Simulations for the SKA

Steve Rawlings on behalf of

Aris Karastergiou, Sadegh Khochfar, Hans-Rainer Klöckner, Francois Levrier, Tom Mauch, Lance Miller, Danail Obreschkow, Richard Wilman (Oxford)

Matt Jarvis (Herts),

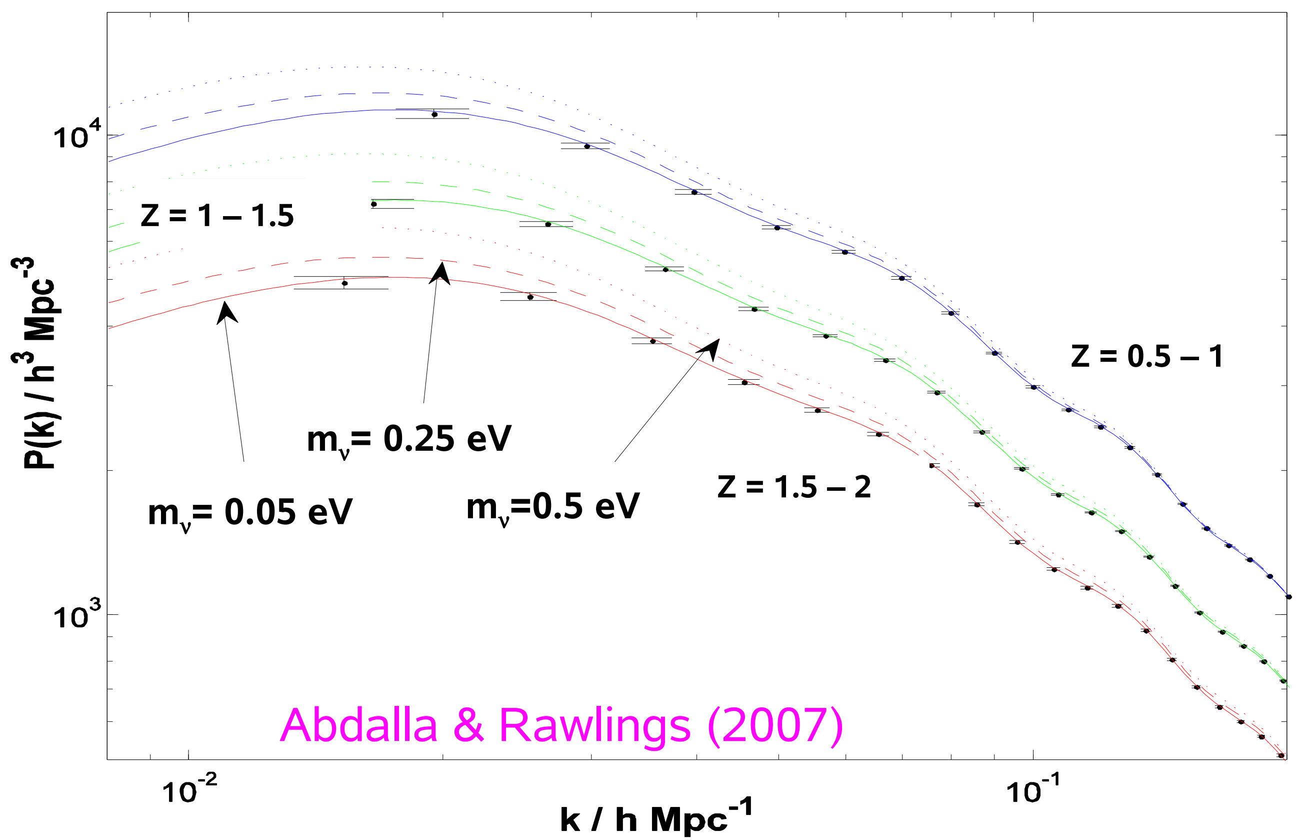
Filipe Abdalla (UCL) + Oxford e-Research Centre

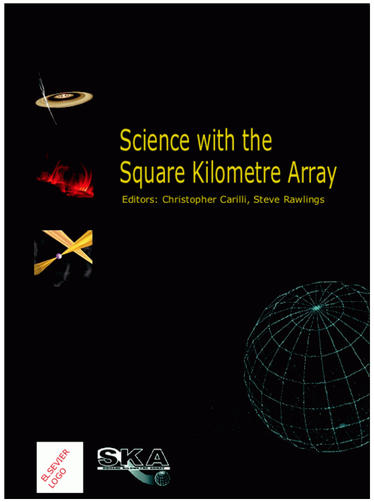
and the SKADS DS2-T1 and DS2-T2 teams

Why simulate the radio sky

- The need for sophisticated simulations
  - demonstrate SKA can beat systematics
  - hence drive technologies that generate transformational SKA science
- SKADS
  - what's been done 2006-2008 and how to use it
  - what will be done before mid-2009
  - what won't be done by then, and future plans

‘Back of Envelope’ Calculations



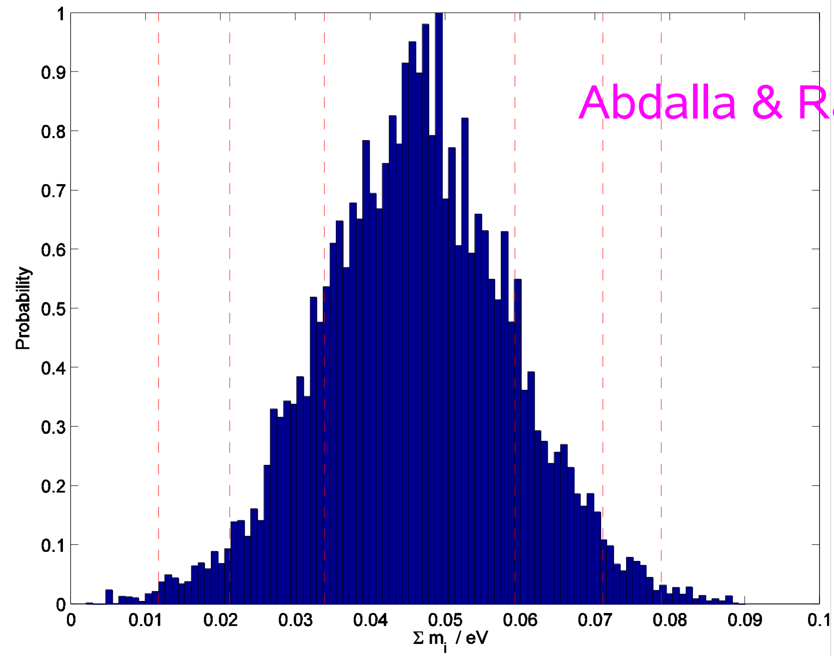


Carilli & Rawlings (2004)

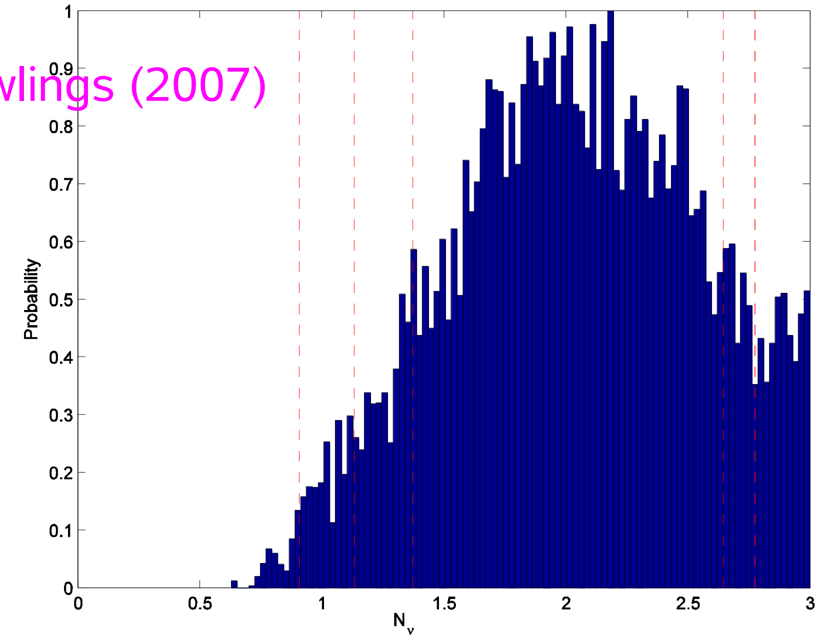
Important to keep adding to the science case

Measurement of Neutrino masses and hierarchies

SKADS



Abdalla & Rawlings (2007)

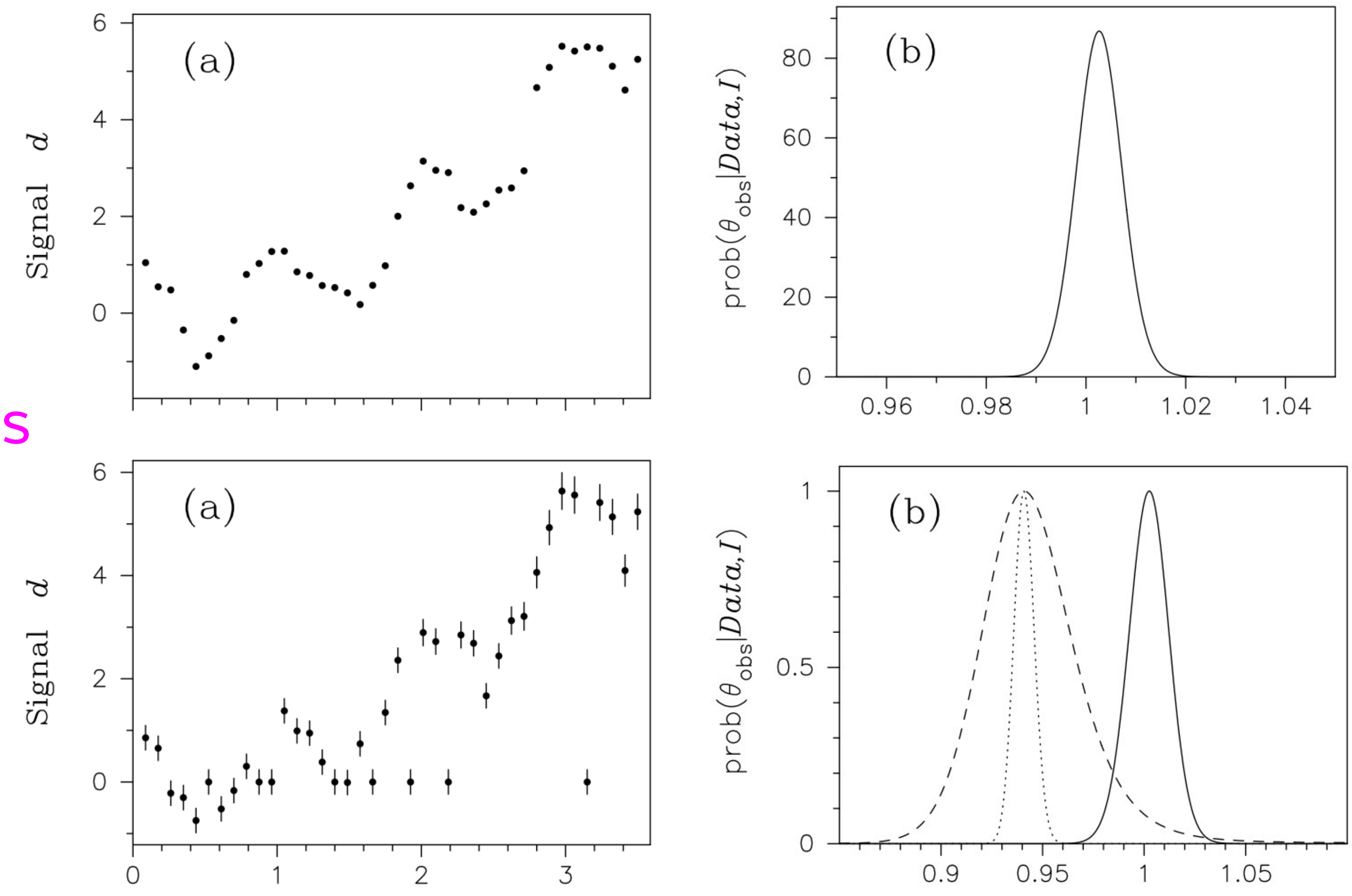


- If Σm_ν is ~0.05 eV, SKA+Planck sufficient and necessary to measure it, and hierarchy must be normal. If $\Sigma m_n > 0.25$ eV, measurement of N_ν too!
- With particle physics experiments, prospect of evidence for sterile neutrinos or time evolution of neutrino mass. But all based on `perfect' (10σ catalogue) SKA survey: no longer adequate

Oxford Physics

Simple (Gaussian error) `e2e' Simulations now inadequate

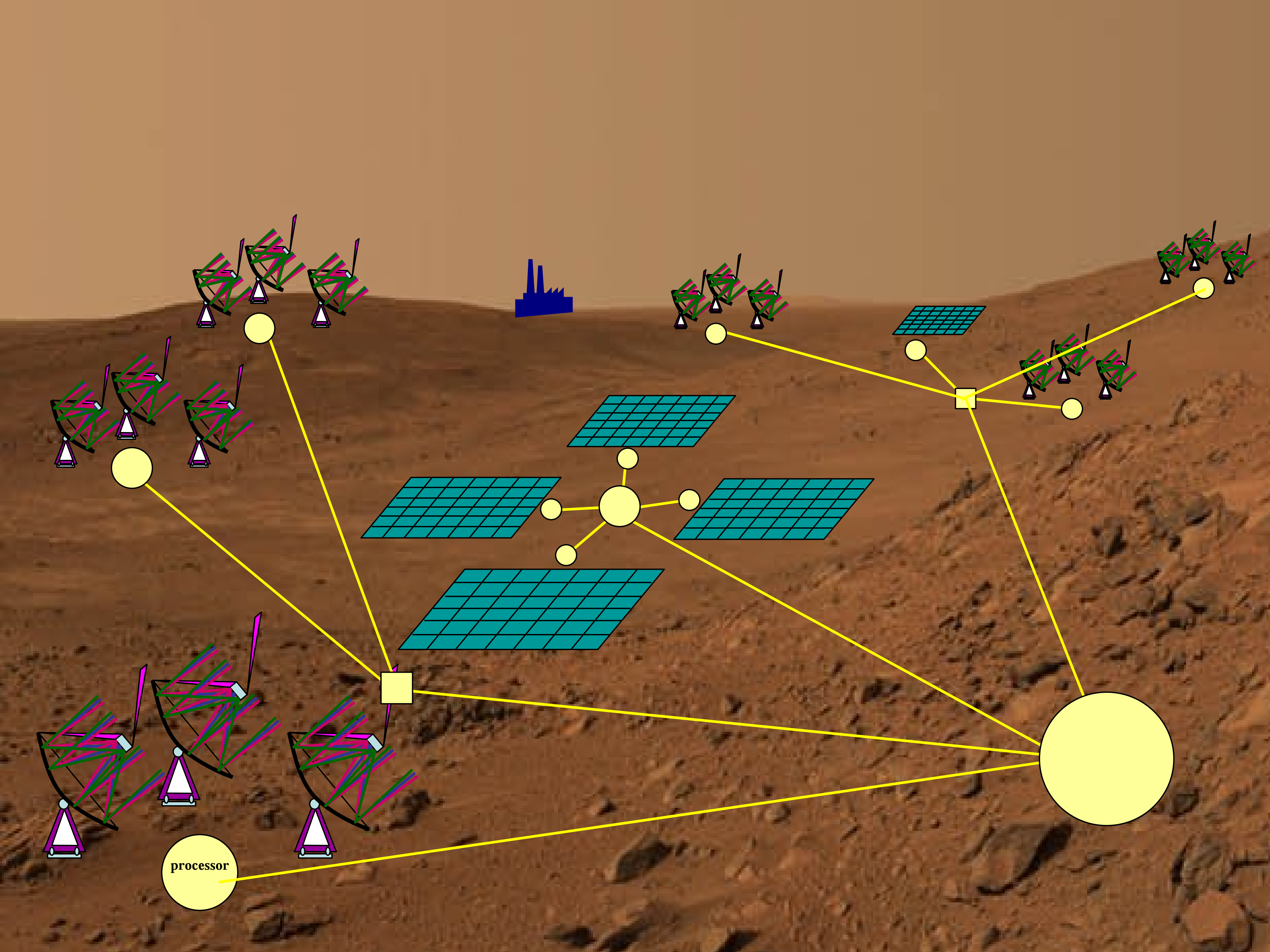
Sivia & Rawlings (2008)



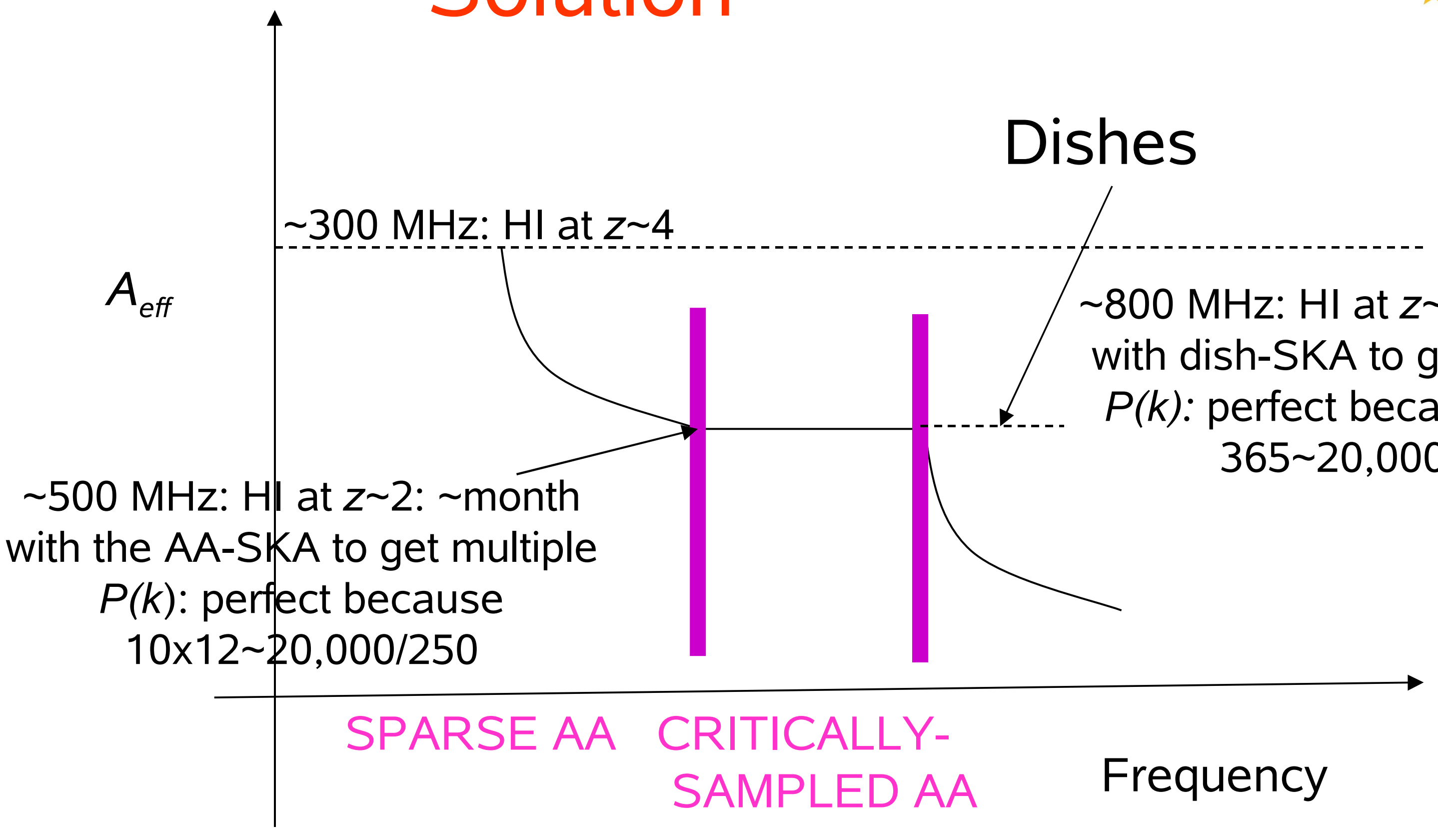
k *w*

Also competitive environment is fast developing

- Spectroscopic All-sky Cosmic Explorer: 5 billion redshifts over 3/4 sky to z~2.5, multi-slit near-IR
- Competition/Synergies: DUNE+SPACE (EUCLID): M-class (300 MEuro), could be as early as 2017


processor

Get the Best Possible Technology Solution



- Collecting area below 500 MHz rises as $(1+z)^2$ for sparse arrays, but only worth going to z~4 because of sky temperature (Braun 2006)
- Use multi-beaming to 'shape' FOV $\alpha\ (1+z)^4$ gives ~constant mapping speed, and a plausible FOV at the lowest frequency

Technology and the Phased Approach

Cosmic volume ~ (time) x (FOV) x $(A/T)^2$ for radio HI

If we relied on observing time, errors on *P(k)* would improve by a factor 10 every 100 years!

Technology I: increase dish FOV by ~100 and `dedicate' array, ~10^5 galaxies to *z*~0.2 (1% SKA pathfinders HI *P*(*k*) by ~2010)

Technology 2: mass production to increase A/T by ~10. ~10^7 galaxies over 20,000 deg^2 to *z*~0.75 (10% SKA in dishes ~2015)

Technology 3: additional technology (500-800 MHz). Further increase in A/T and FOV. ~10^9 galaxies to z~2 (100% SKA in aperture arrays ~2020).

