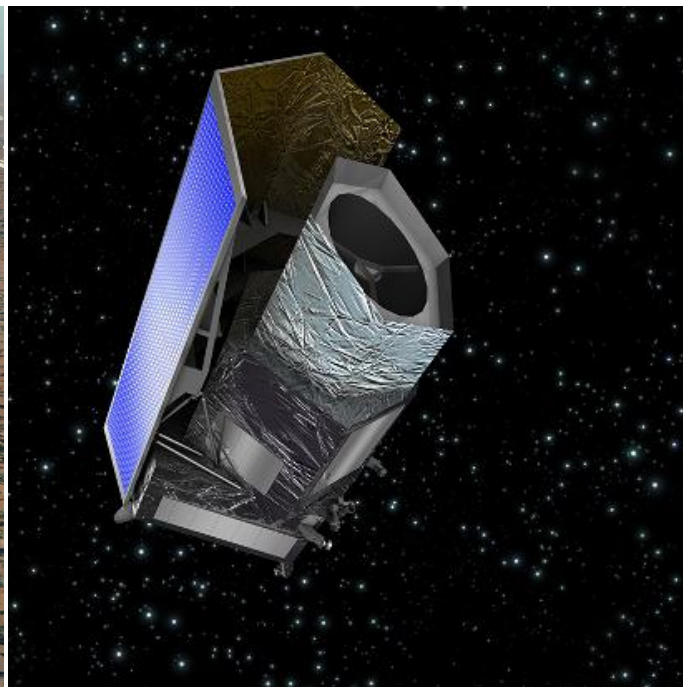
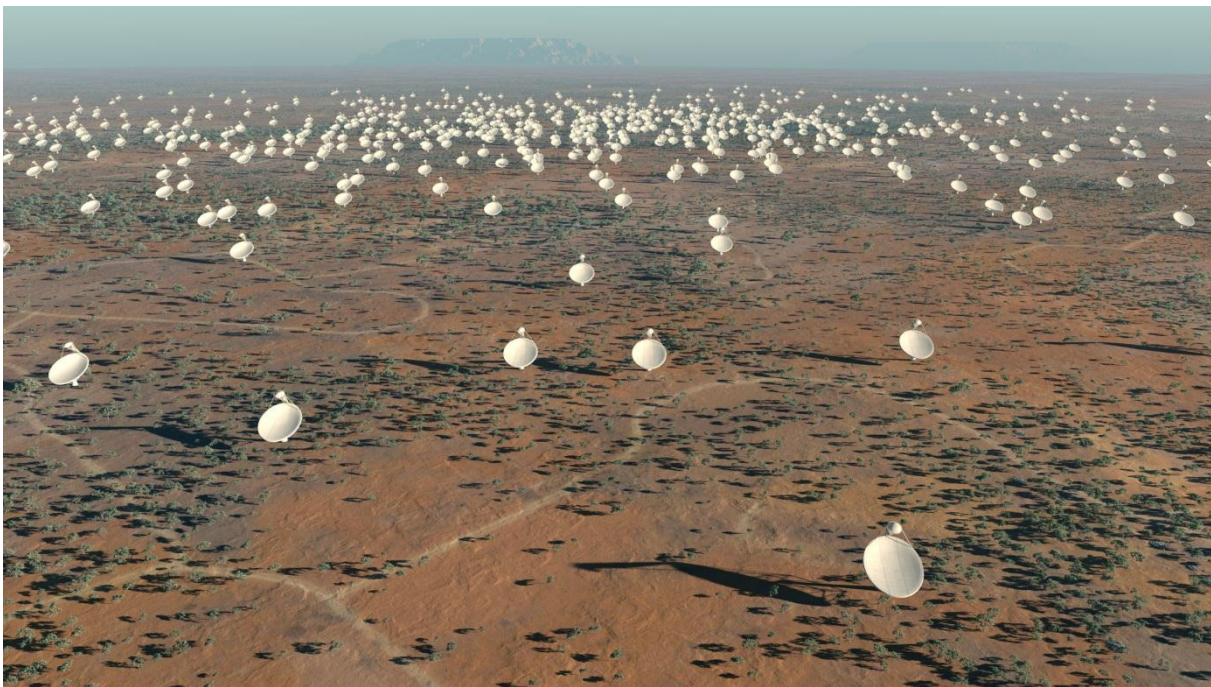
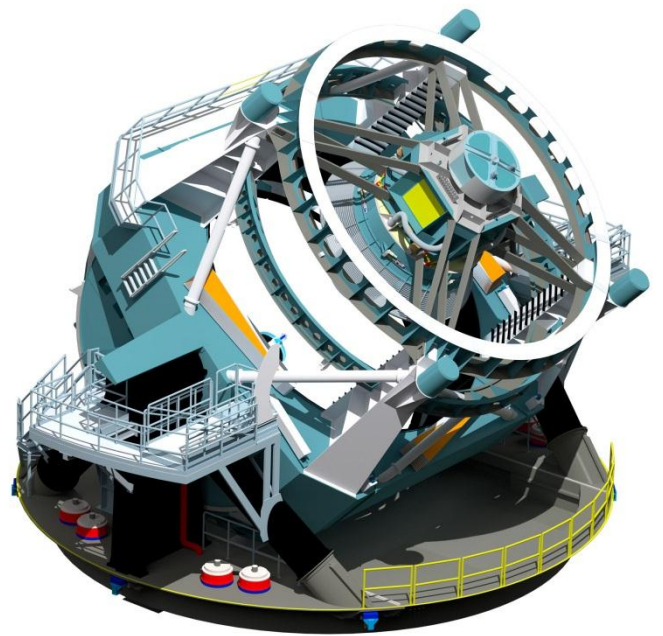
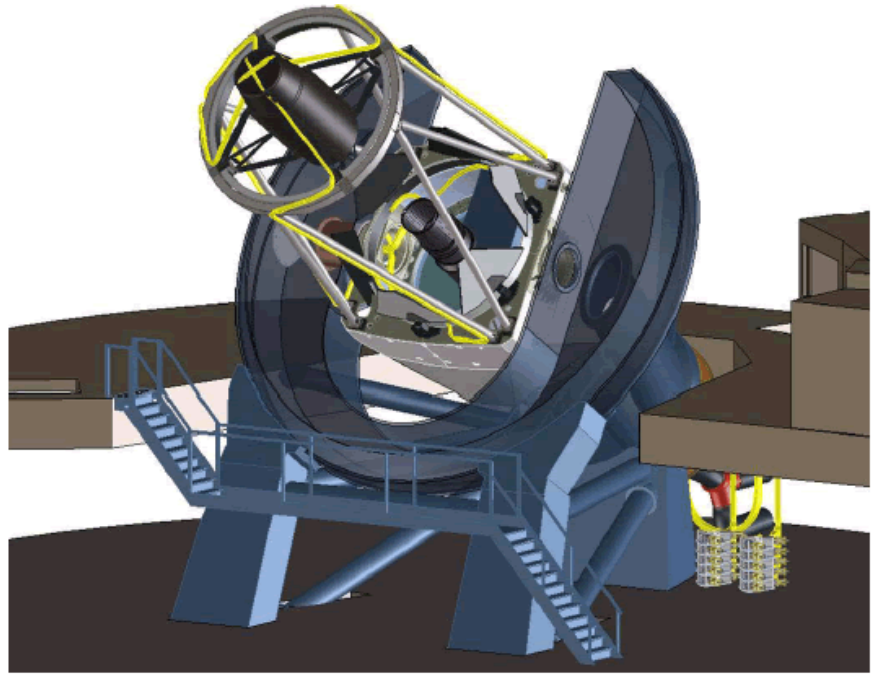
Institute for Computational Cosmology

Durham University




Millennium Workshop 2012





Future large surveys

- Photometric
e.g. Pan-STARRs, DES, LSST, Euclid-VIS
 - Spectroscopic
e.g. BOSS, BigBOSS, Euclid-NIS
 - Multi-wavelength
e.g SKA (HI)
-  Wide-field (>10,000 sq deg), wide redshift ($z=0-3$)
z-surveys: 10-50 million galaxies
imaging surveys ~billions of galaxies

Why build a mock?

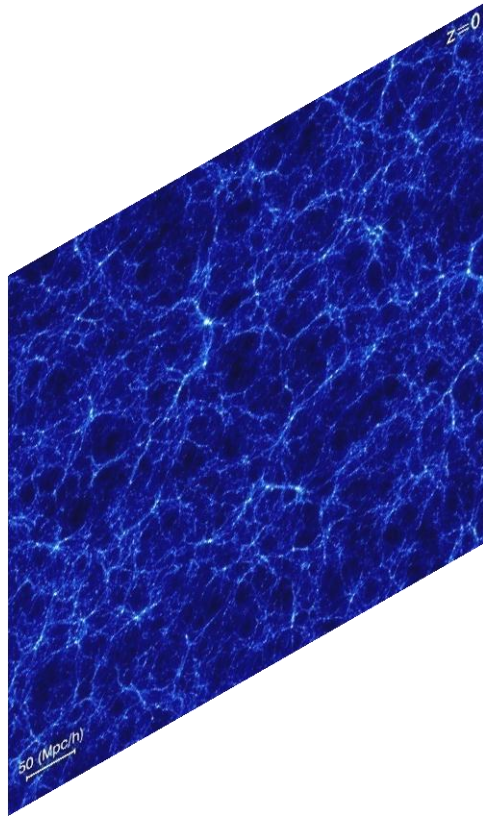
- Test galaxy formation models
- Test algorithms - validation
- Test processing pipelines
- Assess survey performance (FoM)

 Large surveys need mocks now!

What is a mock catalogue?