Technology 4: Benefit from increased processing power as you integrate (300-800 MHz). >10^{10} galaxies to z~4 (2020-2030)

Two types of SKADS extragalactic simulations

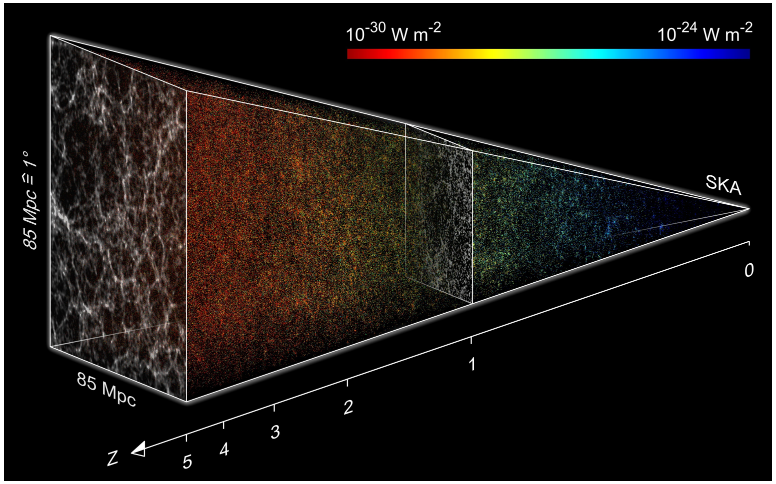
- ***'Line' semi-analytic approach (Obreschkow PhD): SAX.*** DM haloes from Millennium Simulation. Ascribed HI and H_2, star formation rates and AGN properties: provides insufficient FOV for SKADS benchmark

  4x4 deg² out to ***z~20 (HI/CO)***

  ~10^7 **objects**

  star-formation continuum OK

  AGN continuum still needs work



- ***'Continuum' semi-empirical approach (Wilman):SEX.***

  DM density field evolved under linear theory, populated with objects from known radio luminosity functions, and with other important physics (e.g. non-linear structures, source models) `pasted on'

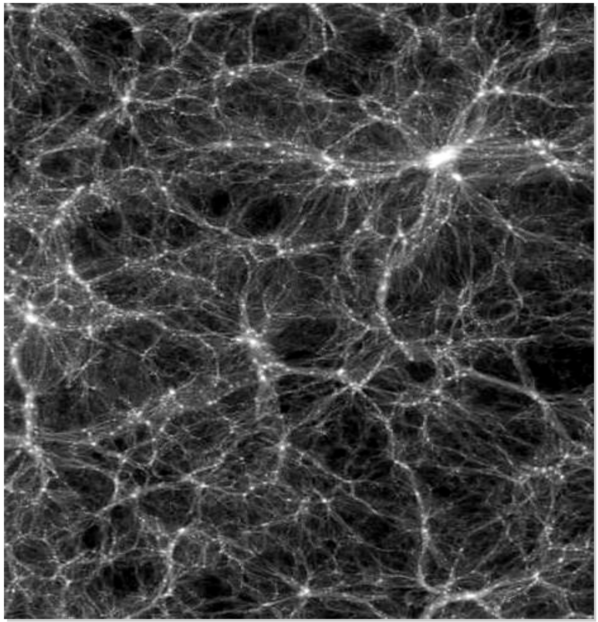
  currently 20x20 deg² out to *z*~20 , ~2.5×10^8 simulated sources

Semi-Analytic eXtragalactic (SAX) simulations

Millennium Simulation

(Springel et al. 05)

Dark matter

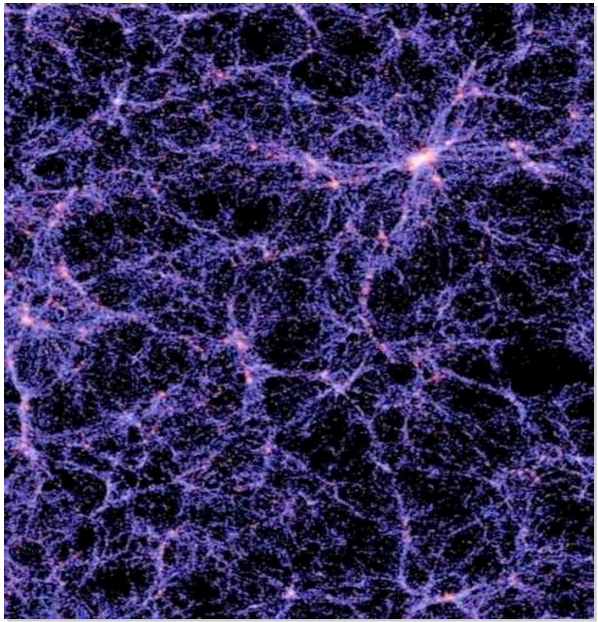


DM haloes, merger trees

Semi-analytics

(De Lucia et al. 06/07)

Visible matter

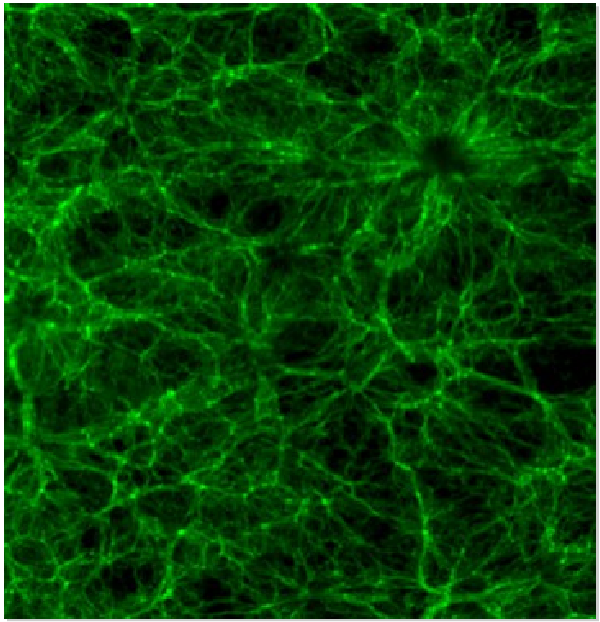


SFR, cold gas mass

Post-processing

(Obreschkow et al. 08)

Neutral atomic hydrogen



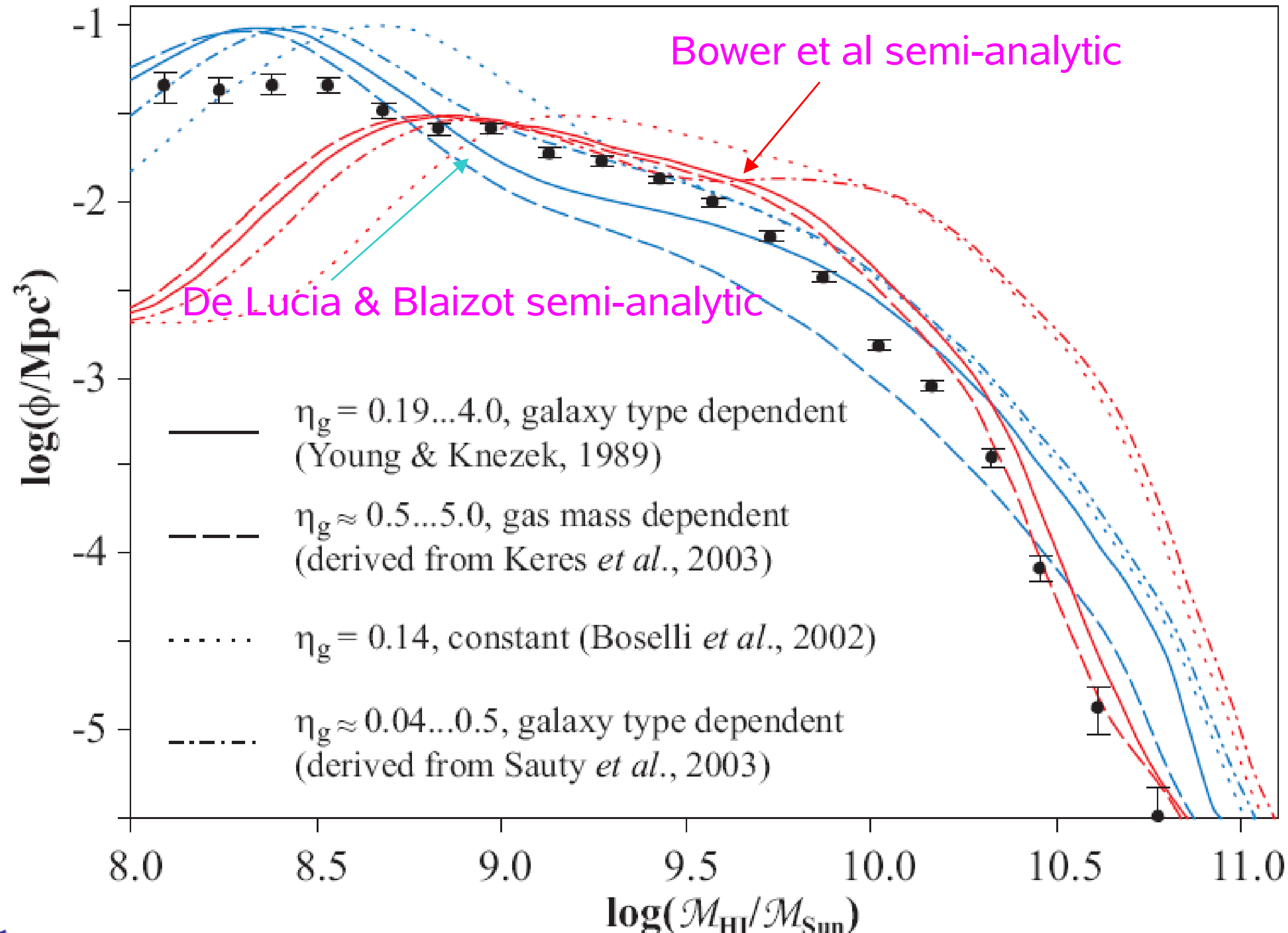
HI from cold gas mass

Key Issue

$$\eta \equiv \frac{\text{H}_2}{\text{HI}}$$

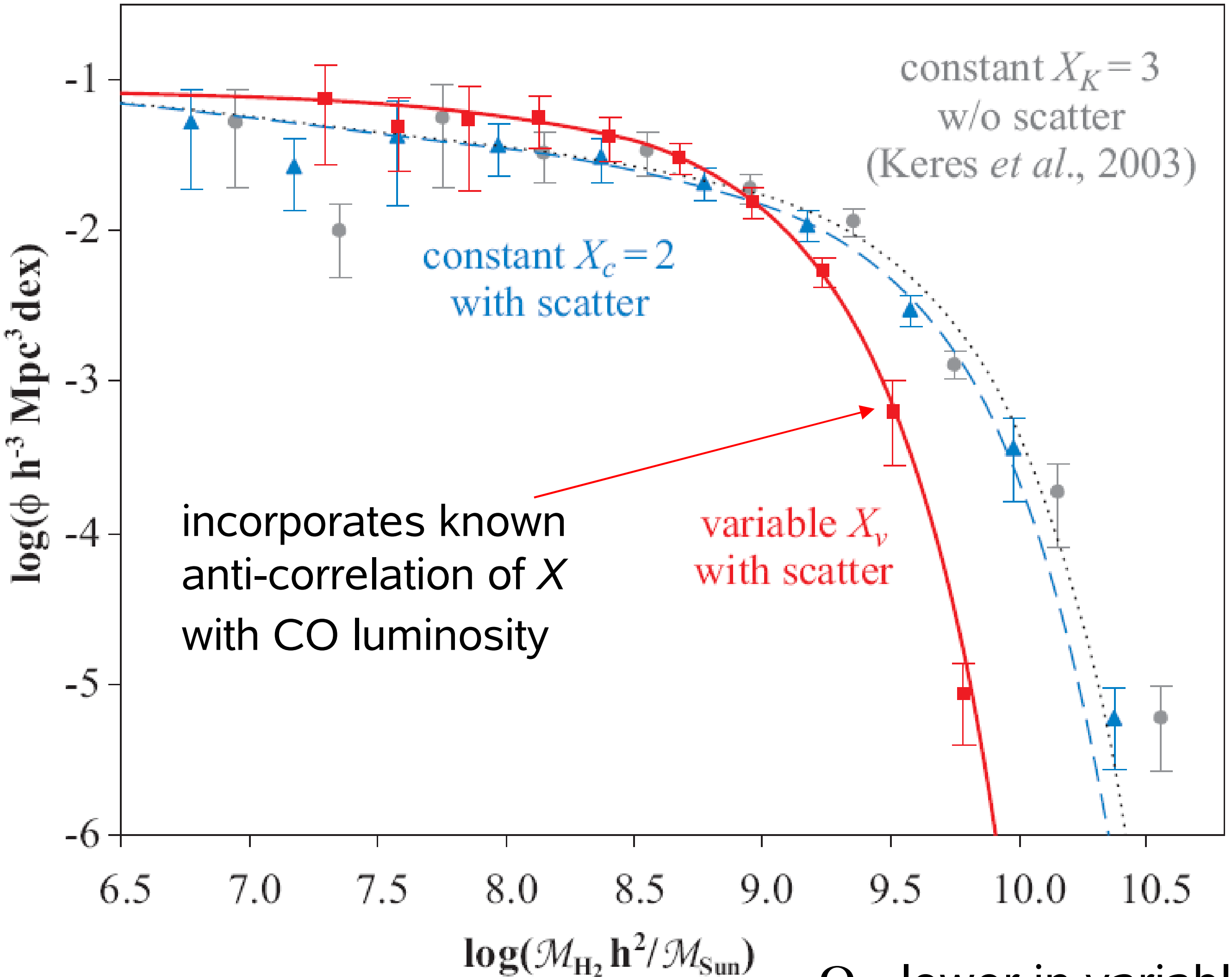
η_g is averaged over a galaxy

Constant η_g inadequate



Obreshkow et al. (2008)

H_2 from the CO LF

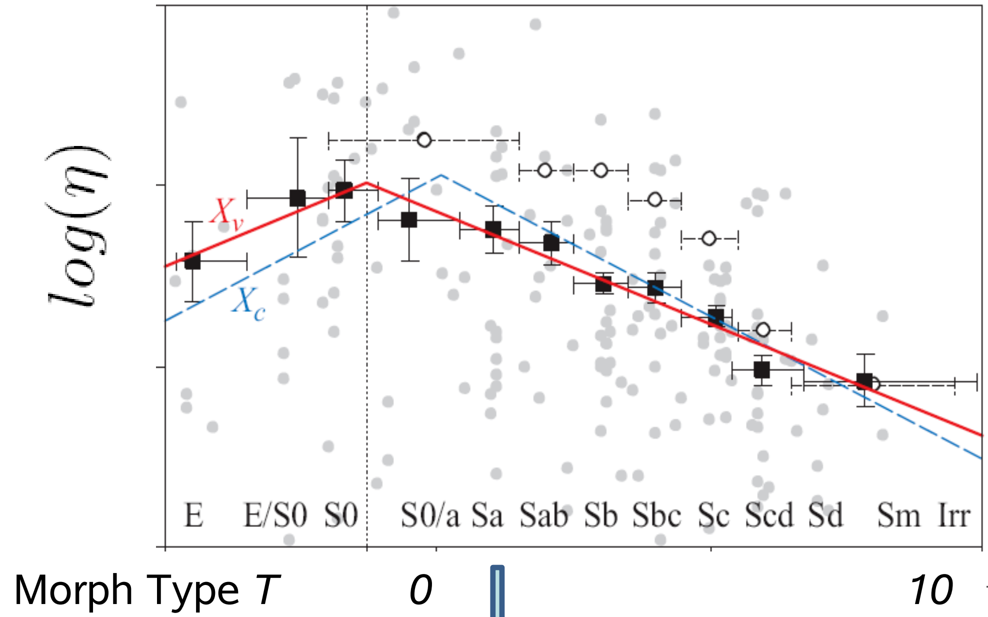


X encodes H_2 to CO ratio

Obreshkow et al. (2008)

Ω_{H2} lower in variable-*X* model

Models for η_g

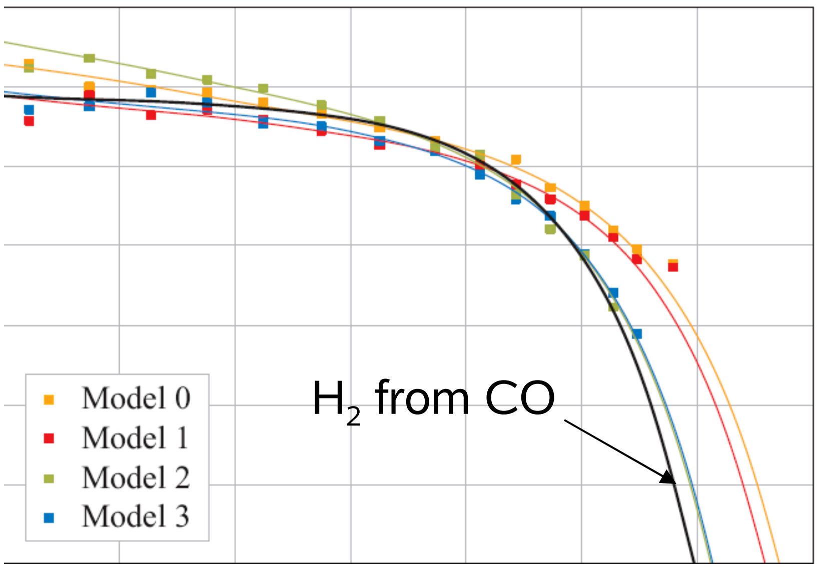


Obreshkow et al. (2008)

Driven by gas pressure

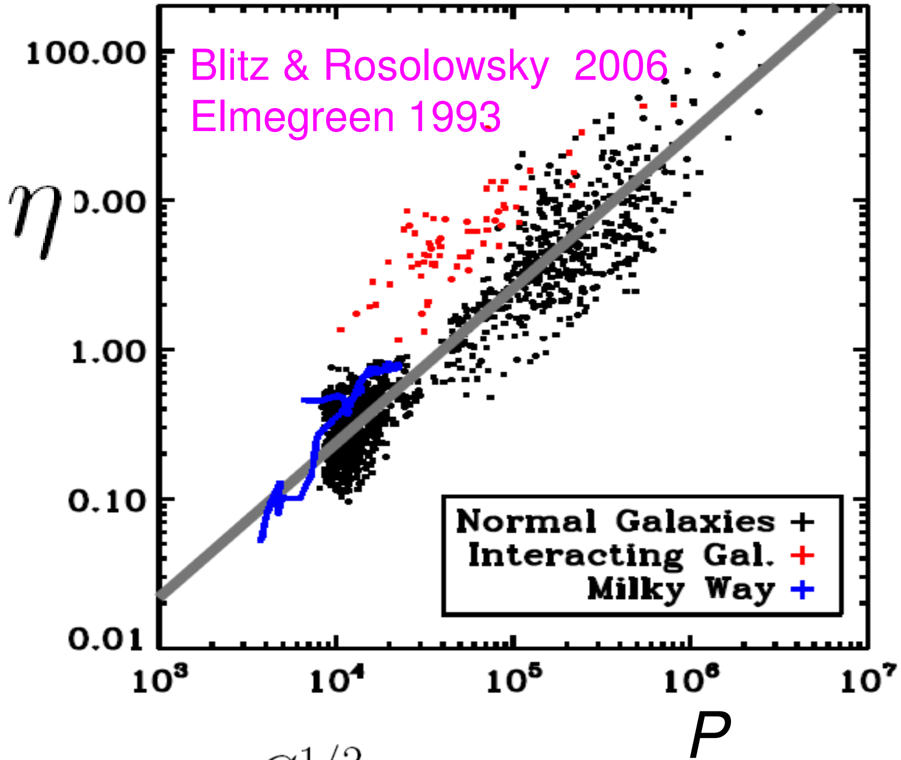
$$log(\eta) = c_0^{sp} + c_1^{sp} T + k_1 \, log(\mathfrak{m}) + \sigma$$