STEP 1

N-body simulation

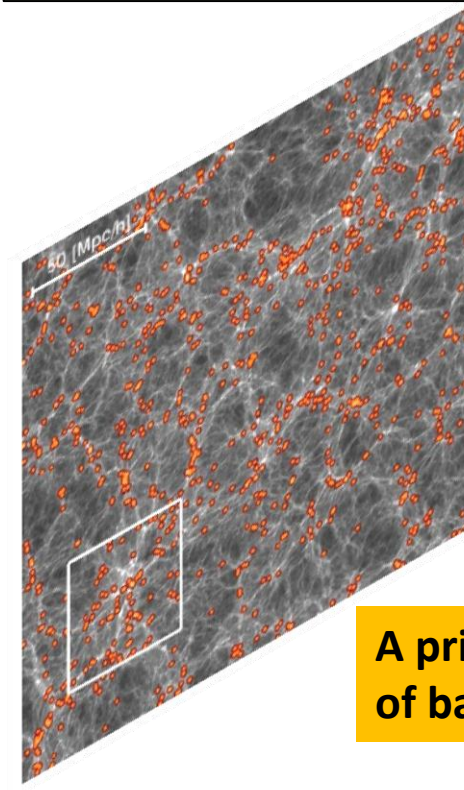


Millennium, Horizon, MICE, MXXL



STEP 2

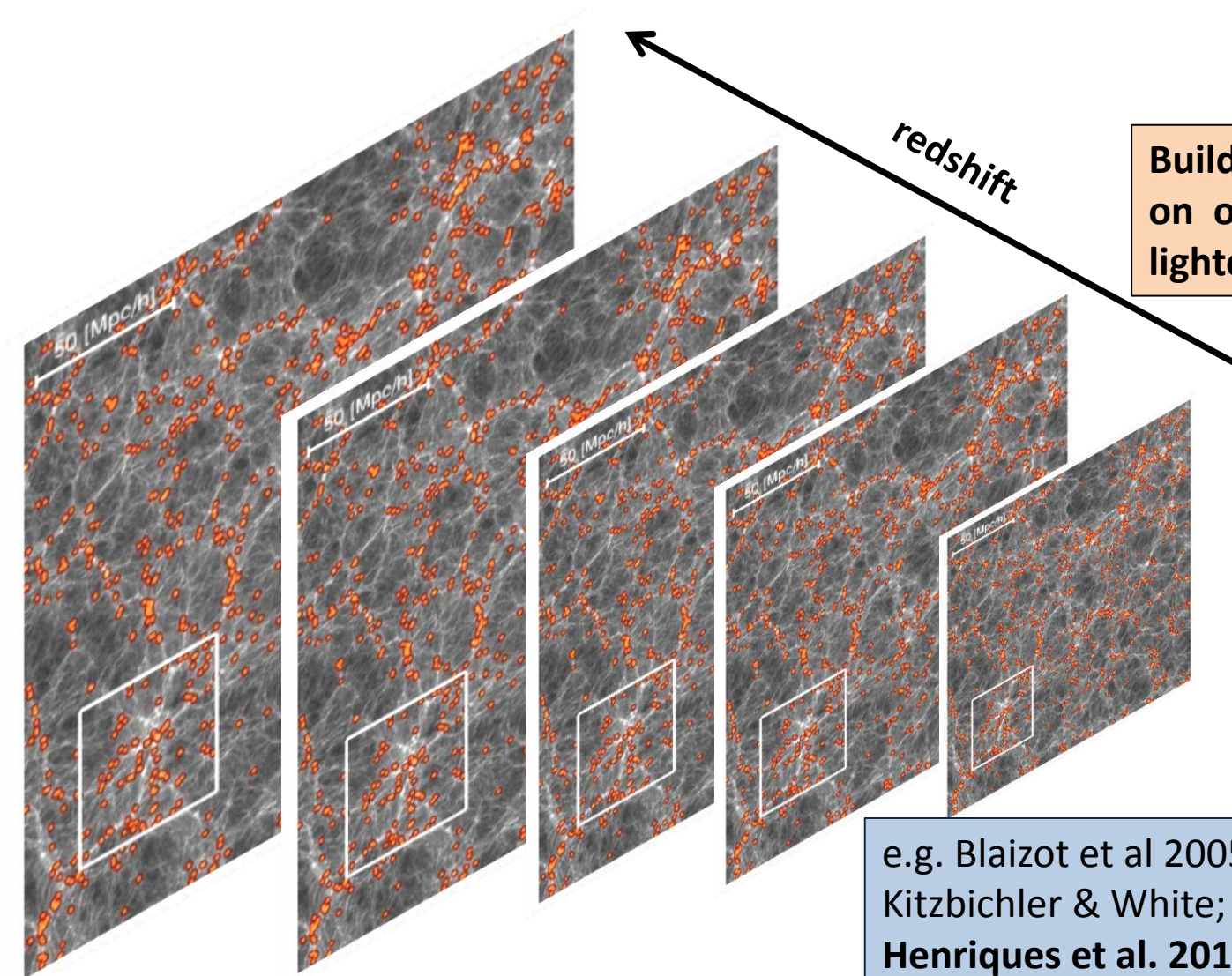
Combine with galaxy formation model



A priori calculation
of baryon physics

e.g. GALFORM; GALICS, LGALAXIES; MORGANA

Constructing observer's past lightcone

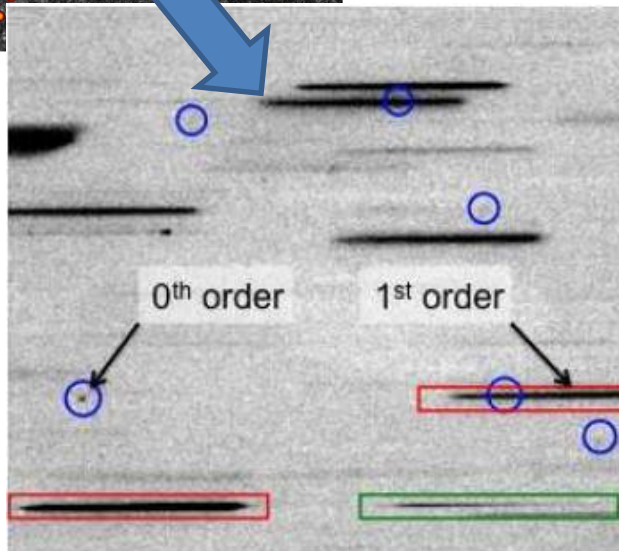
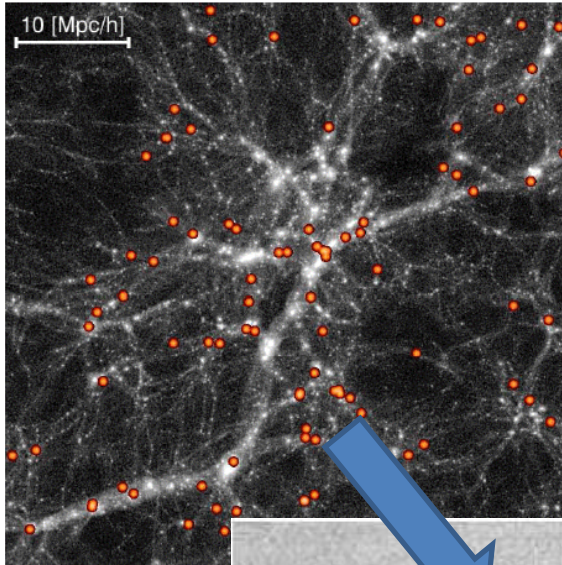


STEP 3

Build galaxy catalogue
on observer's past
lightcone.

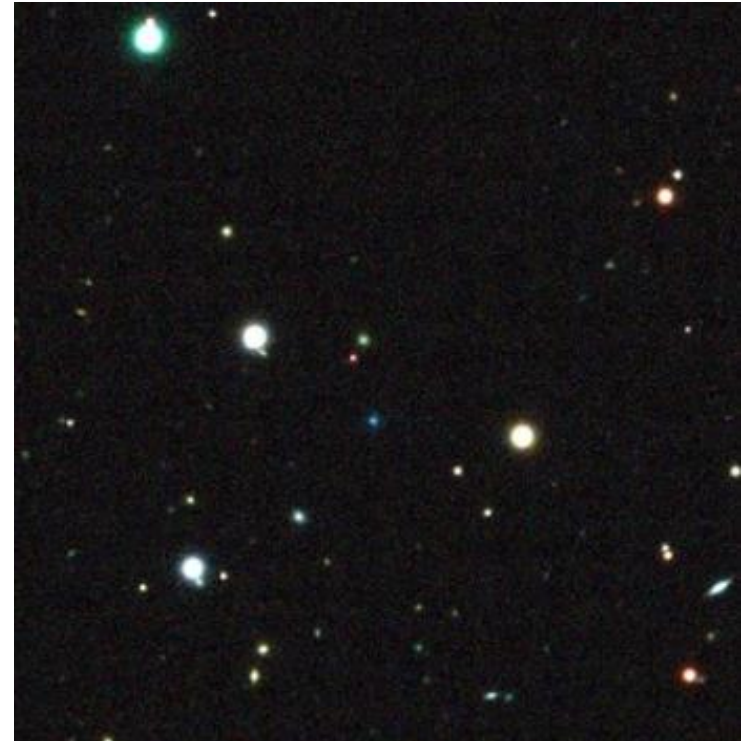
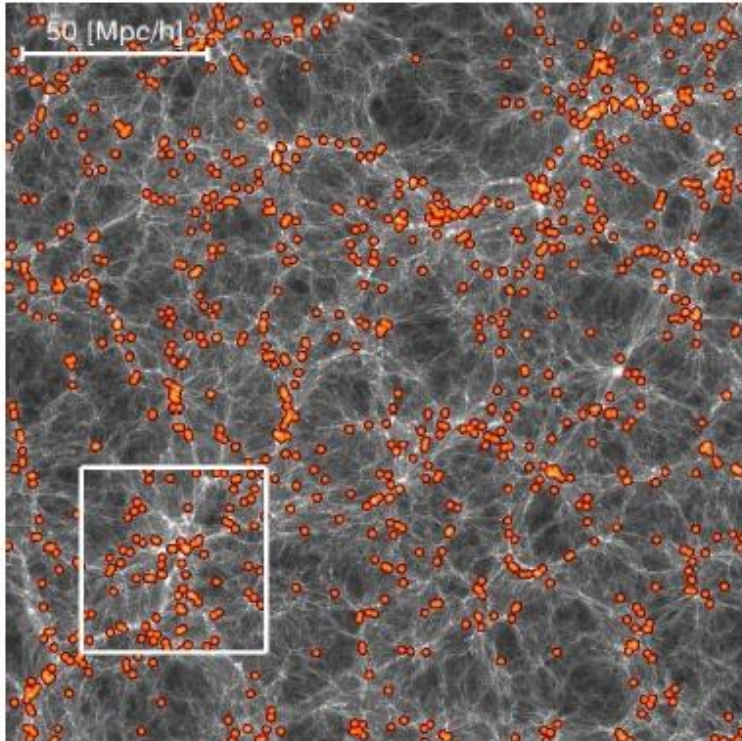
e.g. Blaizot et al 2005; Fosalba et al.
Kitzbichler & White; **Merson et al. 2012**
Henriques et al. 2012