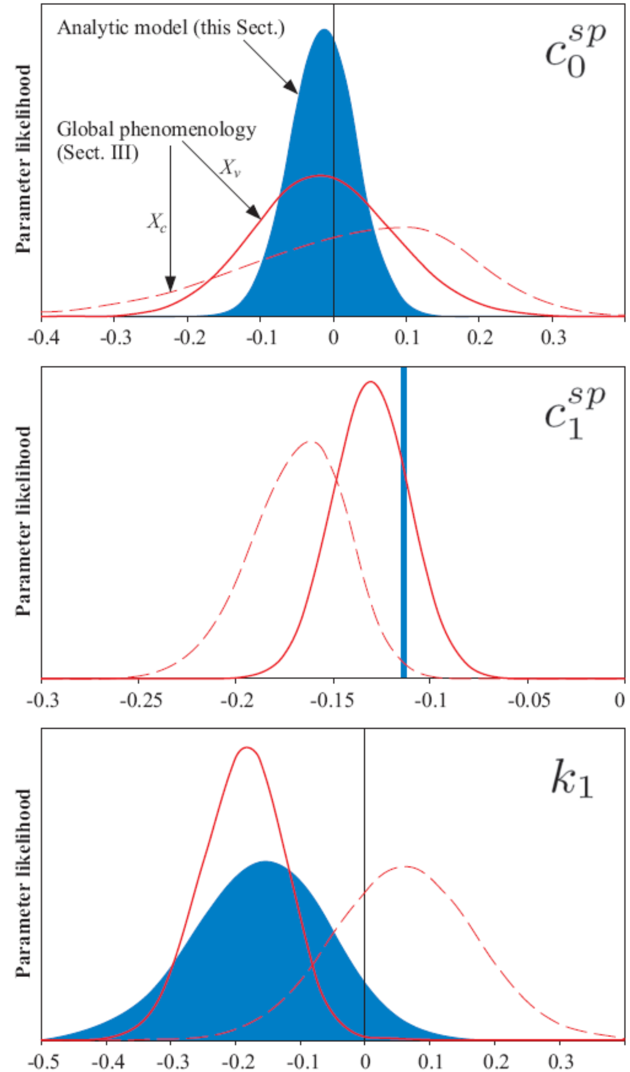
+HIPASS



Free parameters fixed using local sample

η_g is determined by pressure



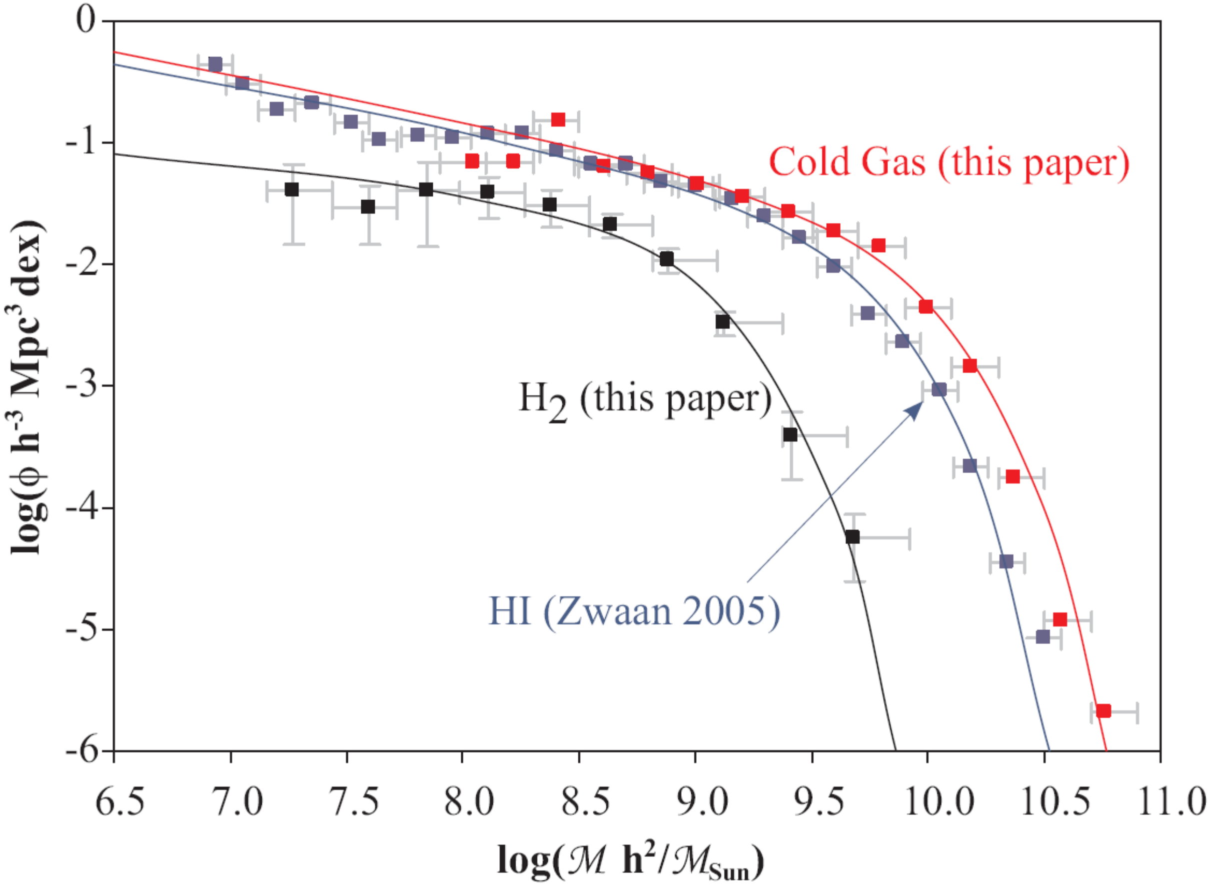


$$P(r) = \frac{G^{1/2}\, v_g}{2^{1/4}\, h_s^{1/2}}\, \Sigma_s^{1/2}(r)\, \Sigma_g(r)$$

$$log(\eta) = c_0^{sp} + c_1^{sp}\, T + k_1\, log(\mathfrak{m})$$

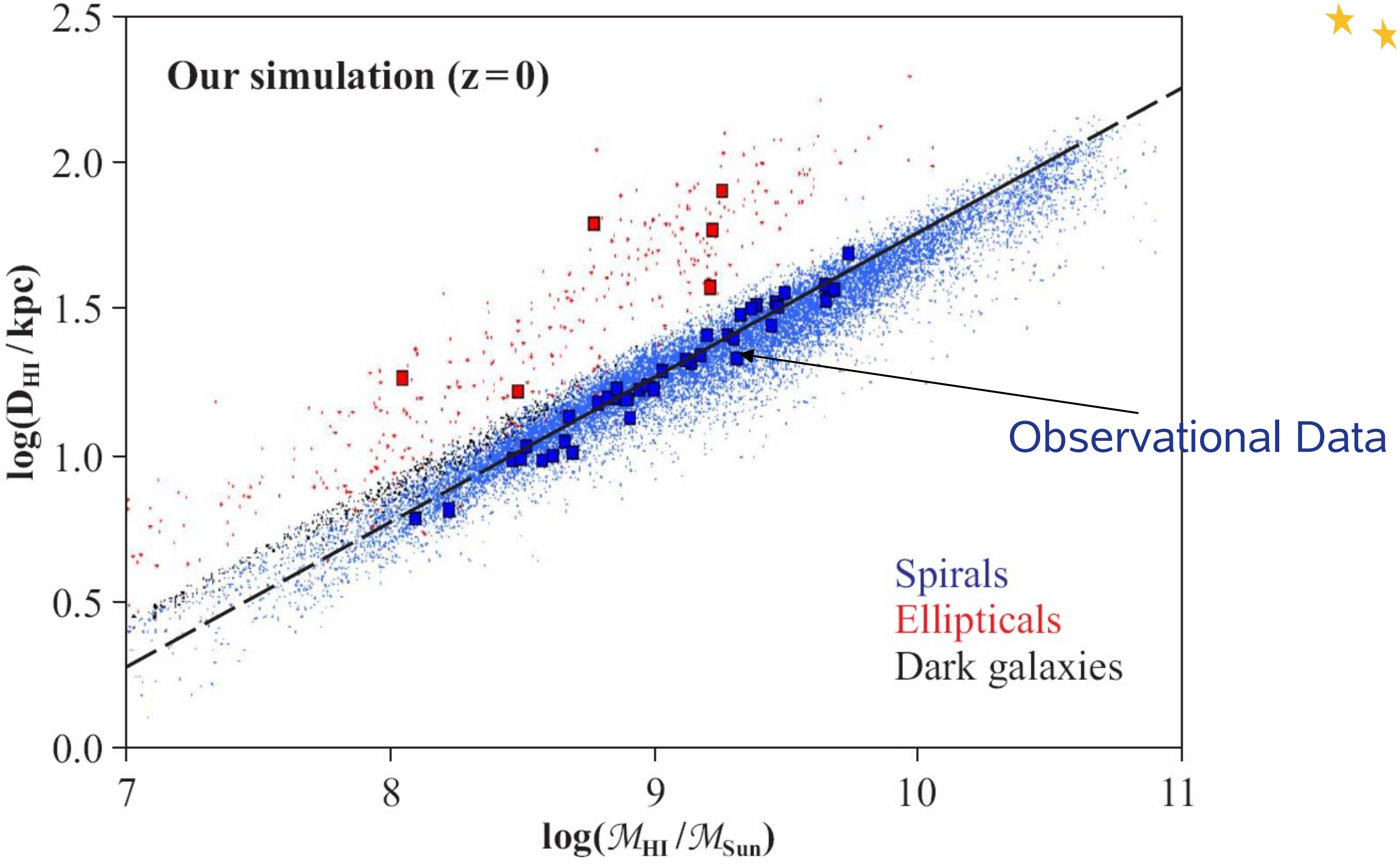
Obreschkow et al (2008)

Physical Model at z=0



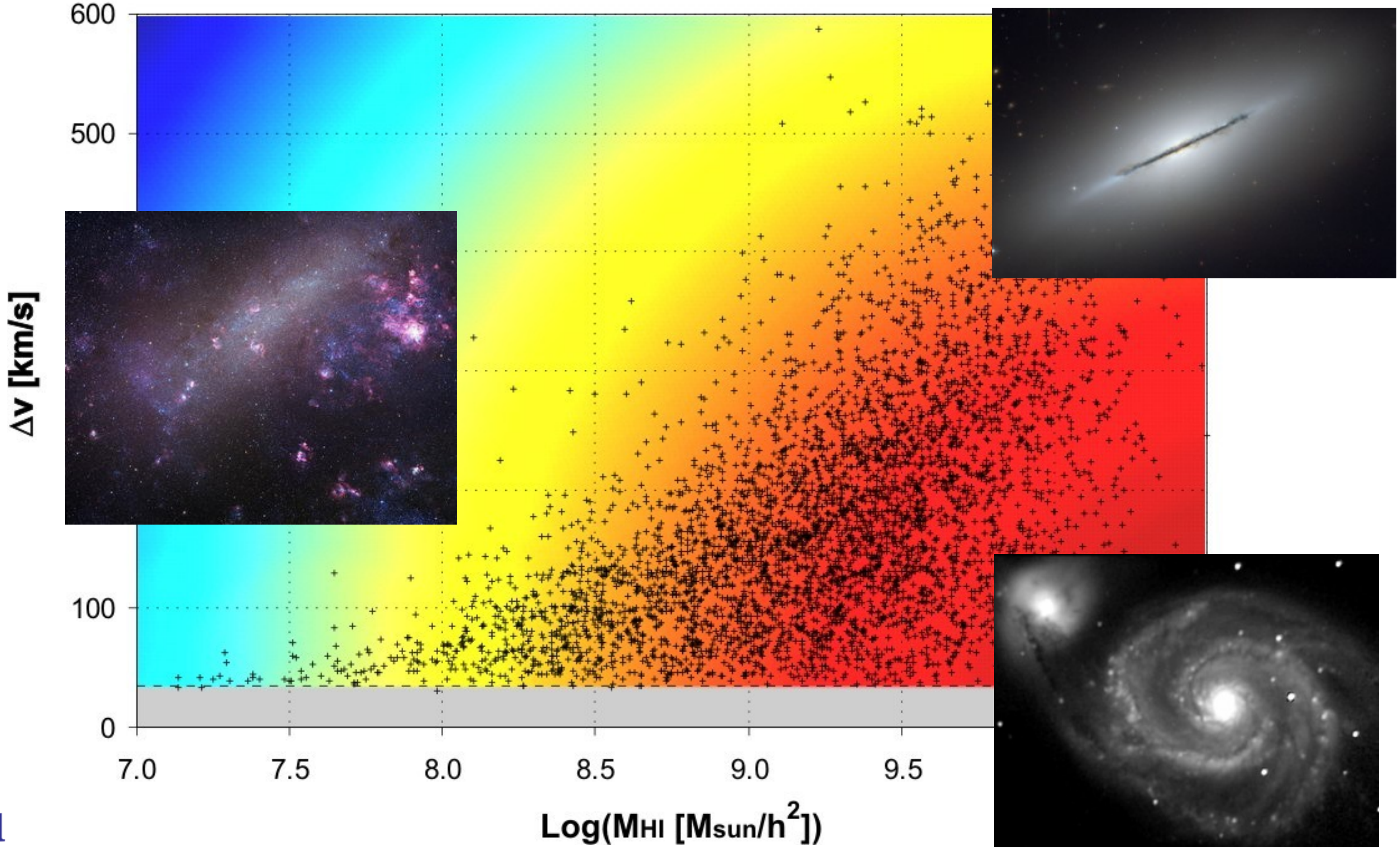
Obreschkow et al. (2008)

Model fits sizes



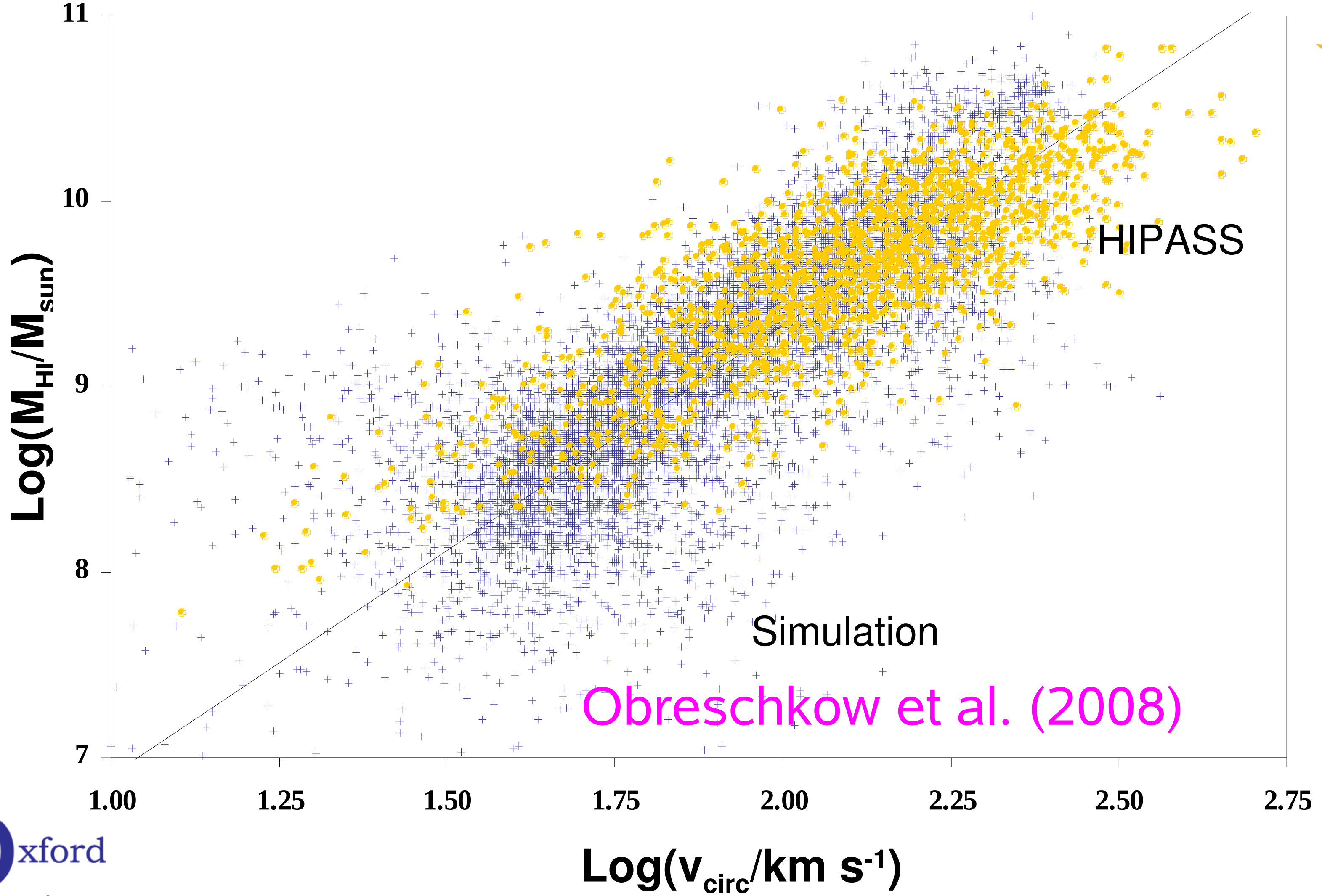
Obreschkow et al. (2008)

After Correction for inclination

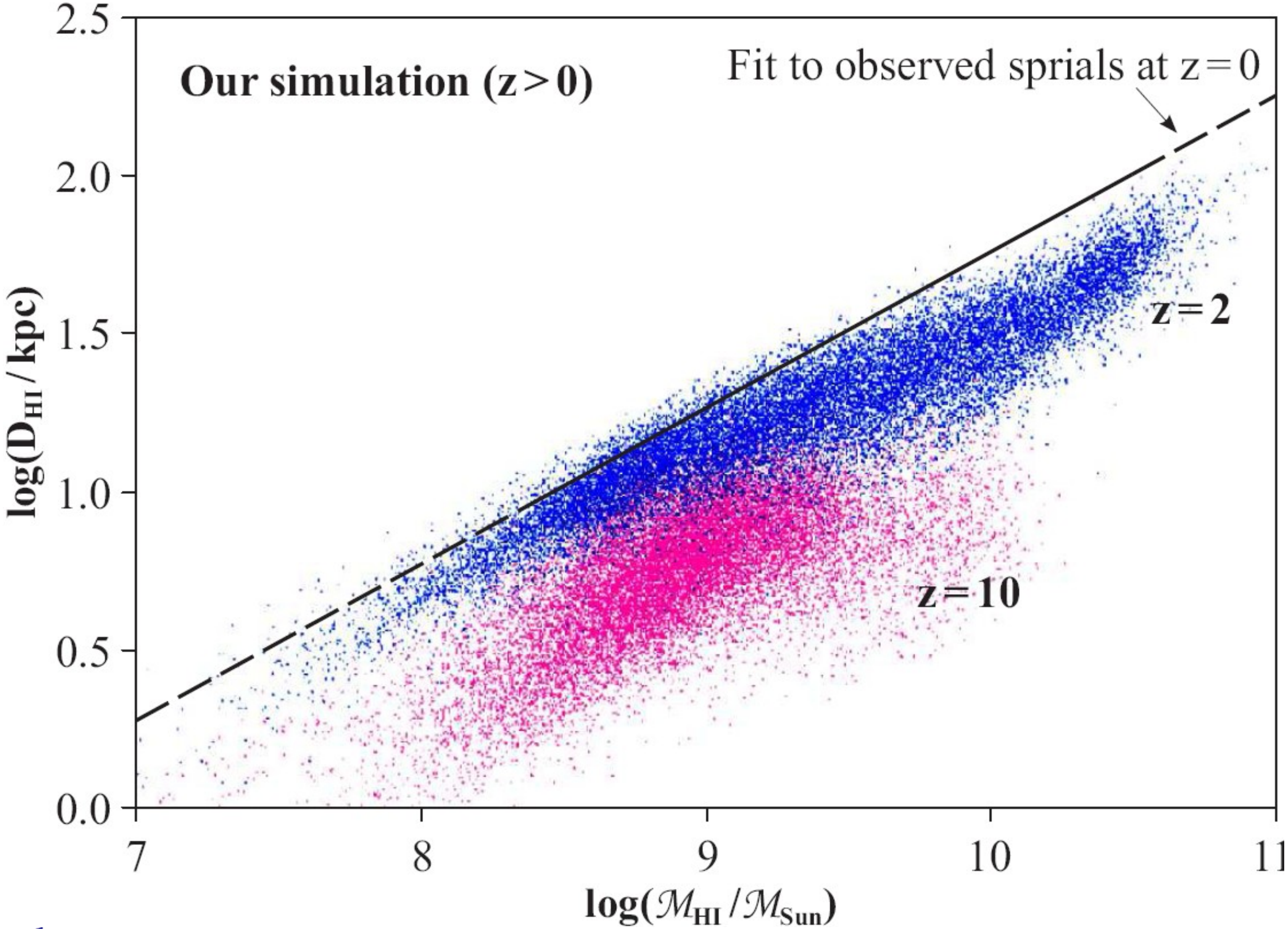


Obreschkow et al. (2008)

Model fits line widths



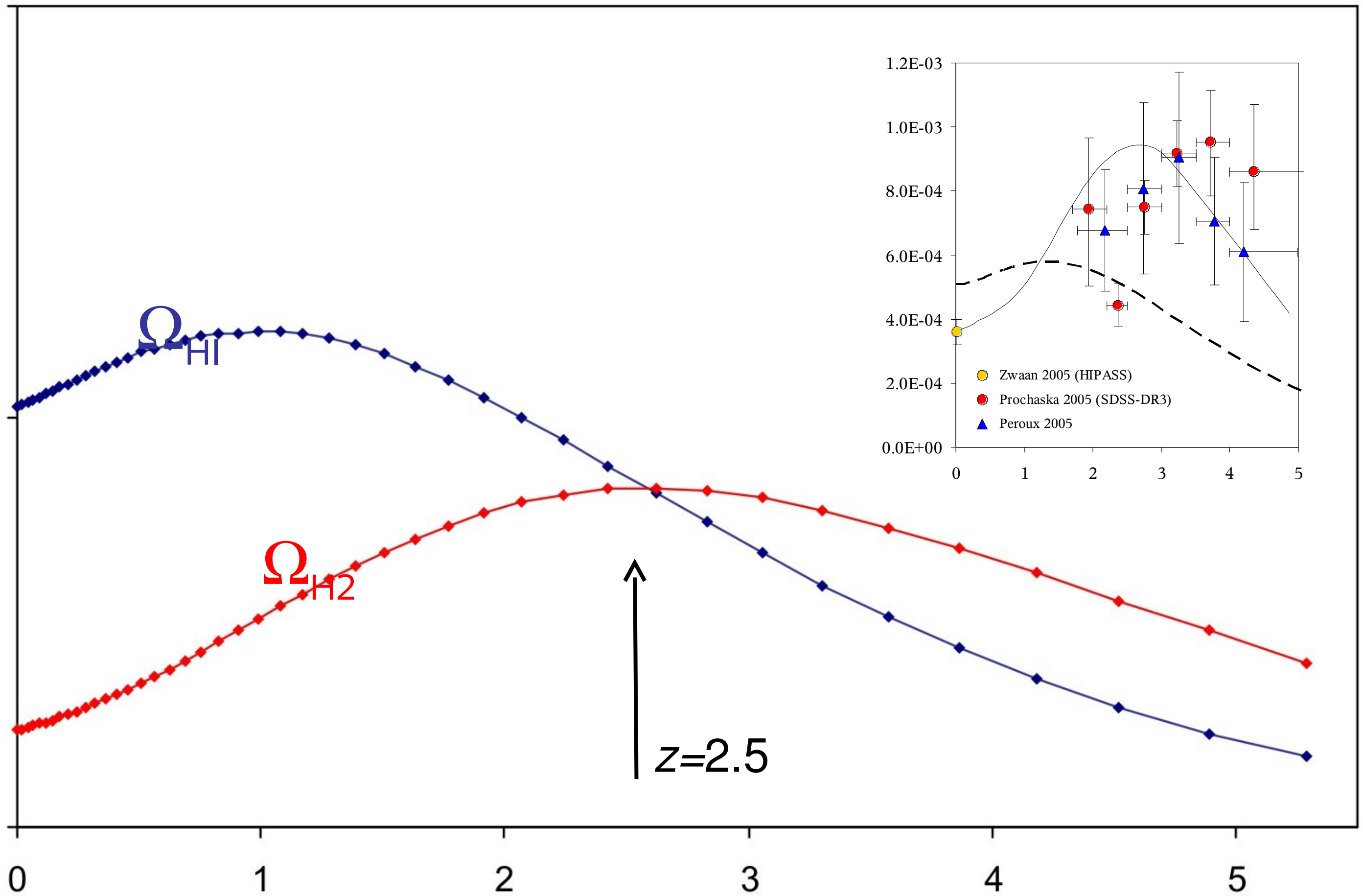
Predictive at high-z



Obreschkow et al. (2008)