Step 4: Apply survey selection: “window function”



- Produce mock survey from idealised catalogue
- Apply angular footprint
- Apply scanning mask
- Incorporate sampling
- Redshift completeness
- Steps 1, 2, 3 generic
- Step 4 is survey specific

Mock catalogues need observables

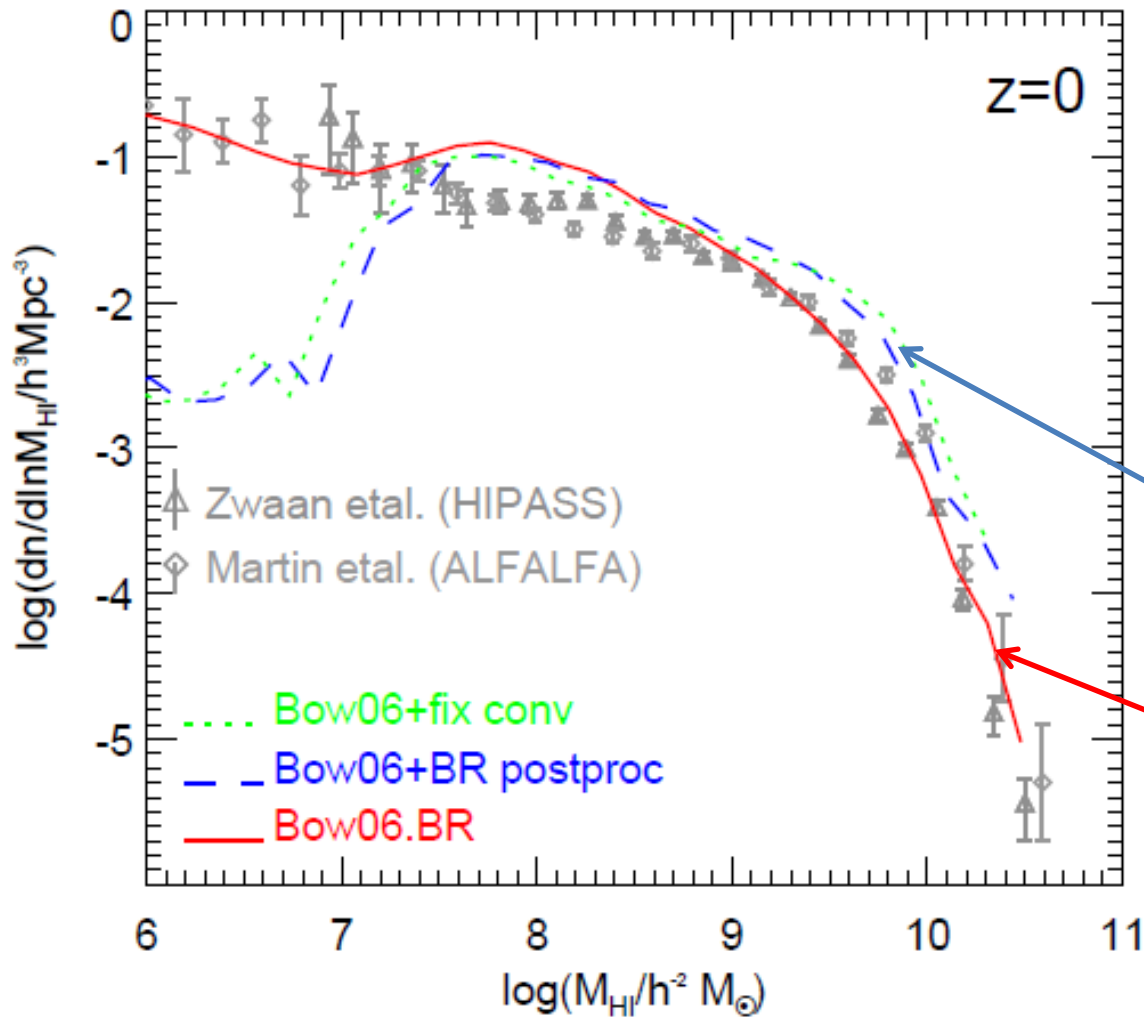


SFR
SFH
Stellar mass
Cold gas mass
Black hole mass



images
Full SED (UV, Optical, FIR, Radio)
Galaxies : stars, gas, AGN

The mass function of atomic hydrogen



New SF model
computes
H2/HI split

Old model + post-processing

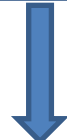
Full calculation

Lagos et al. 2011, 2012

Predict SFR



Use old formula



Apply
extinction?



Line
luminosity

Simple model for emission lines

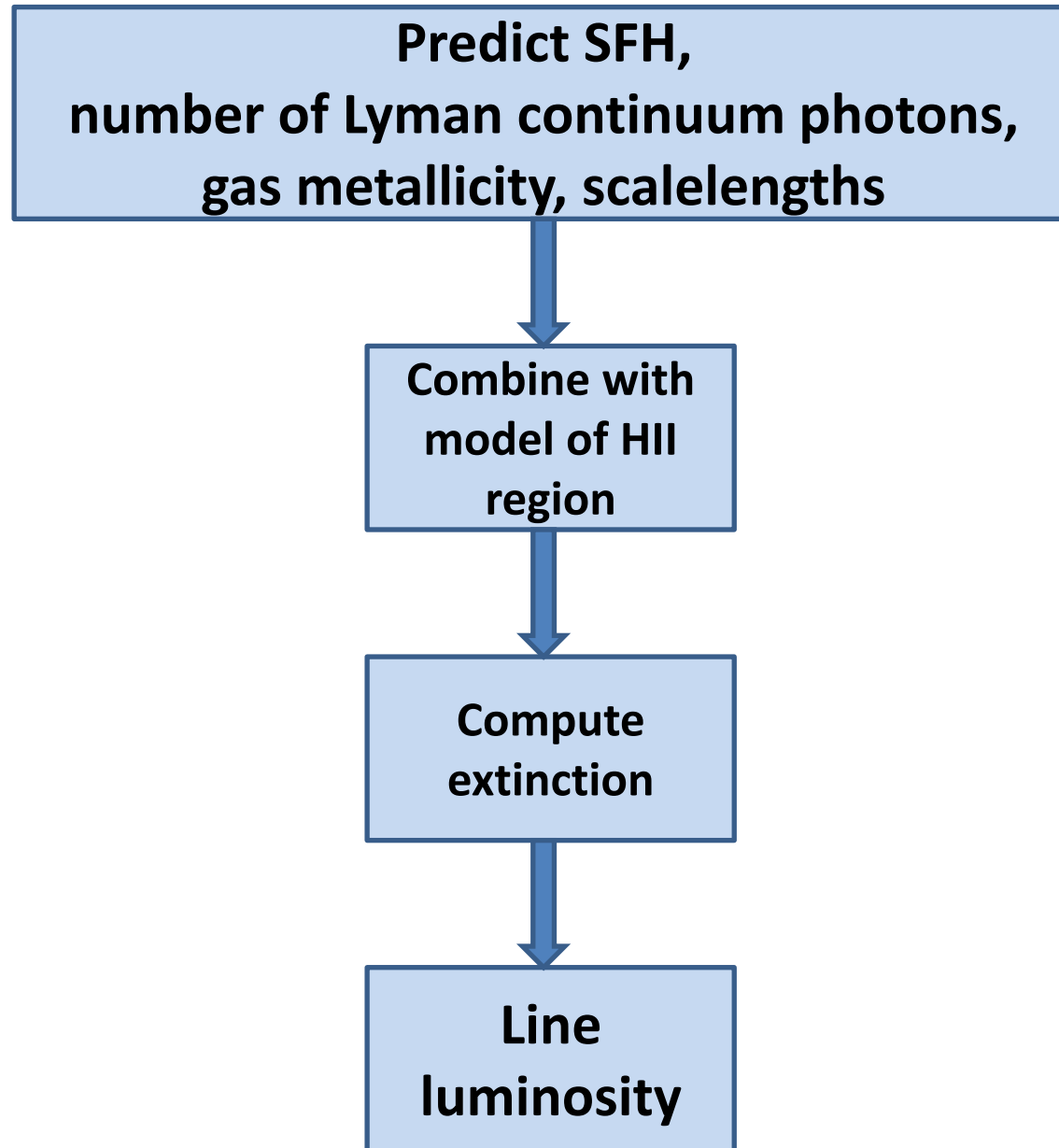
$$\text{SFR}(\geq 10 M_{\odot}) = \frac{L(\text{H}\alpha)}{7.02 \times 10^{41} \text{ ergs s}^{-1}} M_{\odot} \text{ yr}^{-1},$$

$$\text{SFR}(\text{total}) = \frac{L(\text{H}\alpha)}{1.12 \times 10^{41} \text{ ergs s}^{-1}} M_{\odot} \text{ yr}^{-1}.$$