Evolution of HI and H_2

SKADS

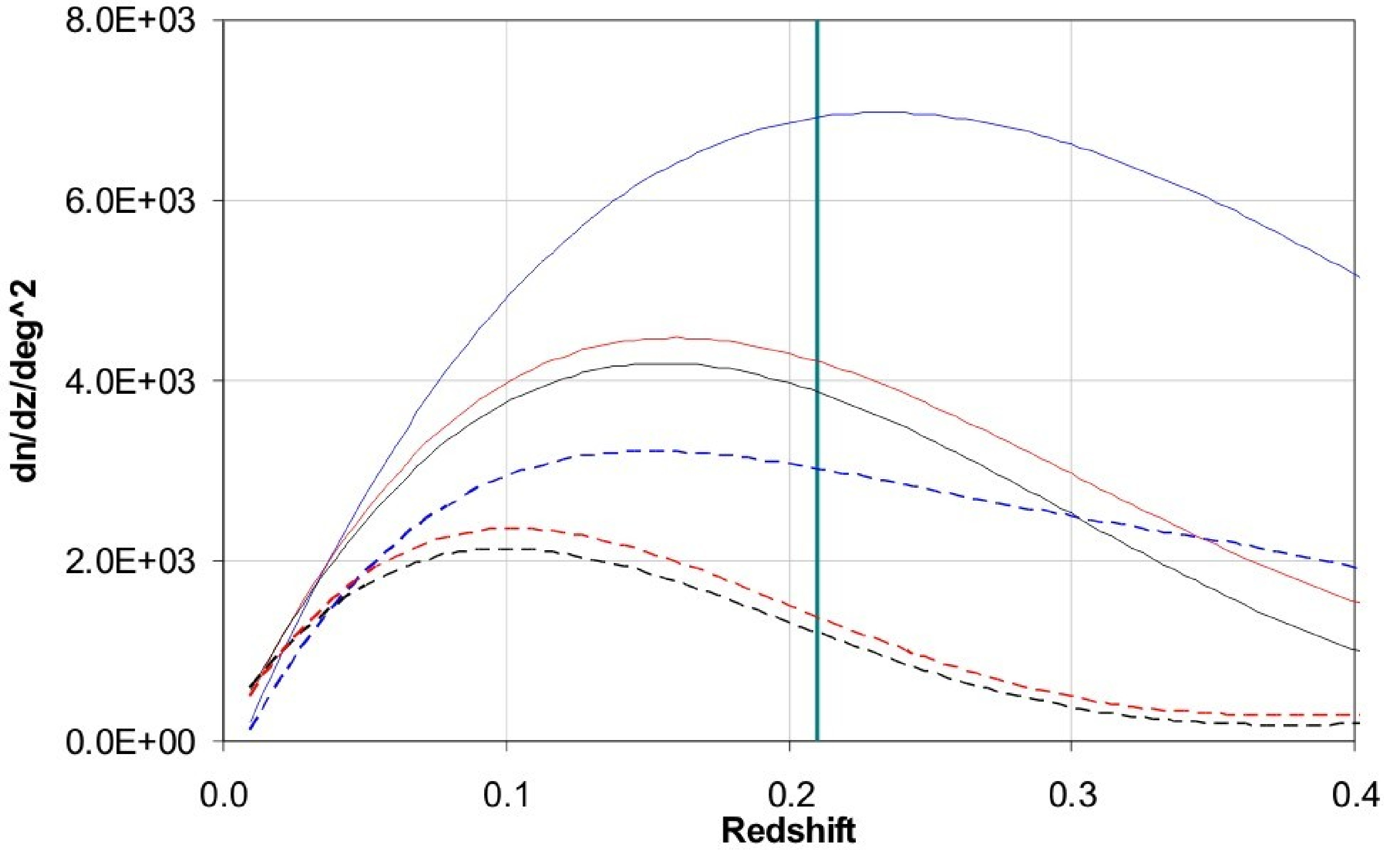


Note link with ALMA

Obreschkow et al. (2008)

Milestone 1: ~1% SKA (~2010)

SKADS



in ~100 days >10^5 galaxies: an ~1000 deg^2 survey to z~0.2;

gives *P(k)* and hence bias

Oxford Physics

Milestone 2: ~1% SKA+FMOS/AAΩ

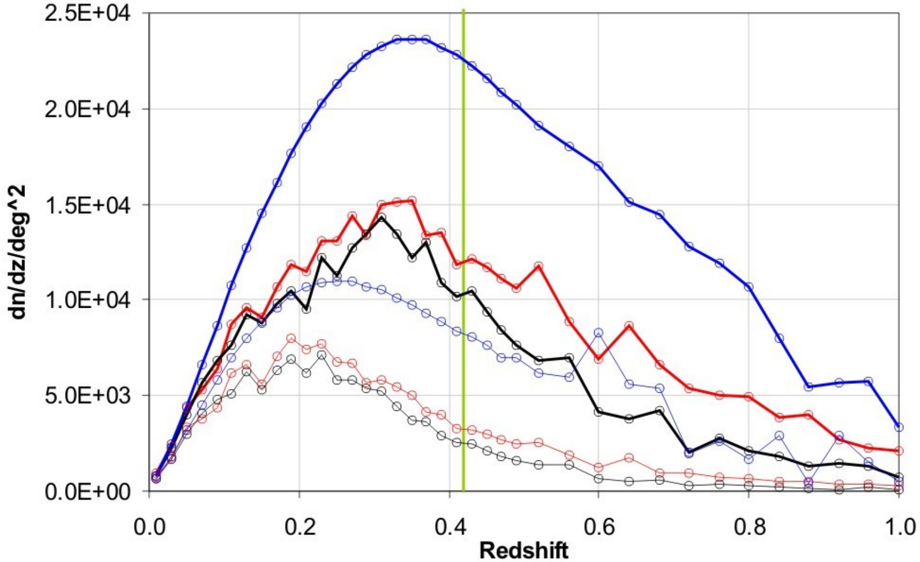
- 'Radio stacking' on optical/near-IR redshifts out to *z~1*

Wiggle-z: ~400 targets per deg^2 over large (~30 deg^2) patches: √12000~100 gain on star-forming subset of galaxies

FMOS: 12,000 deg^{-2}, √12000~100 gain on ALL galaxies in deg^2 patches

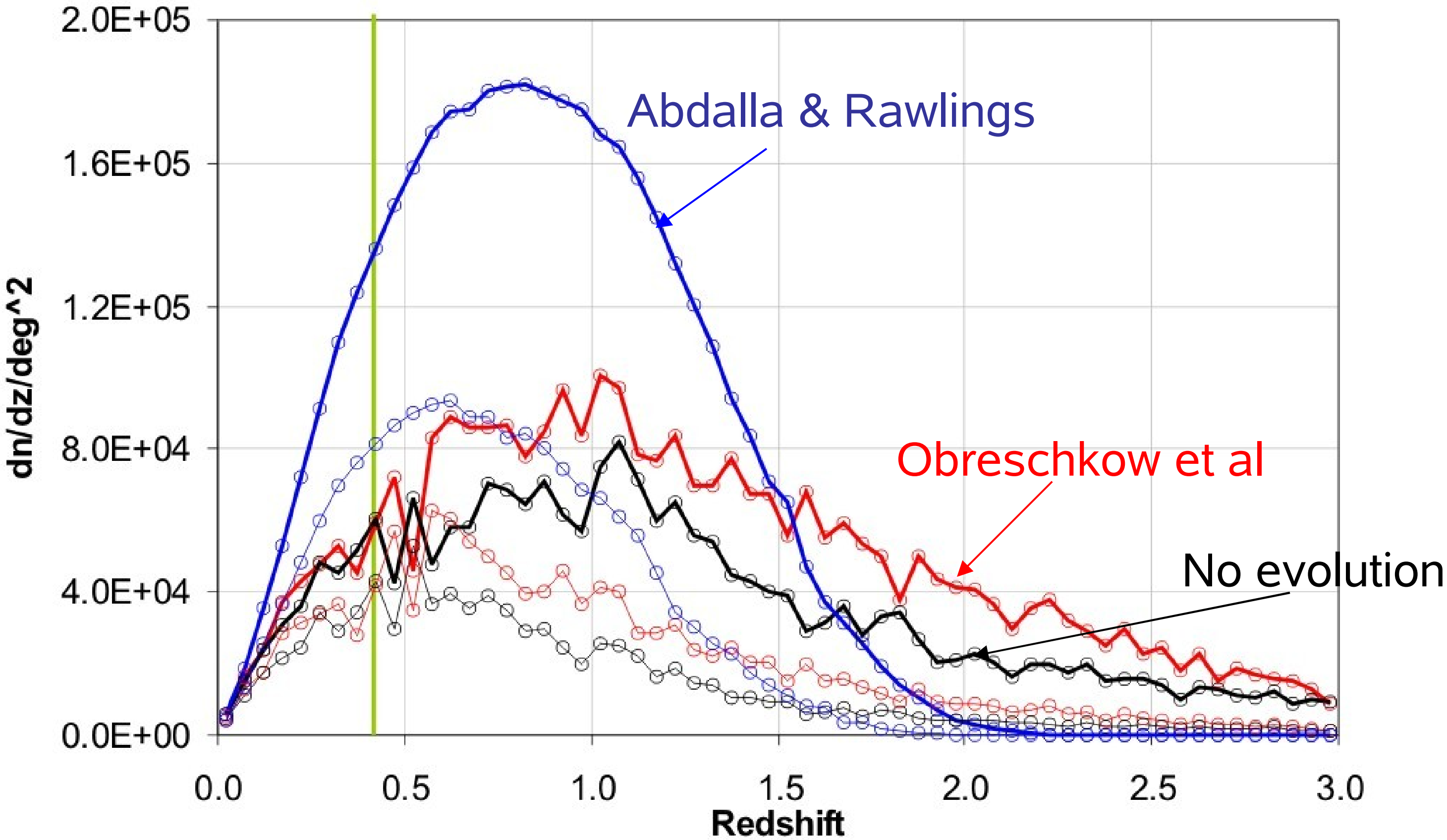
Should provide statistical detections of HI emission to *z~1* (alongside HI absorption experiments to higher *z)*

Milestone 3: ~10% SKA (~2015)



- Dishes: >10^7 galaxies over ~20,000 deg^2 to z~0.75 (or 800 MHz); multiple $P(k)$ to z~0.35, stacked analyses to z~2
- Likely to be similar cosmic volumes to optical redshift surveys (e.g WFMOS: combination very powerful for studying galaxy evolution and bias).
- Survey Volumes then grow, at best, linearly with t, hence $P(k)$ accuracy grows as $t^{0.5}$

Full SKA Predictions

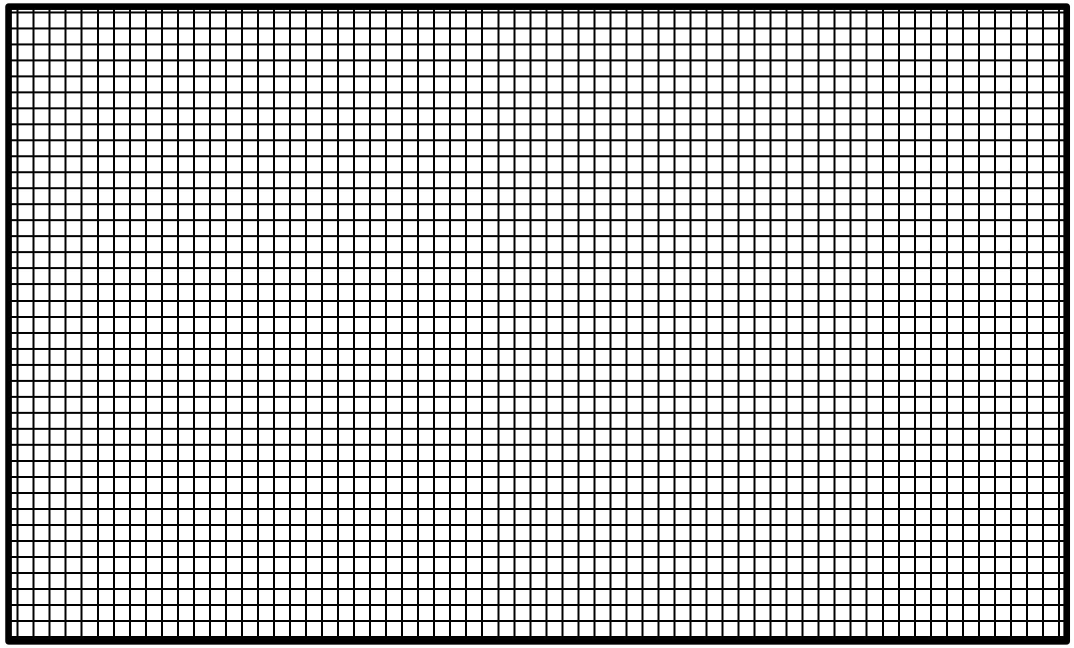


- In ~100 days, phased arrays deliver >10^9 galaxies over ~20,000 deg^2 to z~2 (and multiple *P(k)* to at least *z~1*)

Design requirements of the Semi-Empirical eXtragalactic (SEX) simulation

- Sky area: 20x20 deg^2
- Redshift range: z = 0 – 20
- Flux density limits: 10nJy @ 151MHz, 610MHz, 1.4 GHz, 4.86 GHz & 18 GHz
- Source populations: radio-quiet AGN
  FRI and FRII radio-loud AGN
  normal star-forming galaxies
  starburst galaxies
- Output: online database tool for creating ‘idealised skies’

The Theoretical Framework

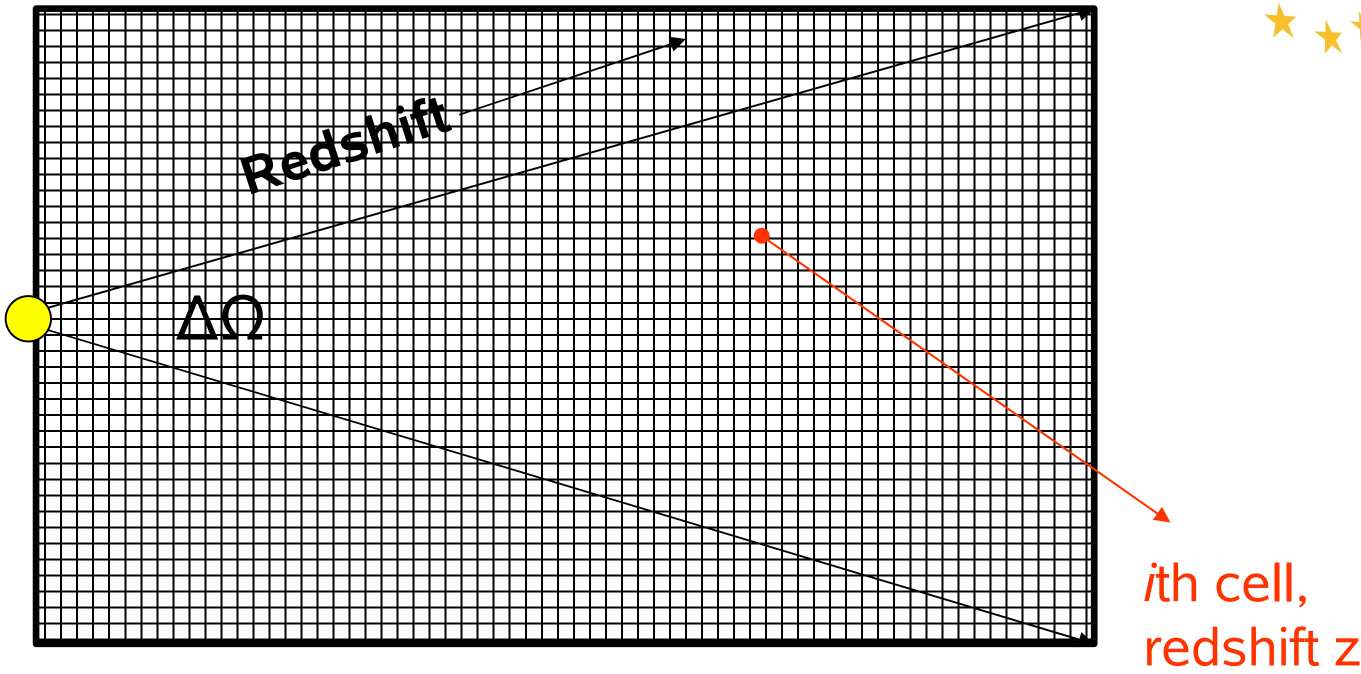


Starting point: a z=0 $(\delta\rho/\rho)_{DM}$ linear theory dark matter density field defined on a 550x550x1500 grid of 5 Mpc/h cells

Cosmology: H_0=70 km/s/Mpc, Ω_M=0.3, Ω_Λ=0.7, σ_8=0.74, BAO *P(k)*

Wilman et al. (2008)

The Theoretical Framework



In the *i*th cell, for each source population:

Wilman et al. (2008)

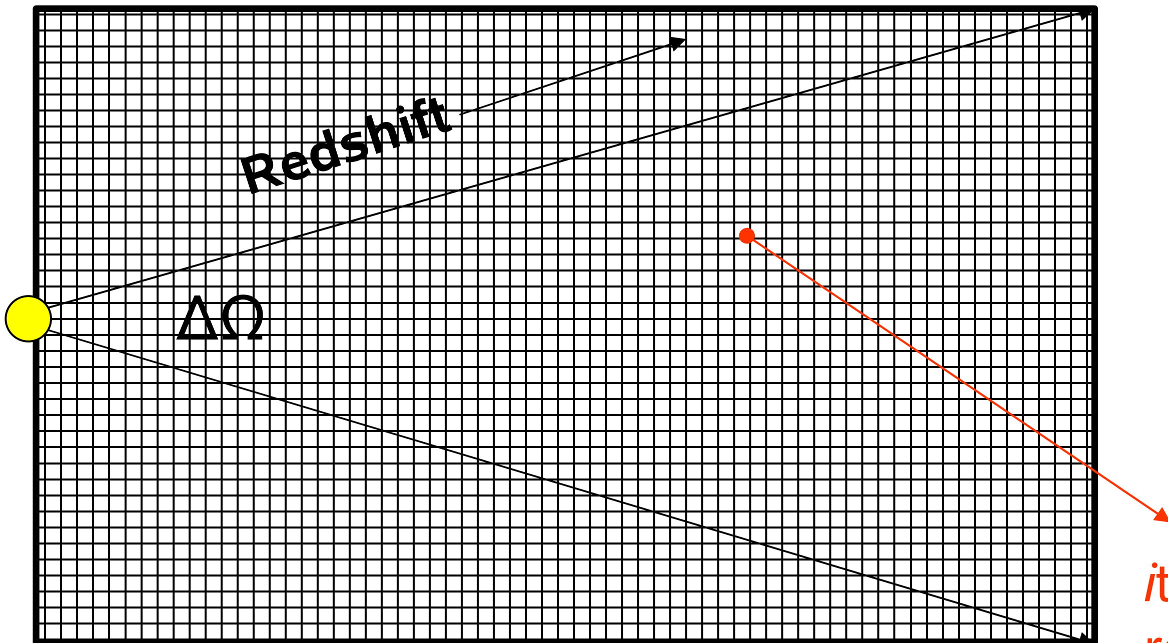
• Compute the mean number of galaxies expected without clustering:

$n_{0,i} = \Phi(L > L_i, z_i)\ \Delta V$

• Modify this by biasing the log-normal density field:

$$n_{CLUS,\,i} = n_{0,i}\ A \exp[\, b(z_i)\ G(z_i)\ (\delta\rho/\rho)]$$

The Theoretical Framework



In the *i*th cell, for each source population:

- Poisson sample the LF at $L > L_i$
- In the limit $(\delta\rho/\rho) \rightarrow 0$, $(\delta n/n) \rightarrow b(z)G(z)(\delta\rho/\rho)$ (i.e. a linear bias model)

Wilman et al. (2008)

Continuum source populations

- Radio-quiet AGN: Hard X-ray AGN LF (Ueda et al. 2003) + X-ray:radio relation (Brinkmann et al. 2000)
- FRI radio sources: Willott et al. (2001) 151 MHz LF
- FRII radio sources: Willott et al. (2001) 151 MHz LF
- Normal star-forming galaxies:

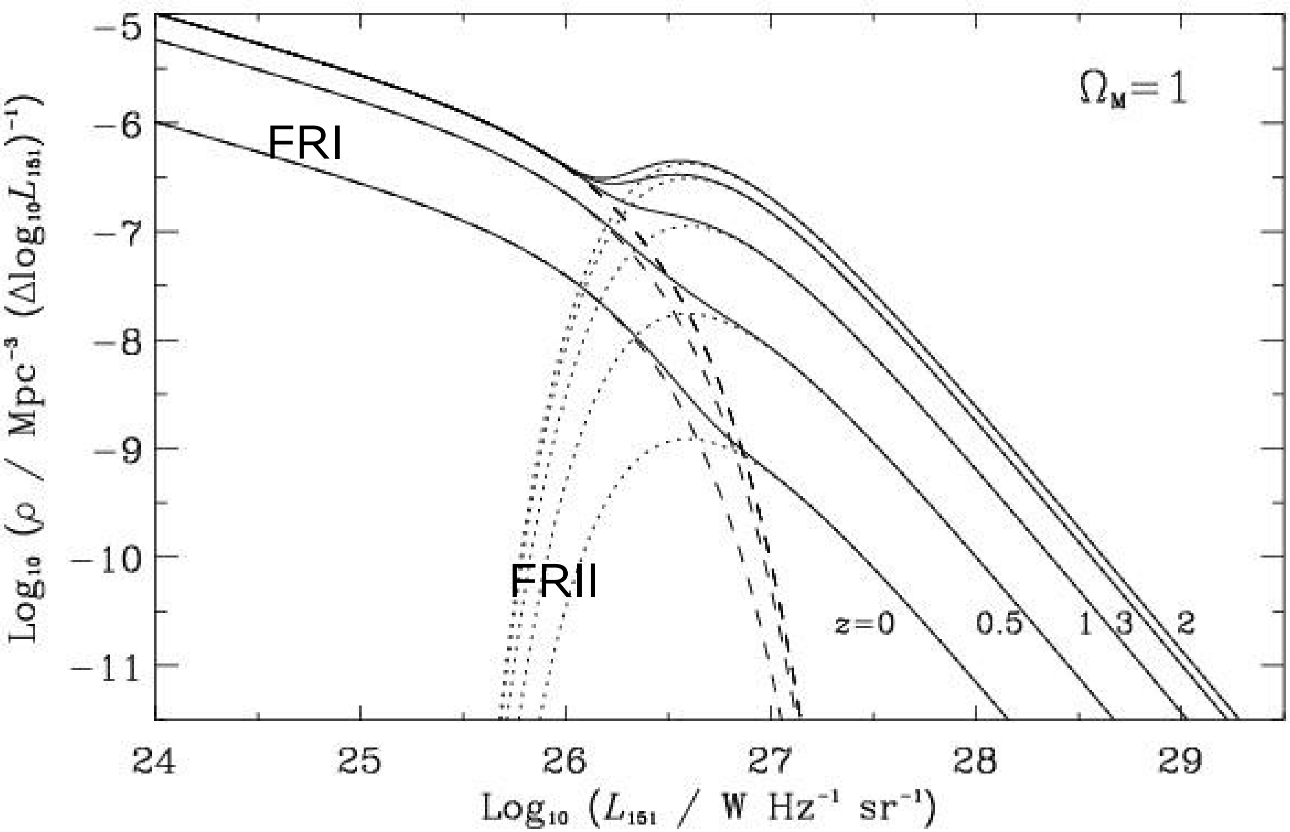
  Yun et al. (2001) 1.4 GHz LF (low-L component) + PLE
- Starburst galaxies:

  Yun et al. (2001) 1.4 GHz LF (high-L component) + PLE

Wilman et al. (2008)

FRI and FRII luminosity functions

Willott et al. (2001) 'Model C' @ 151 MHz



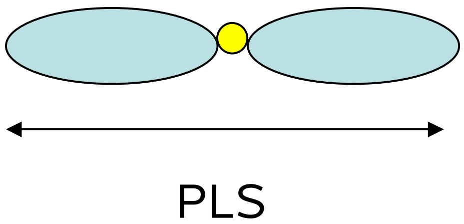
Wilman et al. (2008)

FRI/FRIIs: unification, beaming, structures and spectra

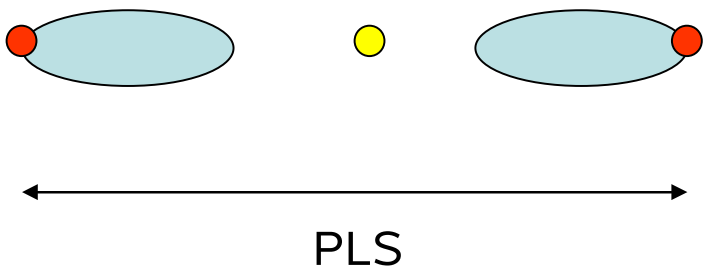
- Assign a true linear size: D_{true} in range $[0, D_0(1+z)^{-1.4}]$, $D_0 = 1$Mpc
- Jet axis-LOS drawn from uniform distribution in cos θ
- Assign intrinsic core:extended luminosity @ 1.4 GHz
  $R_{CL} = 10^x$, $x \sim N(x_{med}, 0.5)$, $x_{med} = -2.6$ (FRI), -2.8 (FRII)
- Apply beaming model to generate $R_{OBS} = R_{CL}\, B(\theta,\gamma)$
  $\gamma=6$ (FRI), 8 (FRII)
- Steep spectrum extended emission, flat-spm core with curvature
- GPS cut-off ($F \sim \nu^{2.5}$) below ν_p: $\log \nu_p = -0.2 - 0.65 \log(D_{true}/\mathrm{kpc})$
  (O'Dea et al. 1998)

Wilman et al. (2008)

FRI/II: unification, beaming, structures & spectra



FRI



FRII:

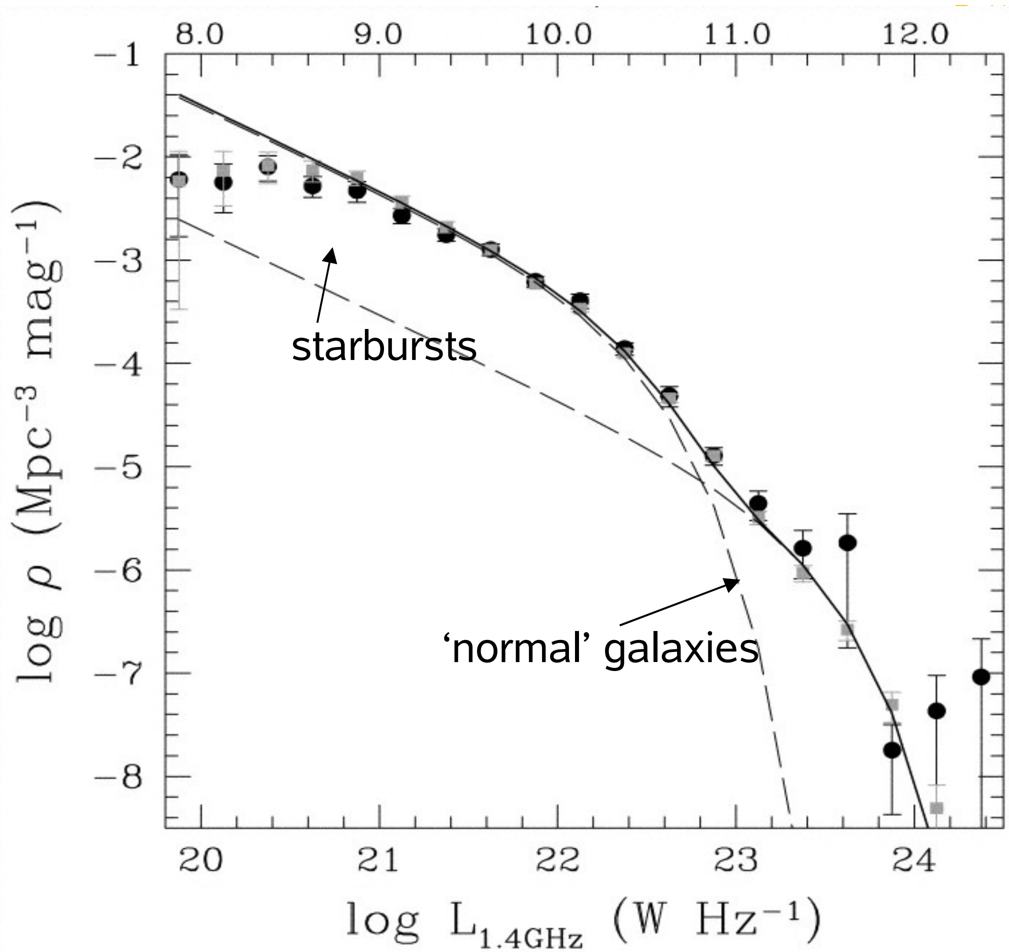
Hotspot:extended flux ratio

$$f_{HS} = 0.4[\log L_{151} - 25.5] \pm 0.15$$

Wilman et al. (2008)

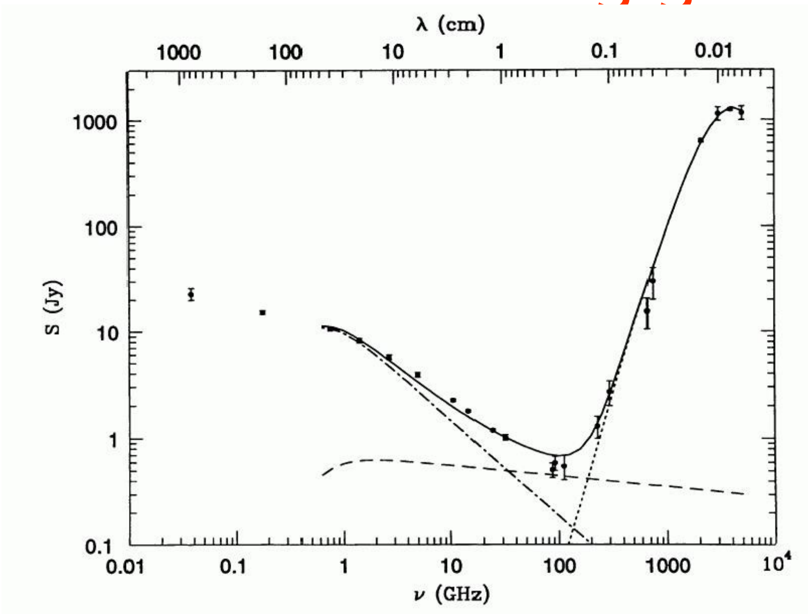
Two populations of star-forming galaxies

- Radio LF of IRAS-selected galaxies from Yun,Reddy & Condon (2001)
- Double Schechter-fn fit representing normal galaxies and starbursts
- We assume LF flattens below $L_{1.4\,GHz} = 10^{20.7}$ W/Hz and integrate down to 10^{18} W/Hz (SFR ~ 10^{-3} M /yr)



Wilman et al. (2008)

Star-forming galaxies: SEDs



M82 radio-far IR SED (Condon 1992)

3 model SED components:

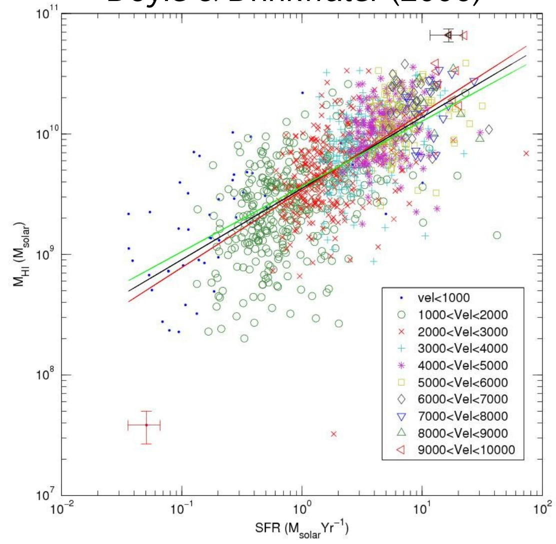
- Non-thermal: $F_\nu \sim \nu^{-0.75}$
- Thermal: $F_\nu \sim \nu^{-0.1}$
- Far-IR thermal dust (45K)

Starbursts: free-free absorption below 1 GHz

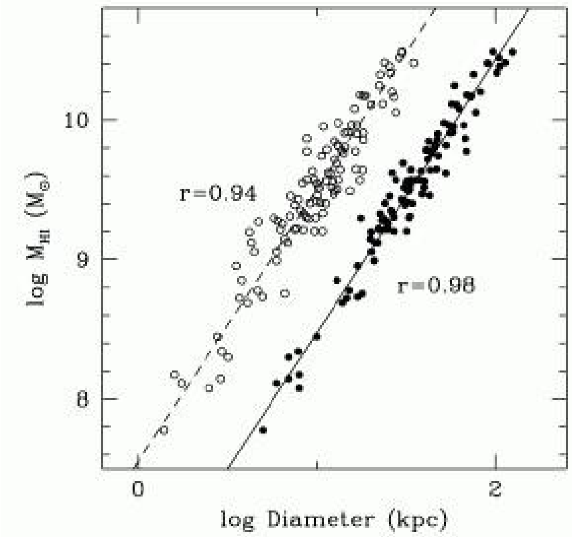
Wilman et al. (2008)

Star-forming galaxies: HI and disk sizes

Doyle & Drinkwater (2006)



Broeils & Rhee (1997)



- Normal galaxies: $\log M_{HI} = (1.96 \pm 0.04) \log D_{HI} + 6.52 \pm 0.06$

Broeils & Rhee (1997)

- Starburst galaxies: $\log D_{cont} = (1+z)^{2.5}$ kpc to z=1.5

  10 kpc at z>1.5

Wilman et al. (2008)

Large-scale structure and biasing

Each population assigned a halo mass which reflects large-scale clustering which is then used to compute b(z)

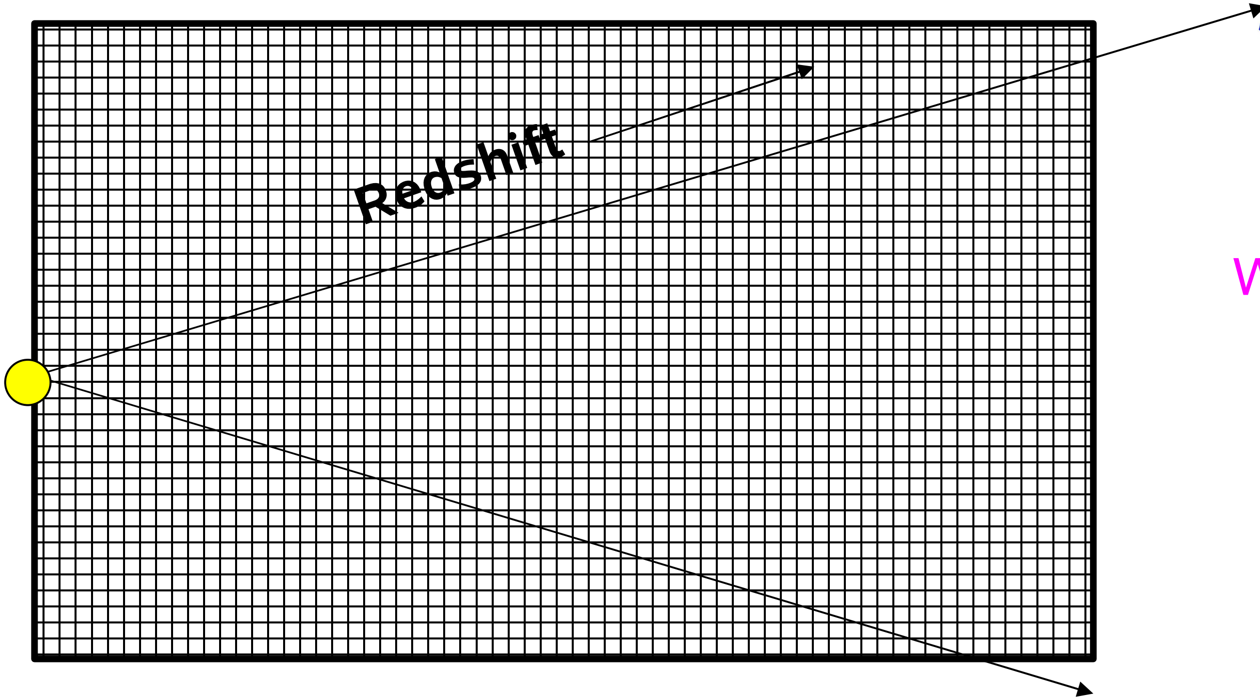
N.B. We are not directly populating galaxy-sized haloes

- Radio-quiet quasars: M_{halo} = 3E12/h M
- FRI radio sources: M_{halo} = 1E13/h M
- FRII radio sources: M_{halo} = 1E14/h M

- Normal star-forming galaxies: M_{halo} = 1E11/h M
- Starburst galaxies: M_{halo} = 5E13/h M

Wilman et al. (2008)

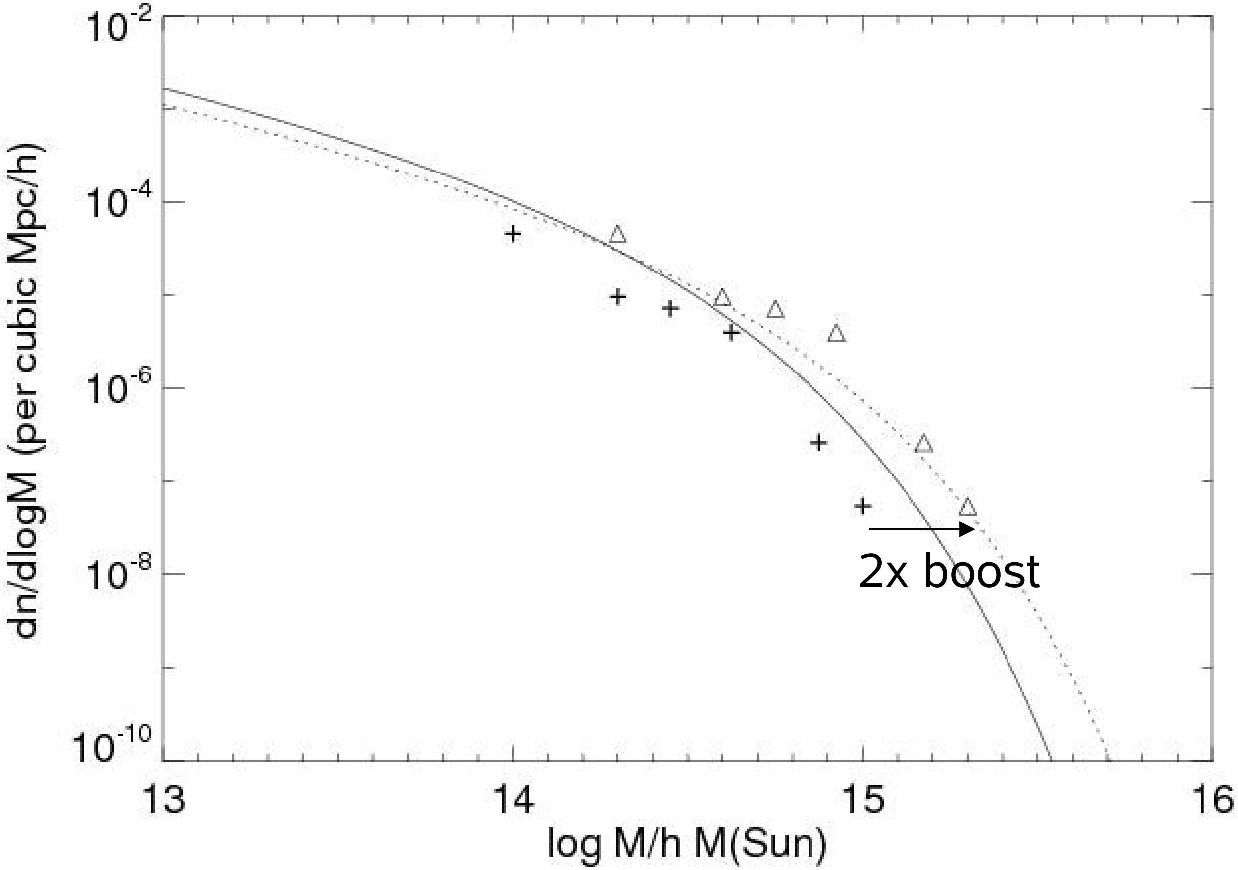
Clusters of galaxies



Wilman et al. (2008)

- Each 5 h^{-1} Mpc cell has mass 10^{13} h^{-1} M → resolution to identify cluster-mass haloes
- Smooth density field on a range of mass scales, 10^{14-16} h^{-1} M , and search for islands of overdense cells with $(\delta\rho/\rho) > 1.66/G(z)$
  - Discreteness of grid and lack of filter-edge interpolation → 'quantised' cluster masses