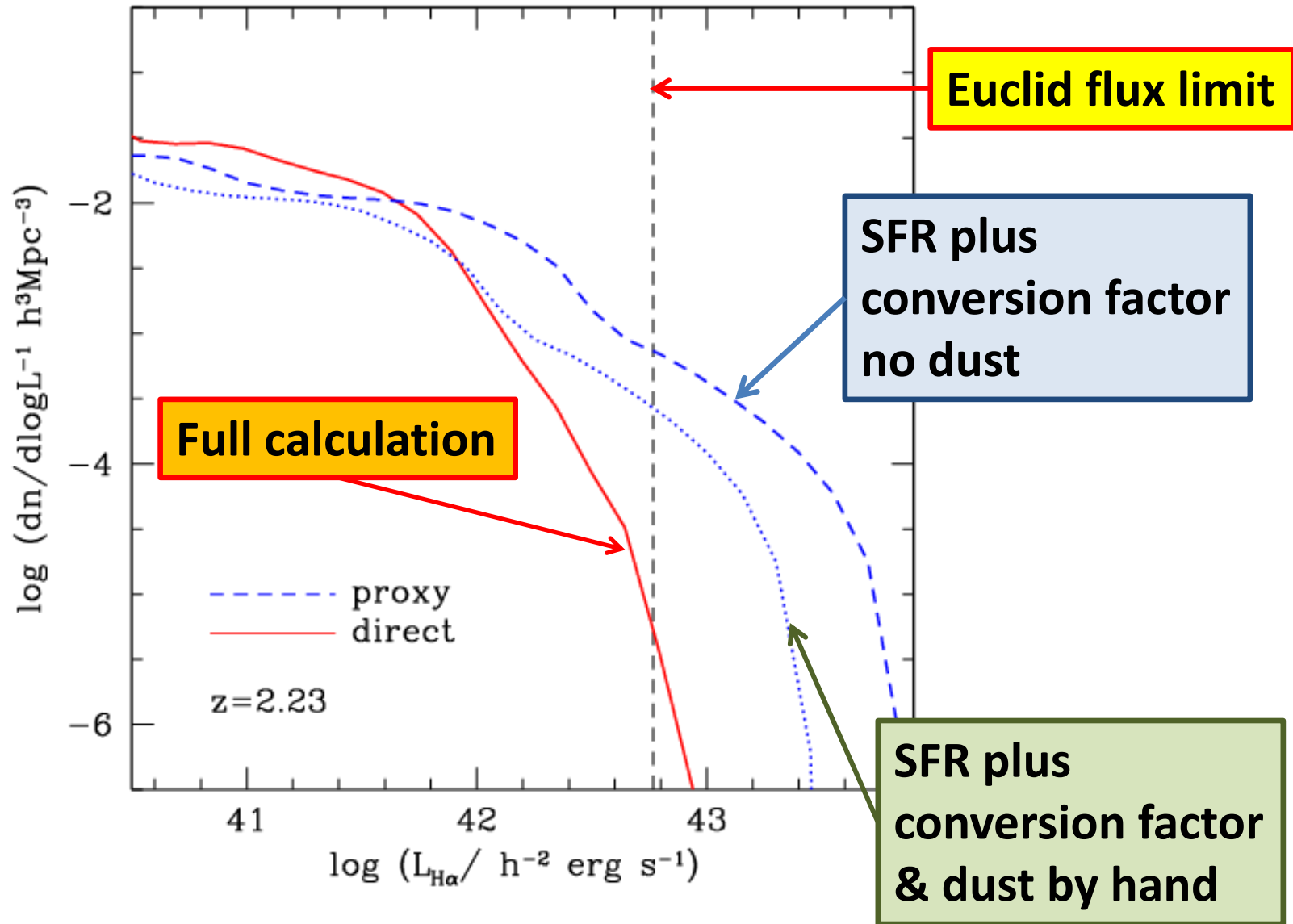
(Kennicutt 1983)

Balmer decrement?
1 magnitude?

Full calculation



Predicting H- α flux



Euclid OU-LE3 requirements for simulations



Cosmological
simulators

Instrument
simulators

Generic needs from Euclid

- Position, redshift
- Emission line properties/spectra
Line flux, equivalent width
- Broad photometry to AB \sim 24-24.5
Euclid NIR
Euclid VIS
Pan-STARRS griz
DES grizy
CFHTLS ugriyz
WFCAM ZYJHK
SDSS ugriz
VISTA-VHS-VIDEO ZYJHKs
- Photometric redshifts

Specific needs: clustering

- 1% $P(k)$ accuracy
- Covariance estimates: $P(k)$ etc
- Initial conditions for reconstruction
- Different cosmologies
- Different galaxy formation models (vary bias)

Specific needs – clusters of galaxies

- DM haloes $M > 1.e+13 M_{\text{sun}}$, $r(\Delta)$, $\Delta=2500, 500, 200$; velocity dispersion along axes from DM particles
- For each galaxy host halo ID, central or sat?
- Simulated images for cluster detection and mass determination through weak/strong lensing

Specific needs: weak lensing

- Galaxies and DM to generate kappa map
- Galaxy shapes with noise (no IA)
- Galaxy shapes with IA
- Shear at each galaxy position
- Image properties:
 - mask, bright stars, chip boundaries, CCD defects, ghosts, variations in depth & background

Specific needs: Time Domain

- Science areas: SNe, exo-planets, solar system
- Time series of VIS and NISP Y, J, H images
- Add fake transients with realistic properties (lightcurves, colours, positions) – properties to be supplied by Time Domain scientists
- Images should be pre-processed (bias corrected, flat fielded and sky subtracted). Individual dithered frames should be available so that an oversampled reference image can be constructed.