Cluster mass function

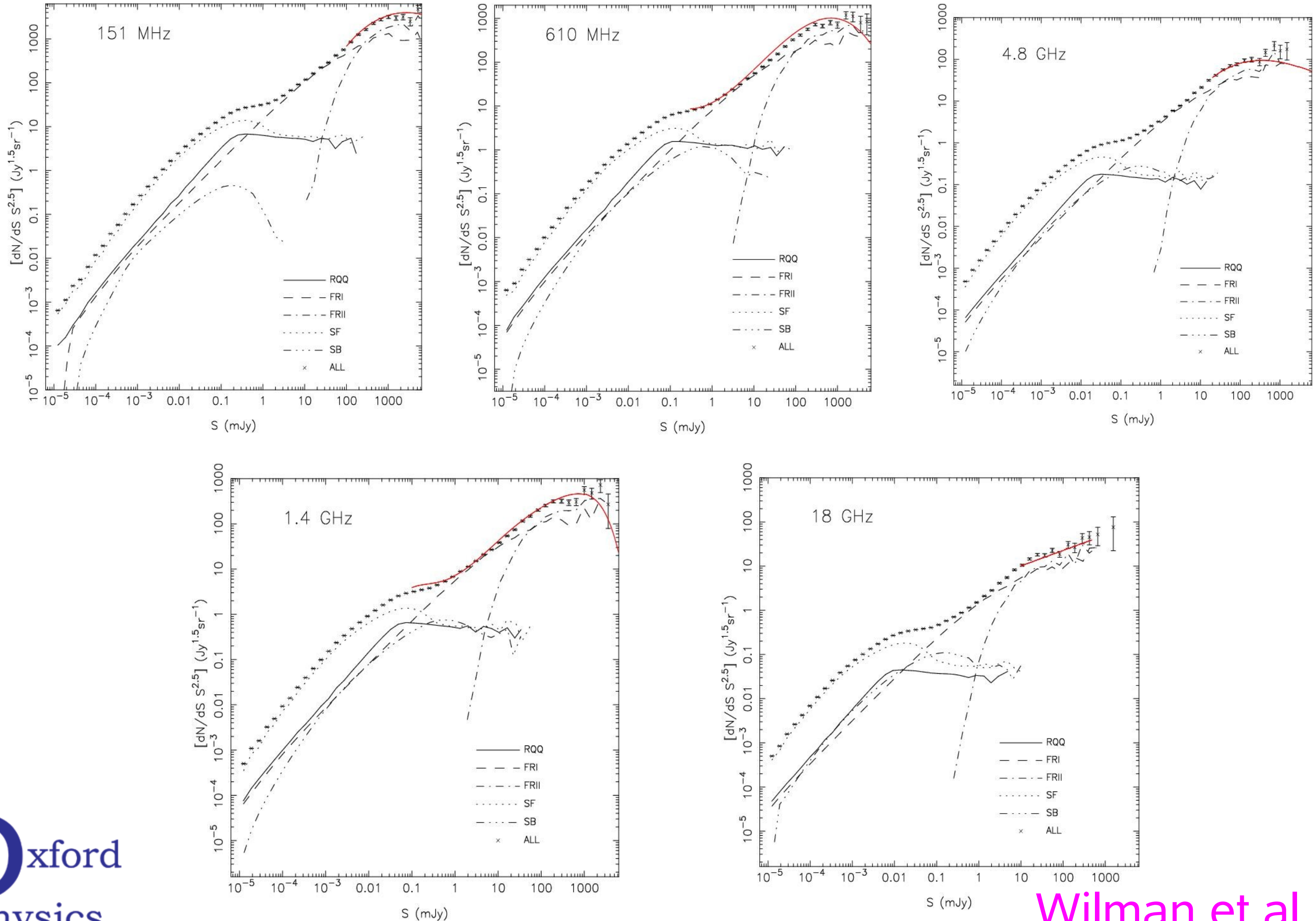


Press-Schechter analytical mass function

Sheth-Tormen

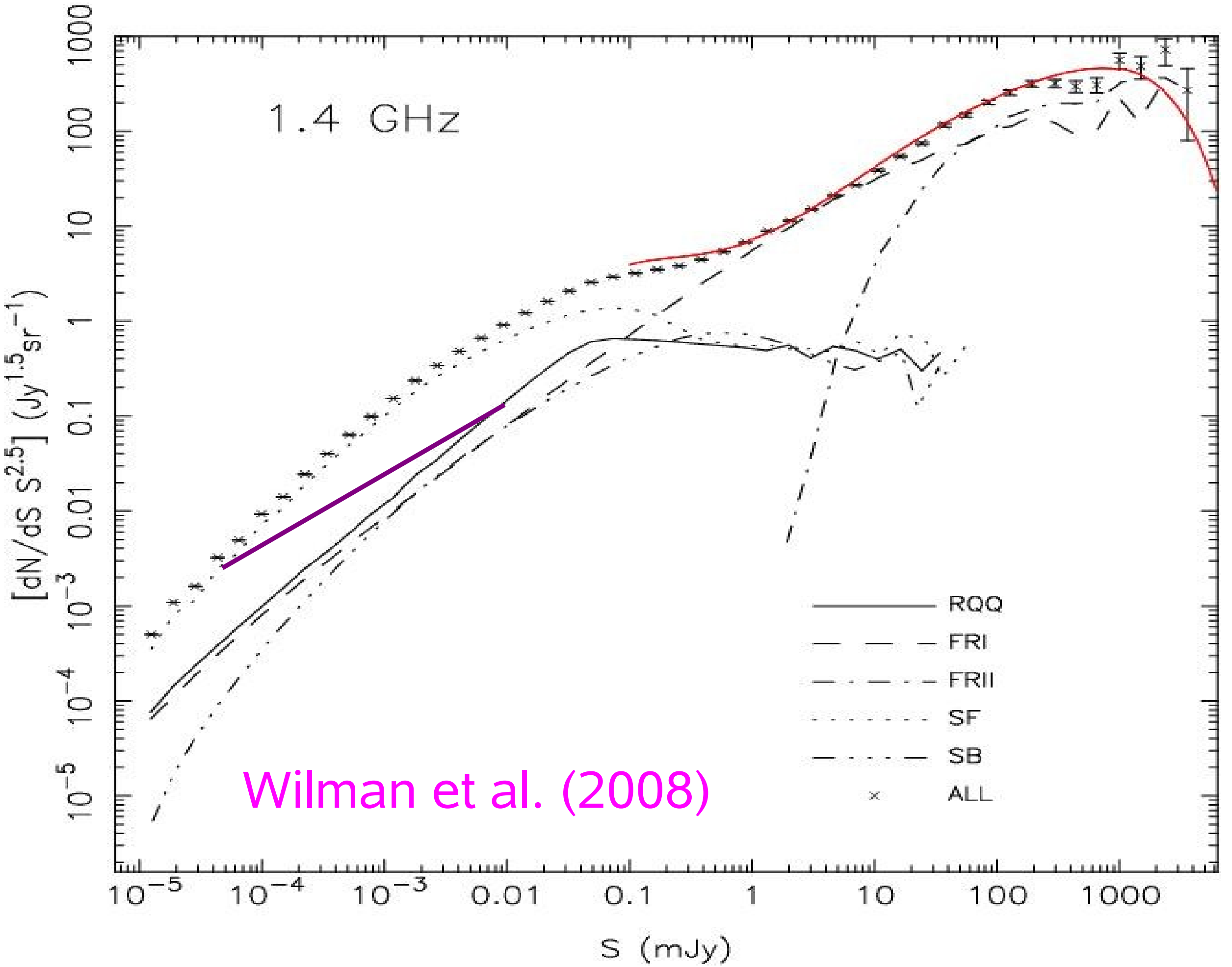
Wilman et al. (2008)

Simulation output: source counts



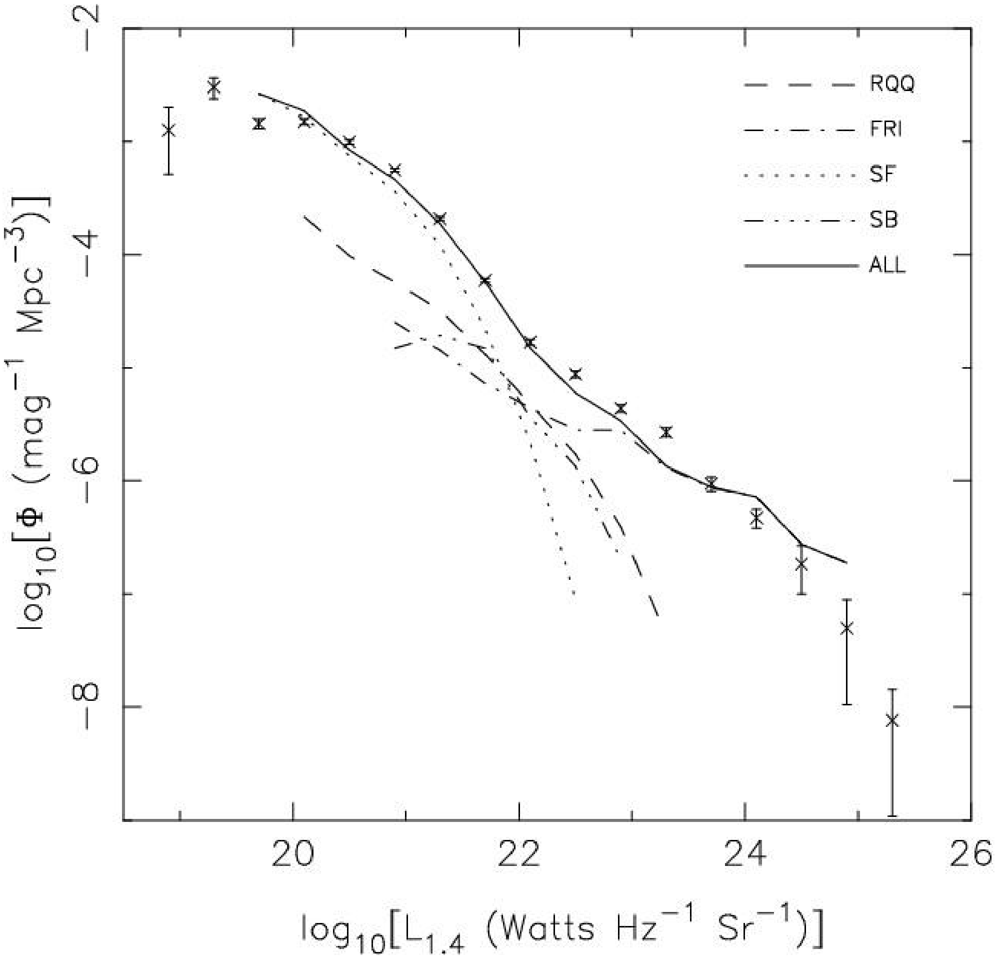
Wilman et al. (2008)

Simulation output: source counts



Hopkins et al. (2000)/Windhorst (2003) model which includes an additional population of ‘normal galaxies’ derived from an optical LF

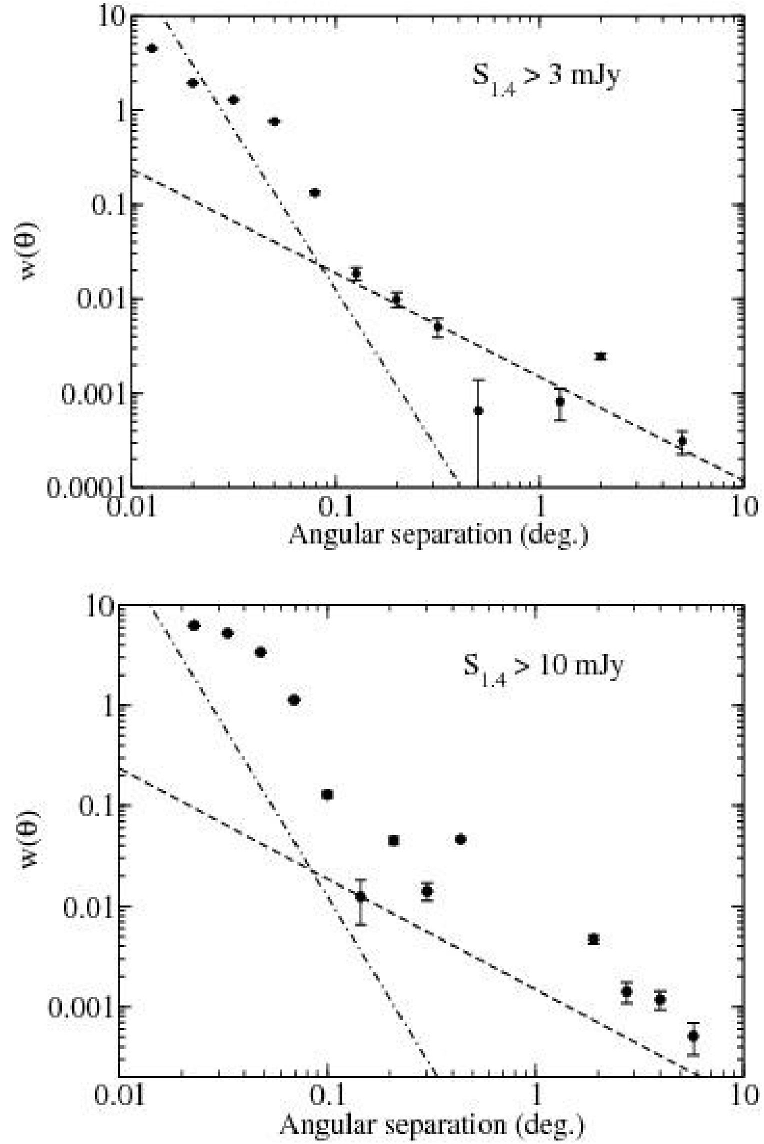
Simulation output: 1.4 GHz local (z<0.3) luminosity function



Data points from Mauch & Sadler (2007)

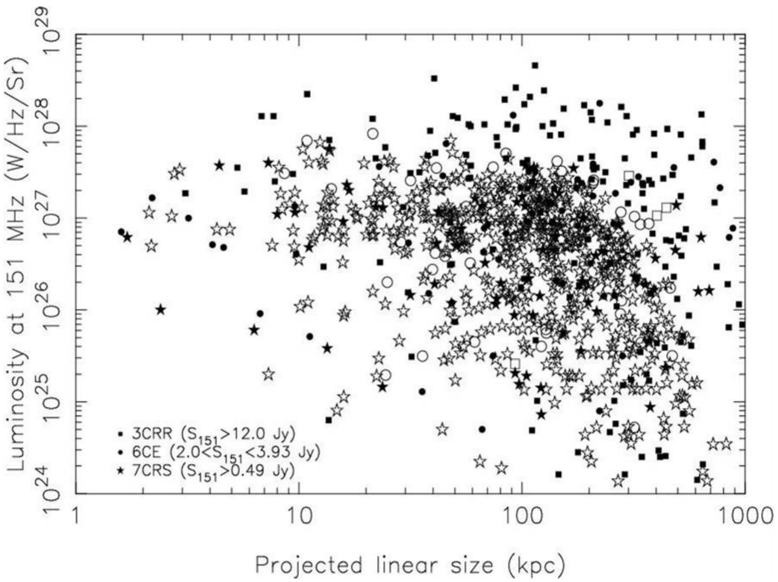
Wilman et al. (2008)

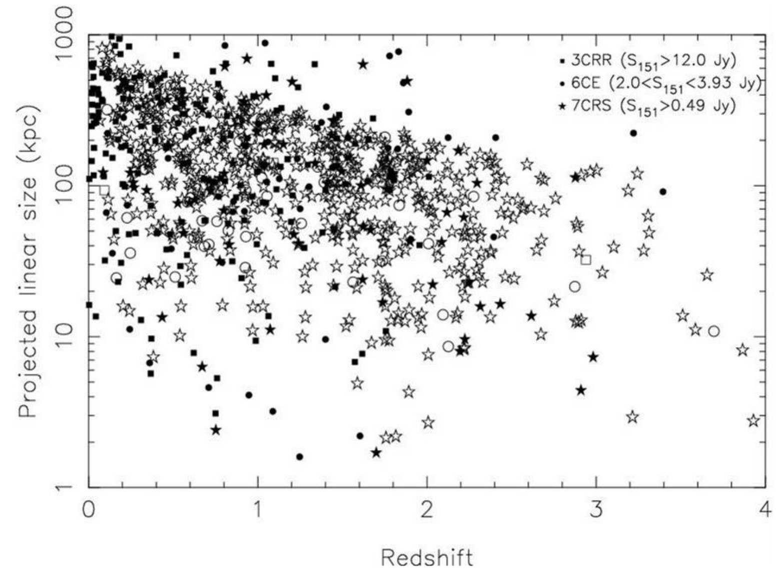
Simulation output: angular clustering



Wilman et al. (2008)

Simulation output: *P-D* & *D-z* diagrams





Relative sky coverage of the surveys:

Simulation: 0.12 sr 3CRR: 4.24 sr 7CRS: 0.022 sr

Wilman et al. (2008)

Catalogue format

- A common standard for S^3 output
- Galaxy index to identify sub-component sources (cores/lobes/hotspots) to parent galaxy
- Source type (AGN/SF) and sub-component type
- Cluster Index
- Millennium catalogue halo and galaxy indices
- Continuum I[QUV] flux densities @ 151 & 610 MHz, 1.4, 4.86 & 18 GHz
- Morphological info: ellipse PA, major+minor axes

MAIN TABLE

#1	Unique source index
#2	Cluster index (0-no member, cluster index)
#3	Galaxy index
#4	SF type (0-no SF, 1-normal SF, 2-SB)
#5	AGN type (0-no AGN, 1-RQQ, 2-FRI, 3-FRII, 4-GPS)
#6	Structure type (1-core, 2-lobe, 3-hotspot, 4-SFdisk)
#7	RA
#8	DEC
#9	Comoving distance (Mpc)
#10	Redshift
#11	Position angle (rad) for elliptical substructures
#12	Major axis (arcsec)
#13	Minor axis (arcsec)
#14-17	log I, log Q, log U, log V @ 151 MHz
#18-21	log I, log Q, log U, log V @ 610 MHz
#22-25	log I, log Q, log U, log V @ 1.4 GHz
#26-29	log I, log Q, log U, log V @ 4.86 GHz
#30-33	log I, log Q, log U, log V @ 18 GHz
#34	log (M_{HI}) (star-forming galaxies only)
#35	cos (viewing angle) (wrt jet axis, FRI and FRII only)

CLUSTER TABLE

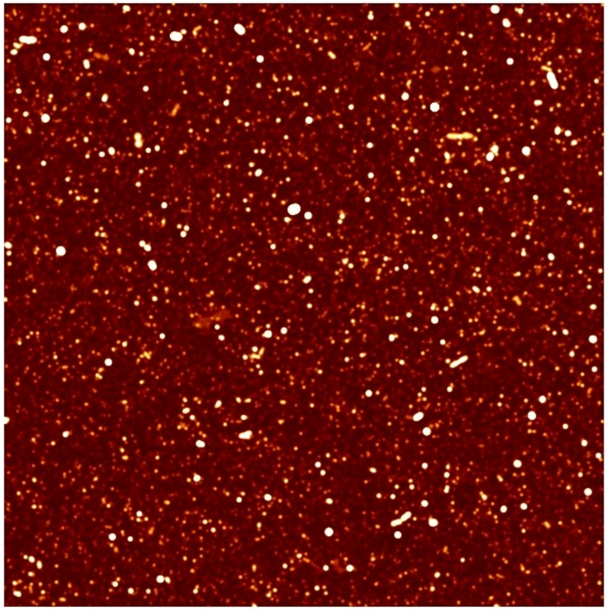
#1	Cluster index
#2	RA centre
#3	DEC centre
#4	Redshift
#5	mass
#6	Virial radius
#7	Cluster virial velocity dispersion

Accessing the simulations

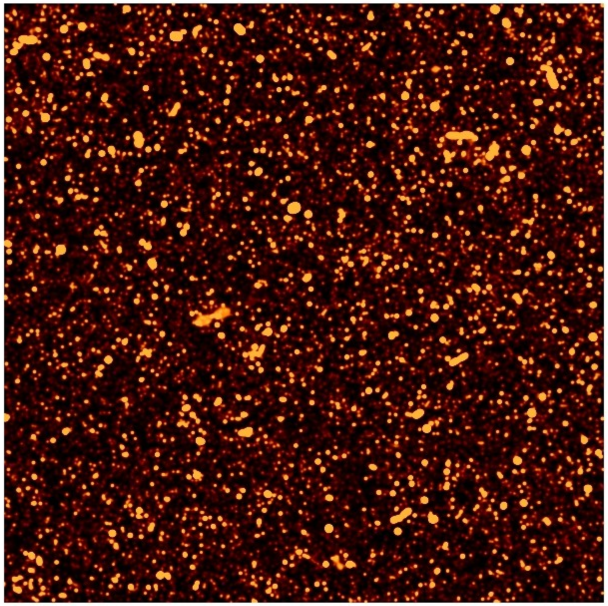
- Central sq deg: catalogues and maps available on the SKADS wiki
- Full 20x20 deg^2: interactive database query tool and mapmaker at

http://s-cubed.physics.ox.ac.uk

Wilman et al. (2008)



1.4 GHz



18 GHz

Limitations of the simulations

- The use of extrapolated luminosity functions
- The lack of star-forming/AGN hybrid galaxies + double counting
- The lack of redshift space distortions
- The lack of small-scale non-linear clustering
- The simplified treatment of galaxy clusters

→ Post-processing by the user is essential!

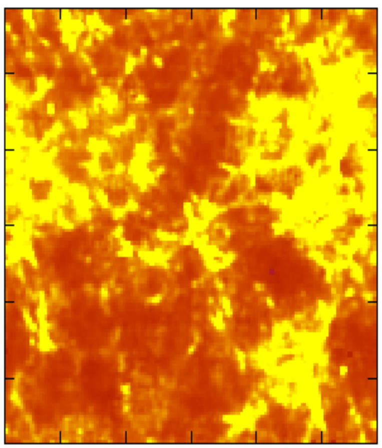
The next step?

- A full hemisphere simulation (20,000 deg^2)
- Star-forming galaxies only
- Out to redshift z=2-3
- Linear theory z-space distortions
- Extension (ii) FRI/II structures: simple prescriptions

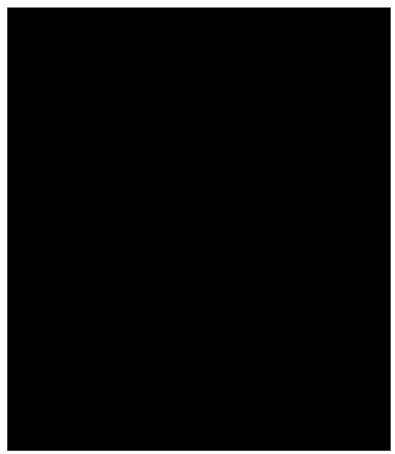
Catalogue format

- A common standard for SKADS extragalactic simulation catalogues
- Galaxy index to identify sub-component sources (cores/lobes/hotspots) to parent galaxy
- Source type (AGN/SF) and sub-component type
- Cluster Index
- Millennium catalogue halo and galaxy indices
- Continuum I[QUV] flux densities @0.15, 0.61, 1.4, 5 and 20 GHz
- HI fluxes and profile descriptions
- Morphological info: ellipse PA, major+minor axes

how does it fit together

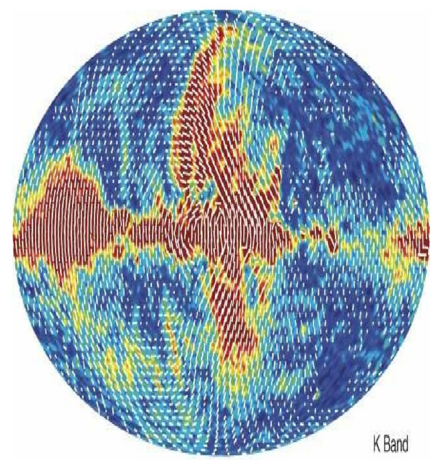


EoR

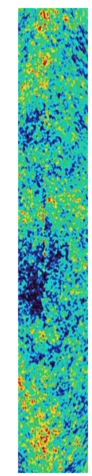


galaxies, clusters

I
Q
U
V

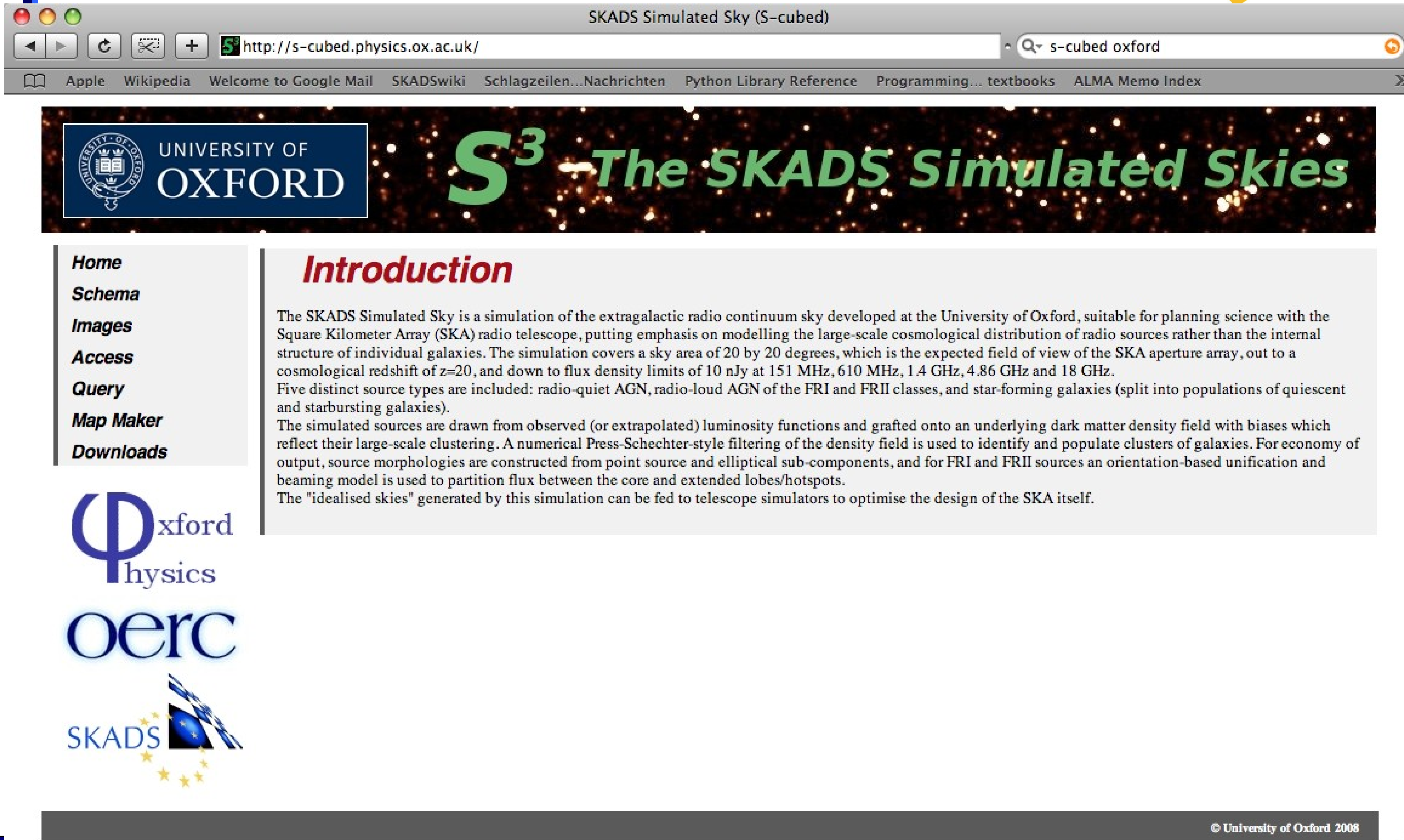


Milky Way



Ionosphere / Troposphere

s-cubed.physics.ox.ac.uk



SKADS Simulated Sky (S-cubed)