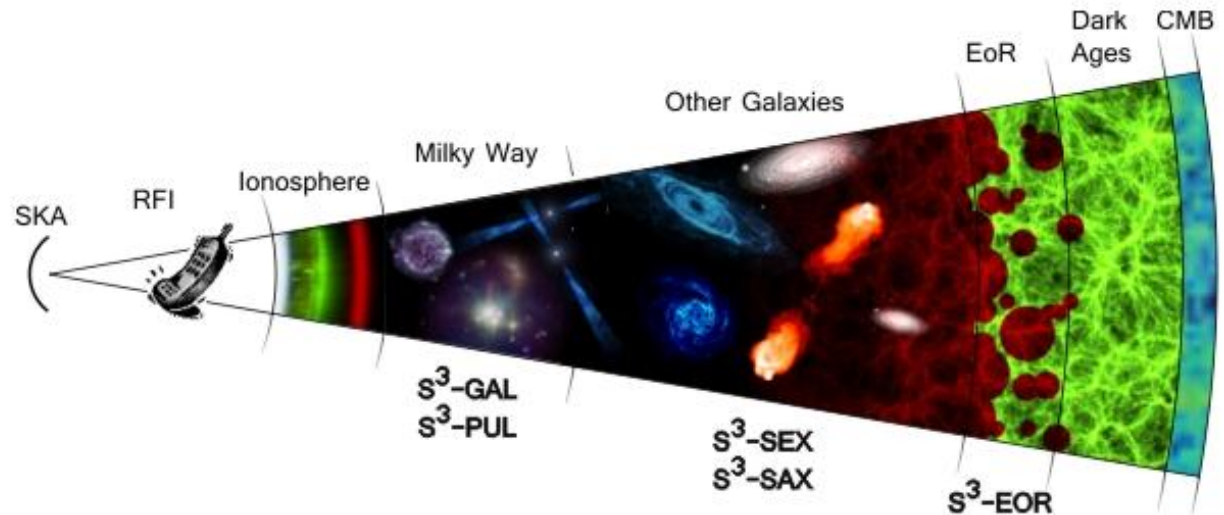
S³ The SKA Simulated Skies

- Home
- S³-SEX
- S³-SAX
- S³-PUL
- S³-GAL
- S³-EOR
- S³-Tools
- Contacts
- Admin

Introduction

The SKA Simulated Skies (S³) are a set of computer simulations of the radio and (sub)millimeter universe, primarily dedicated to the preparation of the Square Kilometer Array (SKA) and its pathfinders. These simulations were led by the University of Oxford as part of the Square Kilometer Array Design Studies (SKADS, [European Community Framework Programme 6](#), contract #011938).

Simulation types



S³ SAX-Sky Database query

- Home
- S³-SEX
- S³-SAX
 - Overview
 - SAX-Box
 - SAX-Sky
 - SAX-Sky query
 - SAX-Media
- S³-PUL
- S³-GAL
- S³-EOR
- S³-Tools
- Contacts
- Admin

SQL database query

Enter your query command, written in [SQL](#) (MySQL), in the text box below. Use the links in the grey box to get example queries.

NOTE: The SQL-server uses a time-out of 2 hours, which limits the complexity of your query and the size of the selected subset of galaxies. To obtain very large subsets, i.e. more than 10^6 galaxies, or the entire database, please contact us [here](#).

Limited select

- [Get the position, apparent redshift, and integrated HI-flux of the 10 closest galaxies in the cone](#)

HI-fluxes of a galaxy cluster

- [This query finds all the galaxies in the most massive cluster in the mock observing cone between \$z=1.2\$ and \$z=1.5\$. For each galaxy the output table gives the position, the apparent redshift, the integrated HI-flux \(Jy km/s\), the HI-peak flux density \(Jy\), the 50% HI-line width \(km/s\), the HI-half mass radius \(arcsec\), the inclination \(rad\), and the extinction corrected absolute blue magnitude. Only galaxies with stellar masses above \$10^9\$ solar masses are retained.](#)

dN/dz Analysis

- [This query evaluates the number of sources per square degree and unit of redshift \$z\$ with an HI-peak flux density above 1 micro Jansky. The evaluation is based on the central \$0.6 \times 0.6\$ square degrees of the mock observing cone and redshift bins of \$dz=0.1\$. The result is given for the range \$z=0-4\$](#)

Optically selected CO-sample

- [This query finds the position, the extinction corrected absolute blue magnitude, and the first five CO-line fluxes of the galaxies in the central square degree at redshift \$z=1.00 \pm 0.05\$, whose extinction corrected absolute blue magnitude is below -22.](#)

Currently available mocks

- Merson et al 2012 arXiv:1206.4046
lightcone catalogues using Millennium
- Euclid mock with H-alpha plus optical and NIR photometry available (600 Mb):

100 sq deg, H-alpha, Euclid NIR, u, g, r, l, z , R, U, V

catalogues limited by H-alpha flux or Euclid NIR photometry

Files:

<http://astro.dur.ac.uk/~d40qra/lightcones/EUCLID>

README:

http://astro.dur.ac.uk/~d40qra/lightcones/EUCLID/EUCLID_100sqdeg.lightcone.readme.pdf

Challenges for next generation mocks

- **Multiple large volume, high resolution simulations – covariance, different cosmologies**
- **Better modelling of observables:**
SFR, M^* , M_{cold} - not good enough for mocks
Euclid, BigBOSS:
emission line redshift surveys
SKA: HI
- **Telescope simulation:**
 - optical, radio: contract out?
- **Serving large catalogues:**
 - different database structures?
 - remote analysis?