http://s-cubed.physics.ox.ac.uk/

S³ The SKADS Simulated Skies

UNIVERSITY OF OXFORD

- Home
- Schema
- Images
- Access
- Query
- Map Maker
- Downloads

Introduction

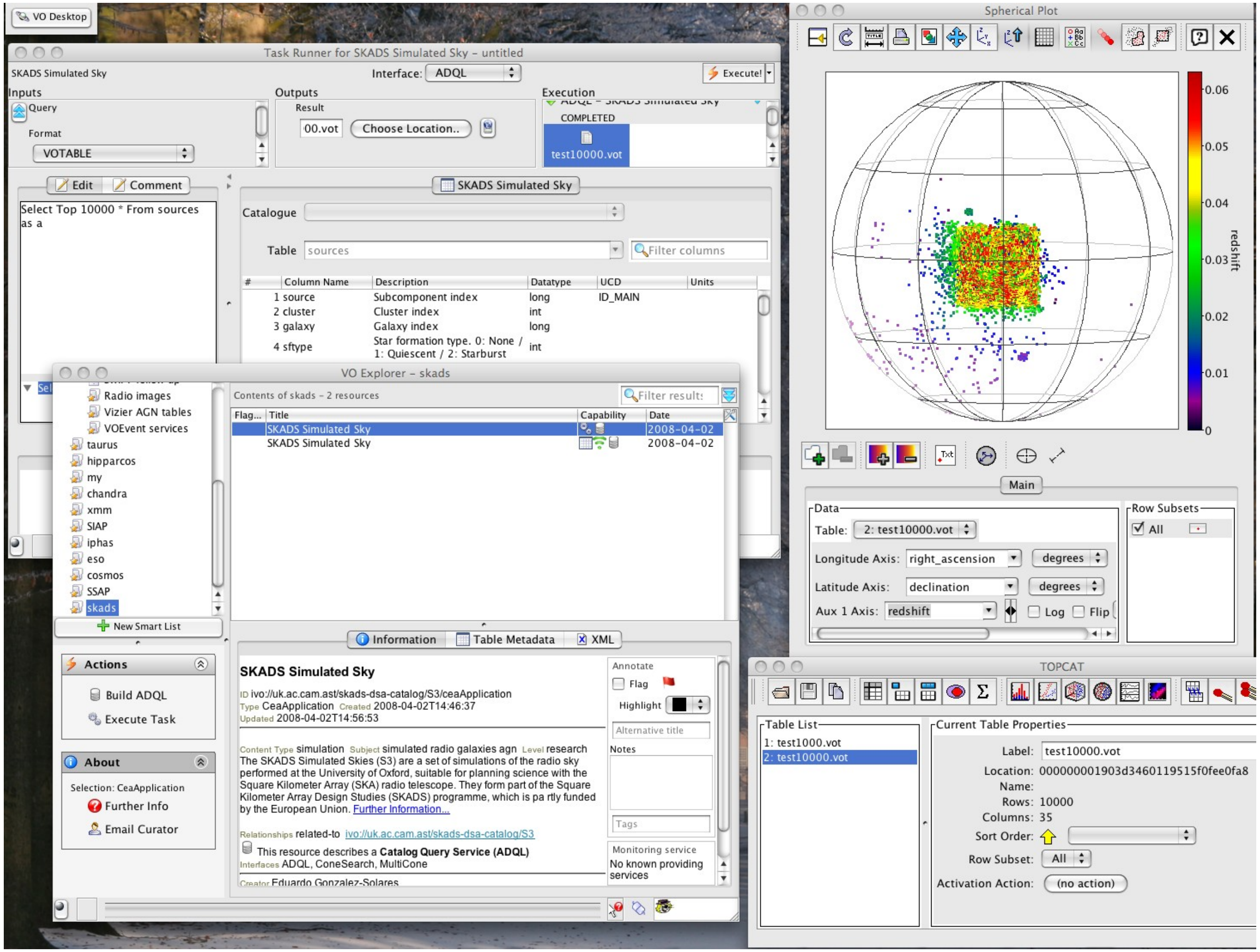
The SKADS Simulated Sky is a simulation of the extragalactic radio continuum sky developed at the University of Oxford, suitable for planning science with the Square Kilometer Array (SKA) radio telescope, putting emphasis on modelling the large-scale cosmological distribution of radio sources rather than the internal structure of individual galaxies. The simulation covers a sky area of 20 by 20 degrees, which is the expected field of view of the SKA aperture array, out to a cosmological redshift of z=20, and down to flux density limits of 10 nJy at 151 MHz, 610 MHz, 1.4 GHz, 4.86 GHz and 18 GHz.

Five distinct source types are included: radio-quiet AGN, radio-loud AGN of the FRI and FRII classes, and star-forming galaxies (split into populations of quiescent and starbursting galaxies).

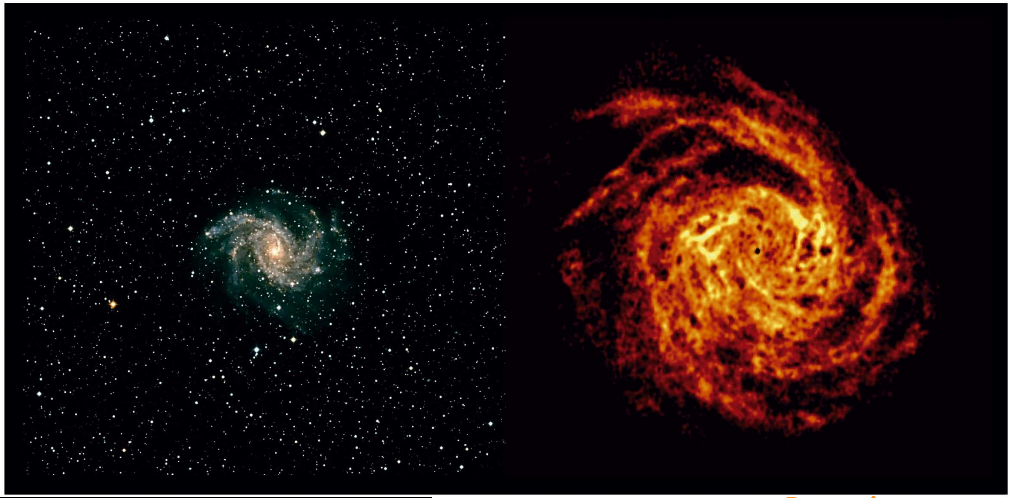
The simulated sources are drawn from observed (or extrapolated) luminosity functions and grafted onto an underlying dark matter density field with biases which reflect their large-scale clustering. A numerical Press-Schechter-style filtering of the density field is used to identify and populate clusters of galaxies. For economy of output, source morphologies are constructed from point source and elliptical sub-components, and for FRI and FRII sources an orientation-based unification and beaming model is used to partition flux between the core and extended lobes/hotspots.

The "idealised skies" generated by this simulation can be fed to telescope simulators to optimise the design of the SKA itself.

...and via the Virtual Observatory!



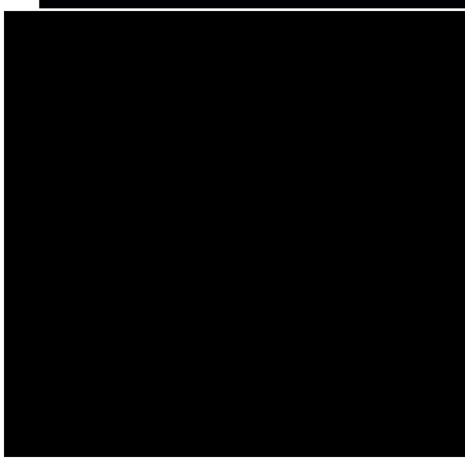
painting on needed



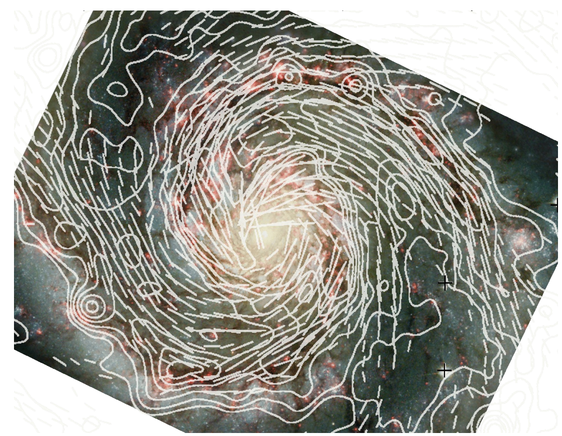
HI

Groningen

B - field configuration



polarization / depolarization (source IQUV & cluster)



Cambridge

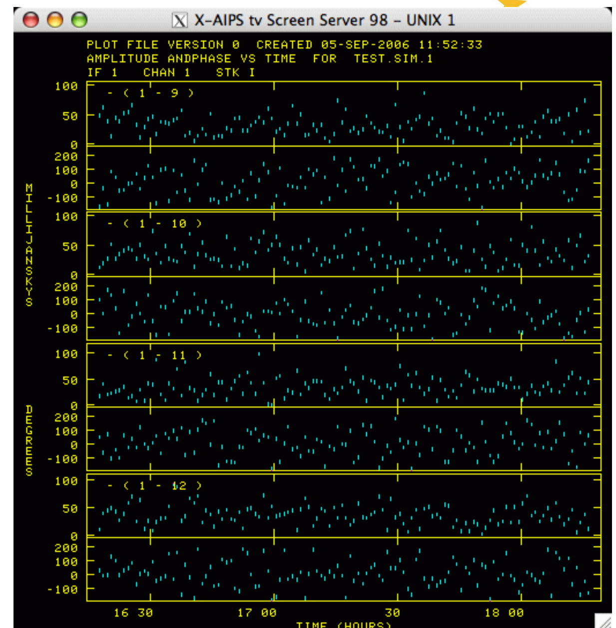
Bonn

Synthesized Catalogues

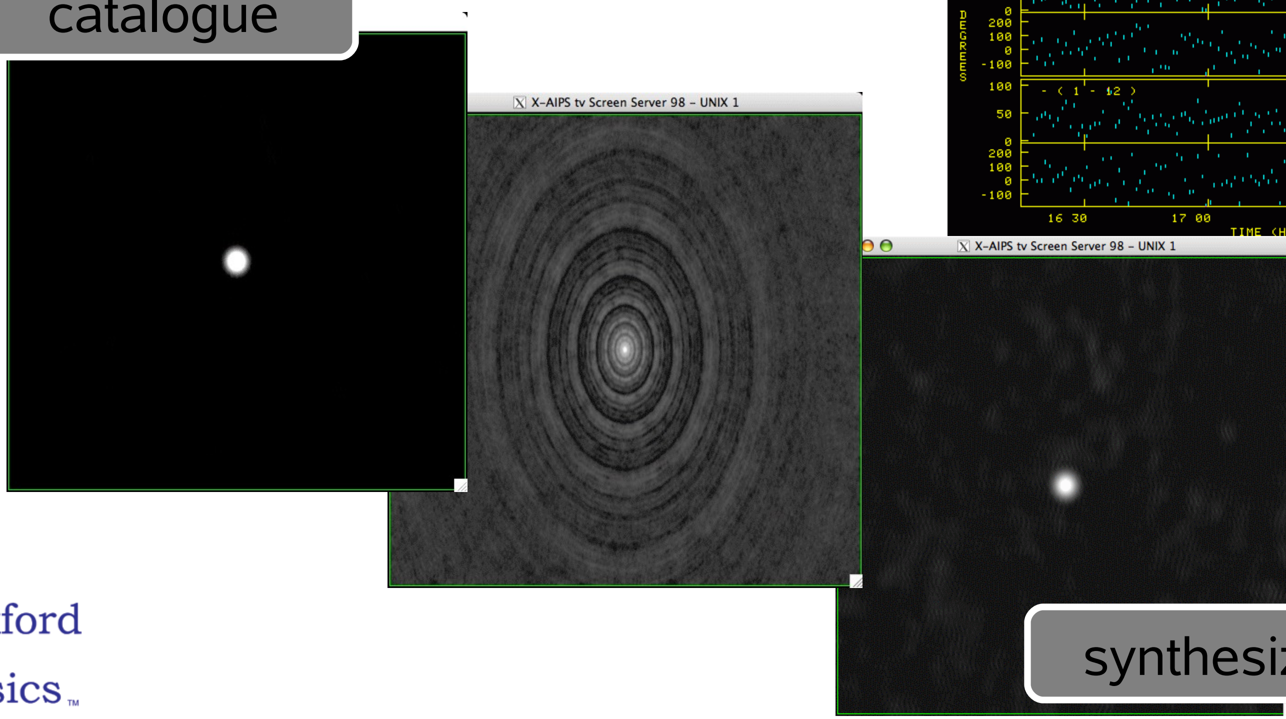
SKADS

Sky Simulation (DS2T1)

Telescope Simulation (DS2T2)



catalogue

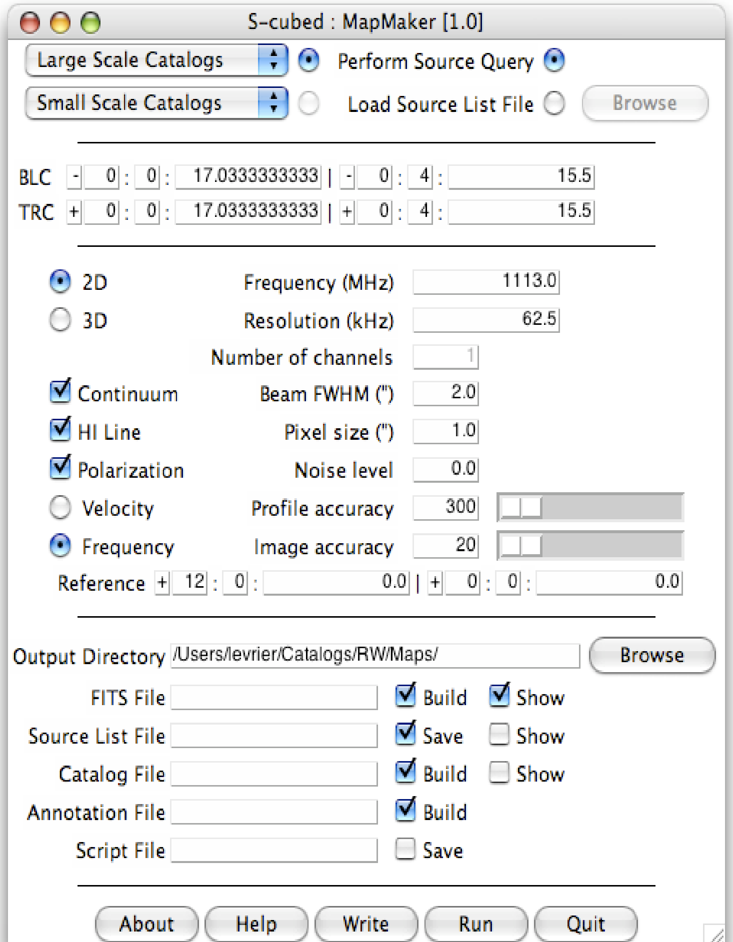


synthesized catalogue

Oxford Physics

Tools: e.g. the MapMaker

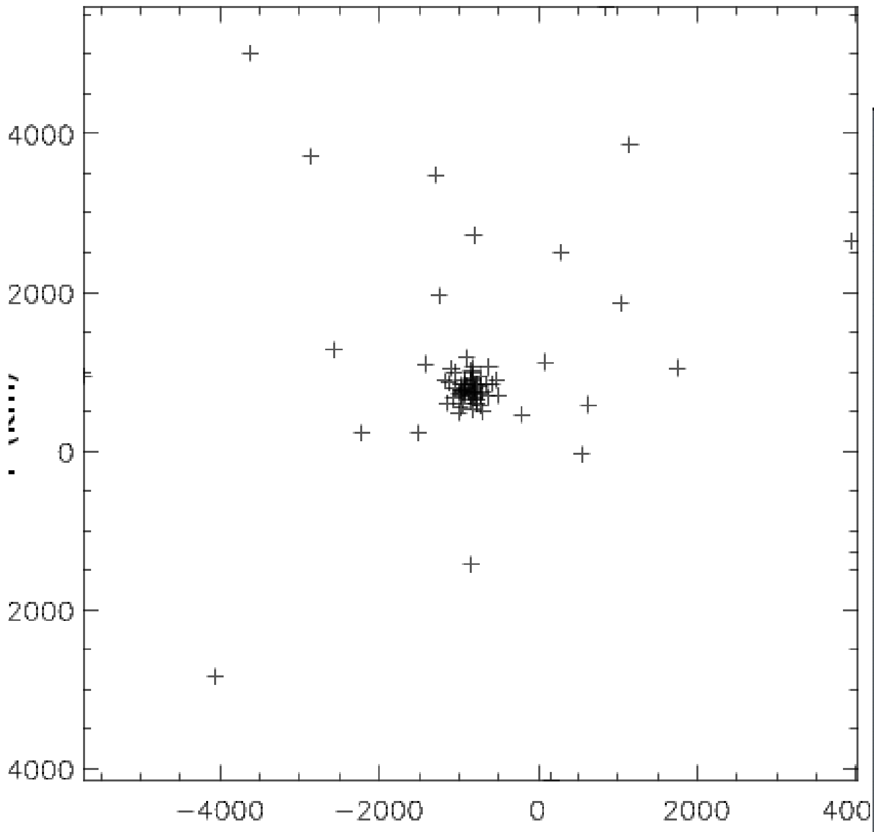
SKADS

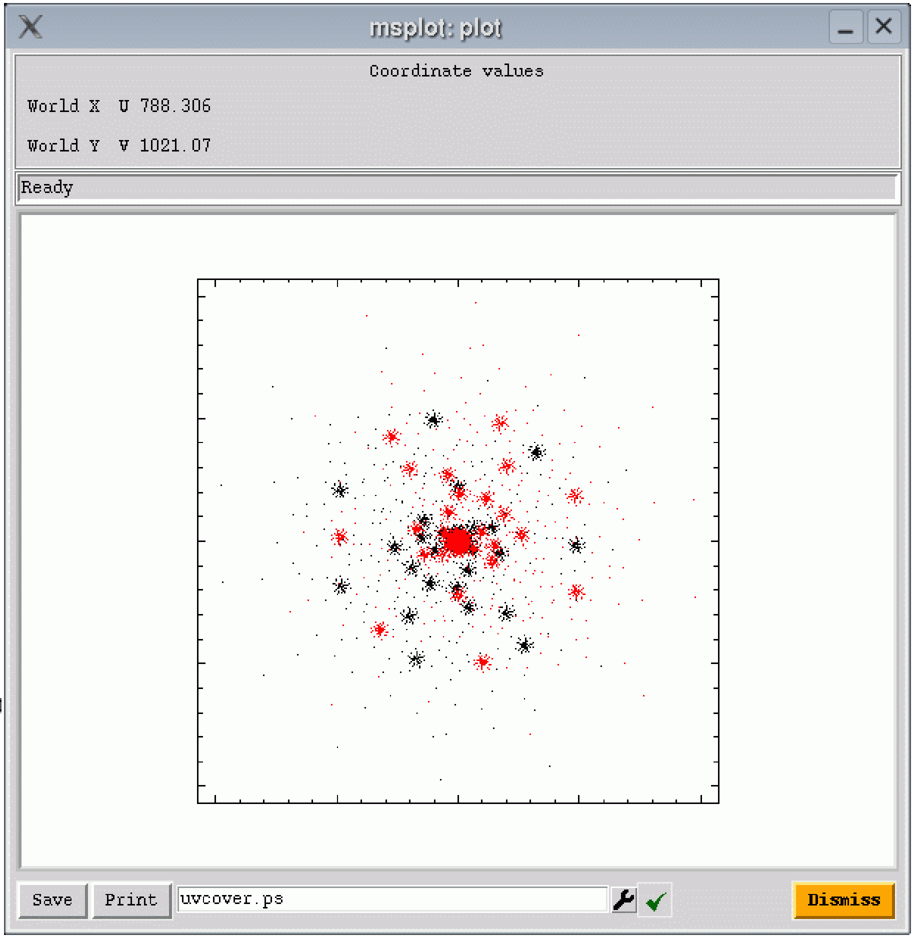


- Standalone application in Python
- SQL/Python interface
- Source list = query + post-processing
- FITS output with annotation file
- 2D images or 3D cubes
- Backend routines to be parallelized and plugged into VO

Oxford Physics

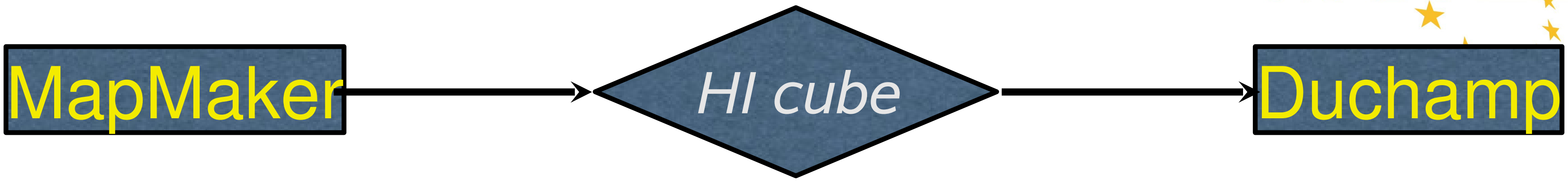
Configuration Library





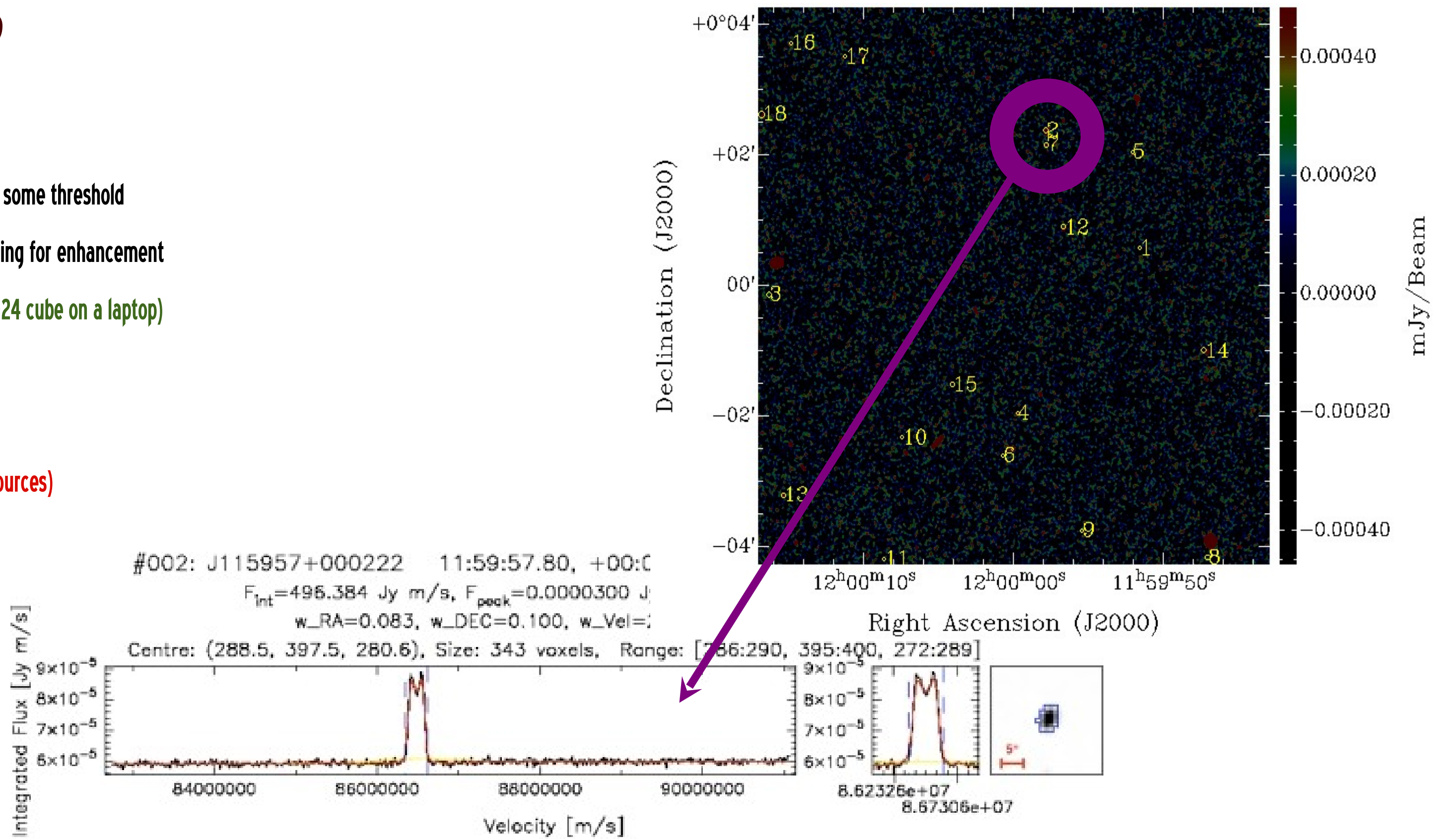
Source HI Extraction

R. Auld (Cardiff)



Duchamp

- ATNF (M. Whiting)
- Locates sets of contiguous voxels above some threshold
- Uses spectral, spatial or wavelet smoothing for enhancement
- Very quick (30 minutes for 512 x 512 x 1024 cube on a laptop)
- VO-compliant
- Memory and process intensive
- Completeness issues (misses high S/N sources)



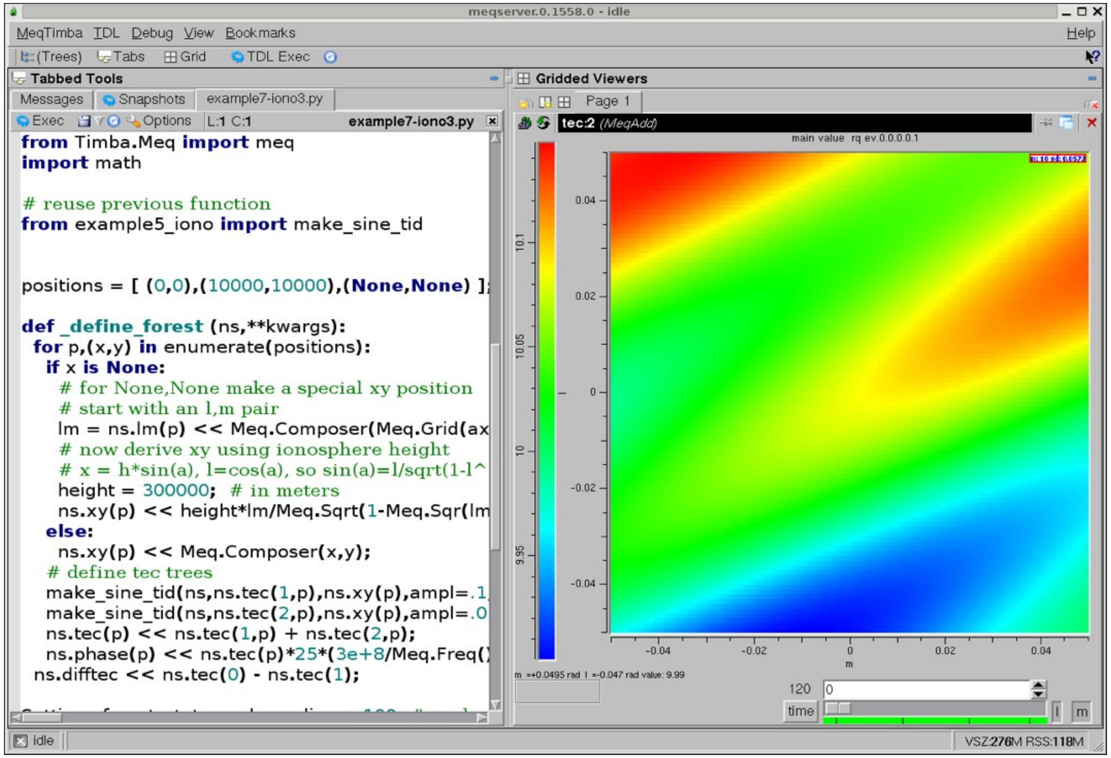
The MeqTrees Software

Developed @ ASTRON [Smirnov, Noordam] with outside contributions [Willis]

Open source code Linux / UNIX (virtual machine available)

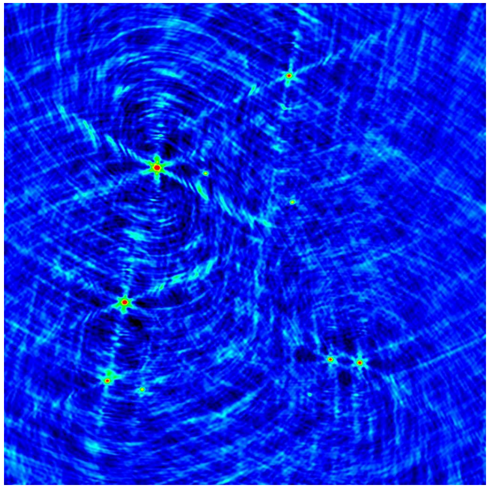
Code repository for beam shapes, ionospheric models, ...

Trees are declared by python scripts and executed by a C++ kernel

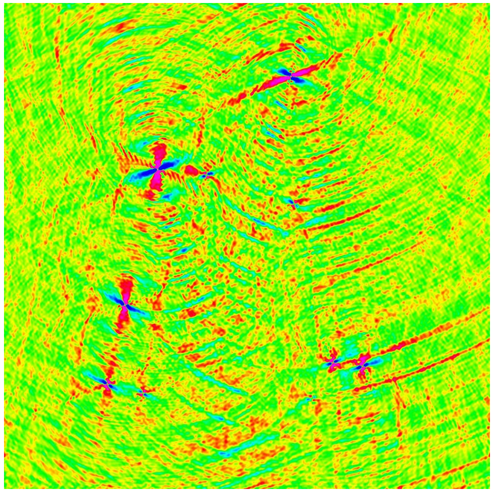


MeqTrees Features

- Powerful calibration : almost any parameter can be solved for
- Fast and smart evaluation on time/frequency grids
- Per-station effects and polarization effortlessly included
- Visual assessment of different models via difference trees



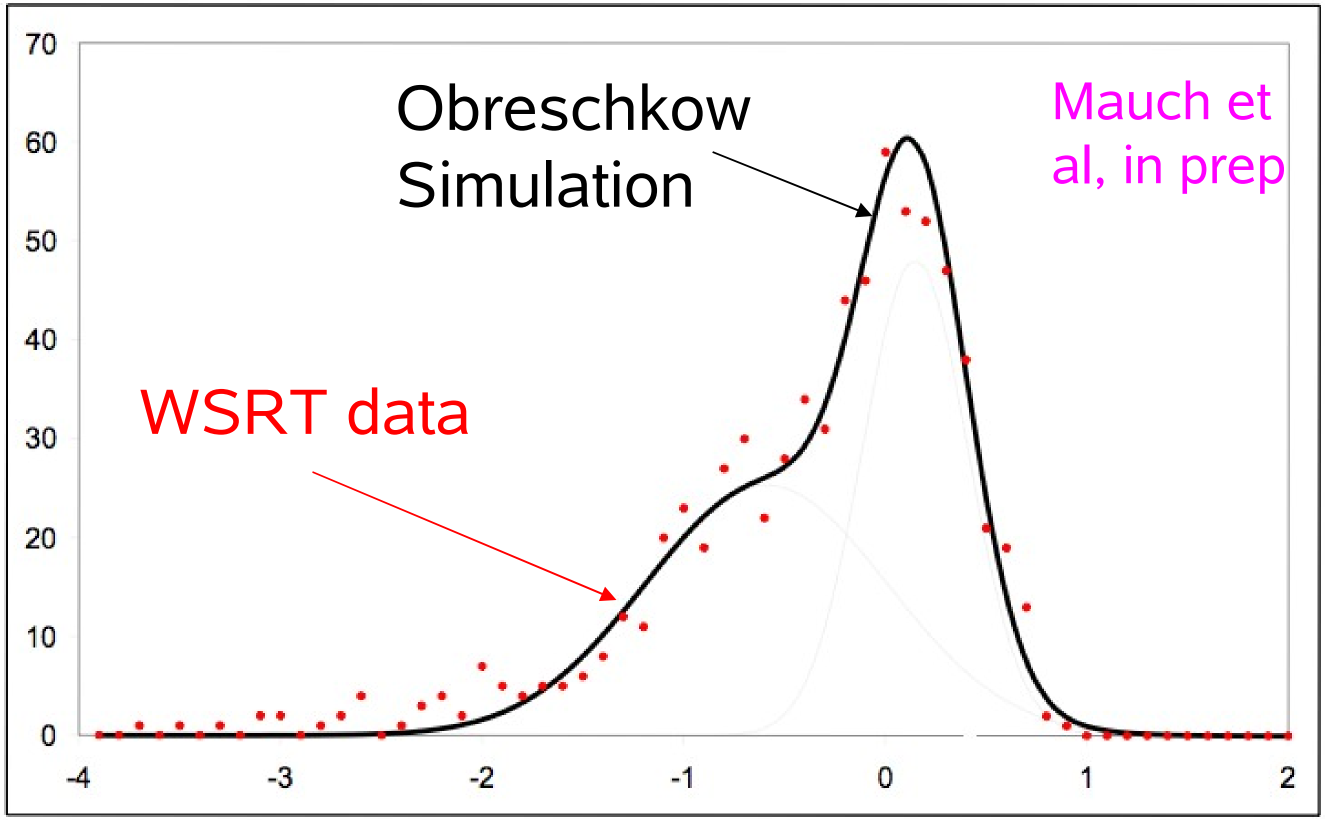
symmetrical vs asymmetrical beam
error level: 2%



average vs per-station beam
error level: 0.02%

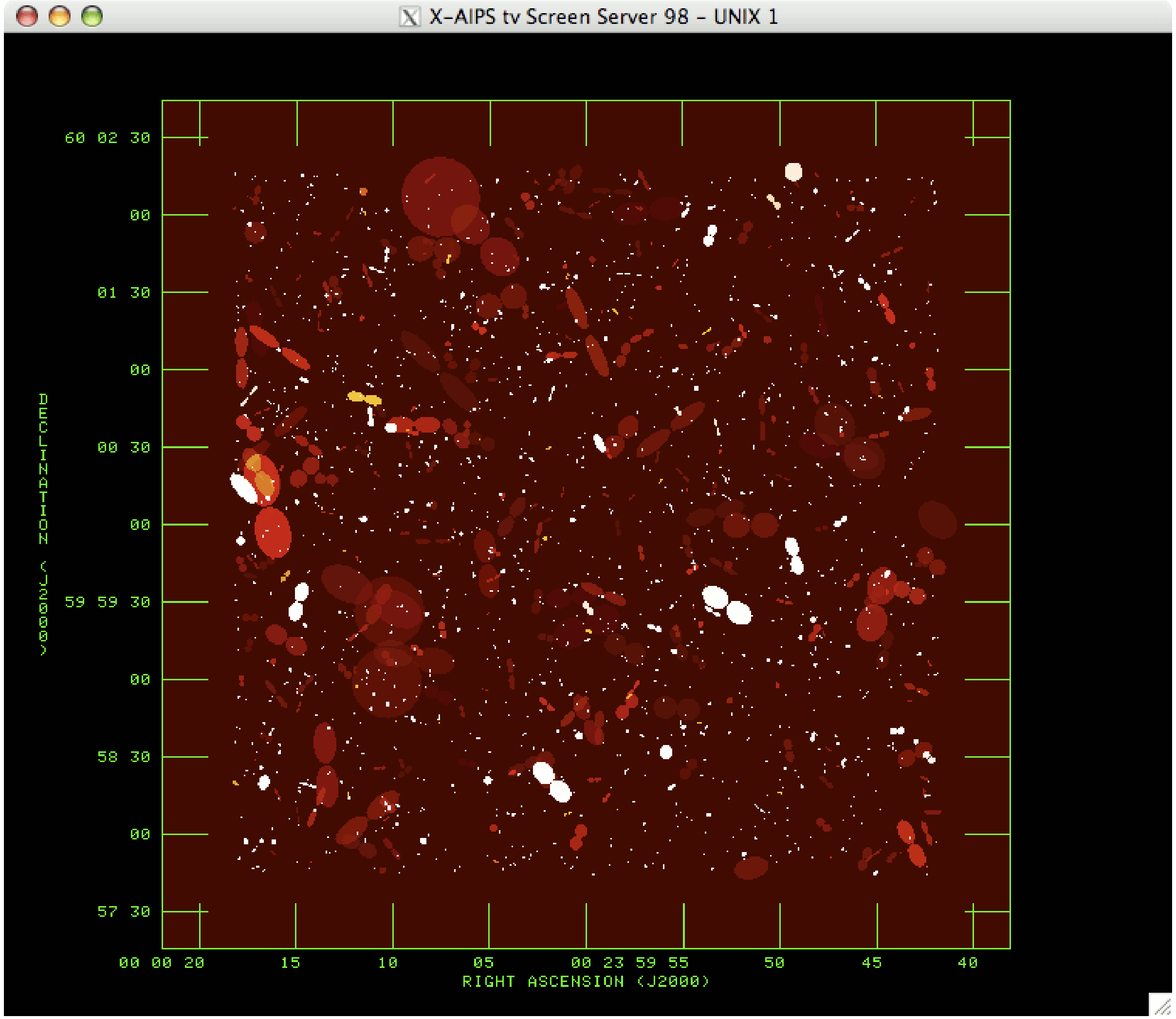
- No parallelization yet : limited number of sources and antennae
- Uses Measurement Sets but does not generate them
- Input skies are source lists, not images

Confontation with Data



Sky simulations will increasingly be checked against real data

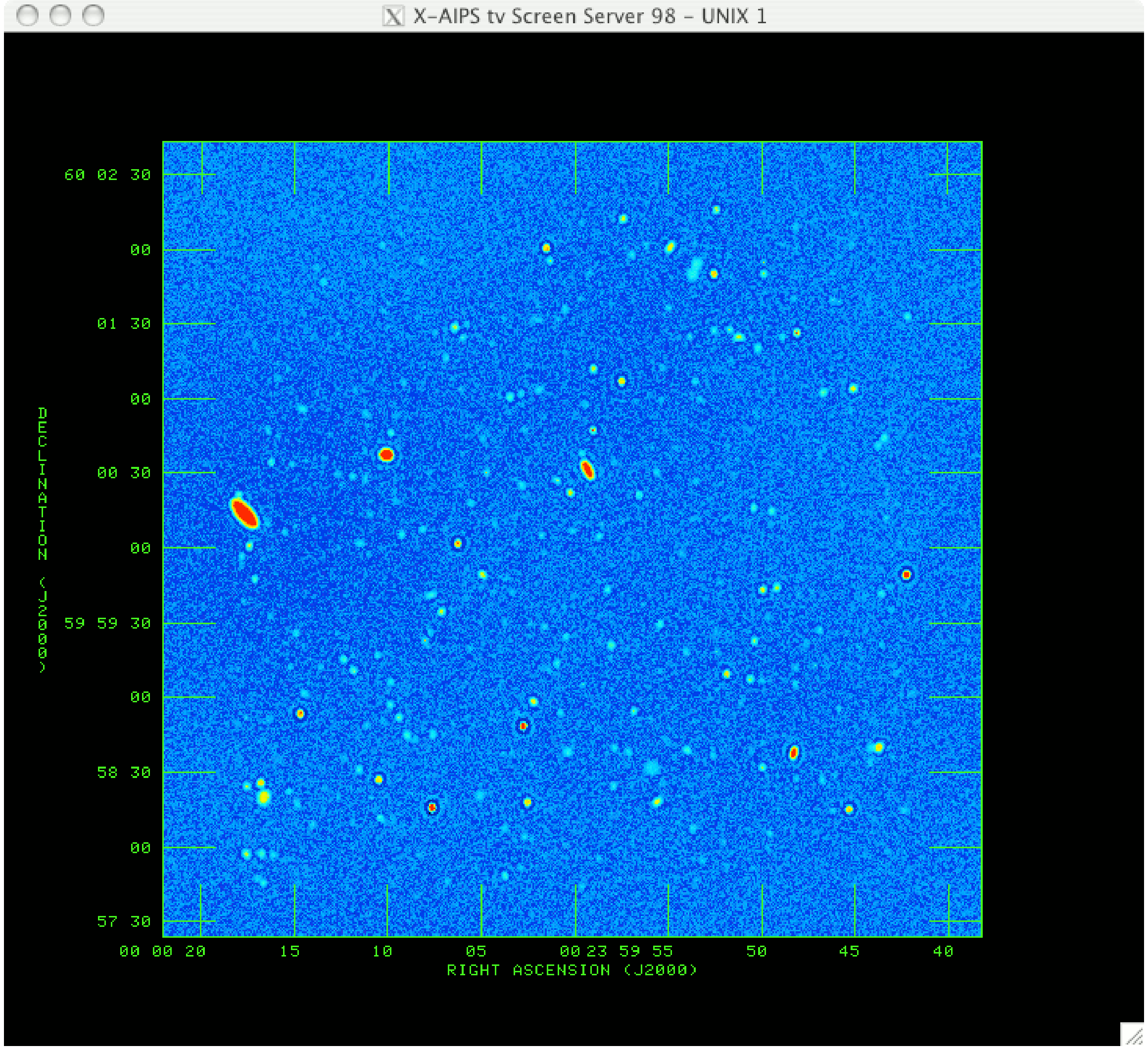
Post-Processing Options



Wilman simulation

Lensing model: Ben Metcalf

Sky as viewed by a simulated interferometer



SKADS

Wilman simulation

+ E-LOFAR simulation (Heywood)

STFC permitting, by~2010

Lensing model: Ben Metcalf

Concluding Remarks

- reference skies exist http://s-cubed.physics.ox.ac.uk

- please use them (for SKA simulations and real science) and provide feedback

- e2e data simulations now urgently needed, and tools will appear on this same web site as they are developed (map-maker, post-processing options, configurations, instrument simulators)

- must be linked to functional simulators and cost model

- must keep an eye on synergies with other facilities, e.g. EUCLID

• The Future: PrepSKA + Path2SKA: FP7 Sep2

Path2SKA - being organised by SR on behalf of ESKAC - strongly linked to international SKA project