Tools for measuring galaxy space densities from HI surveys

Martin Zwaan - ESO

HI stacking at z=1.3

DRAFT VERSION FEBRUARY 1, 2016
Preprint typeset using LATEX style emulateapj v. 5/2/11

THE GAS MASS OF STAR-FORMING GALAXIES AT $Z \approx 1.3$

NISSIM KANEKAR[1], SHIV SETHI[2], K. S. DWARAKANATH[2]

Draft version February 1, 2016

ABSTRACT

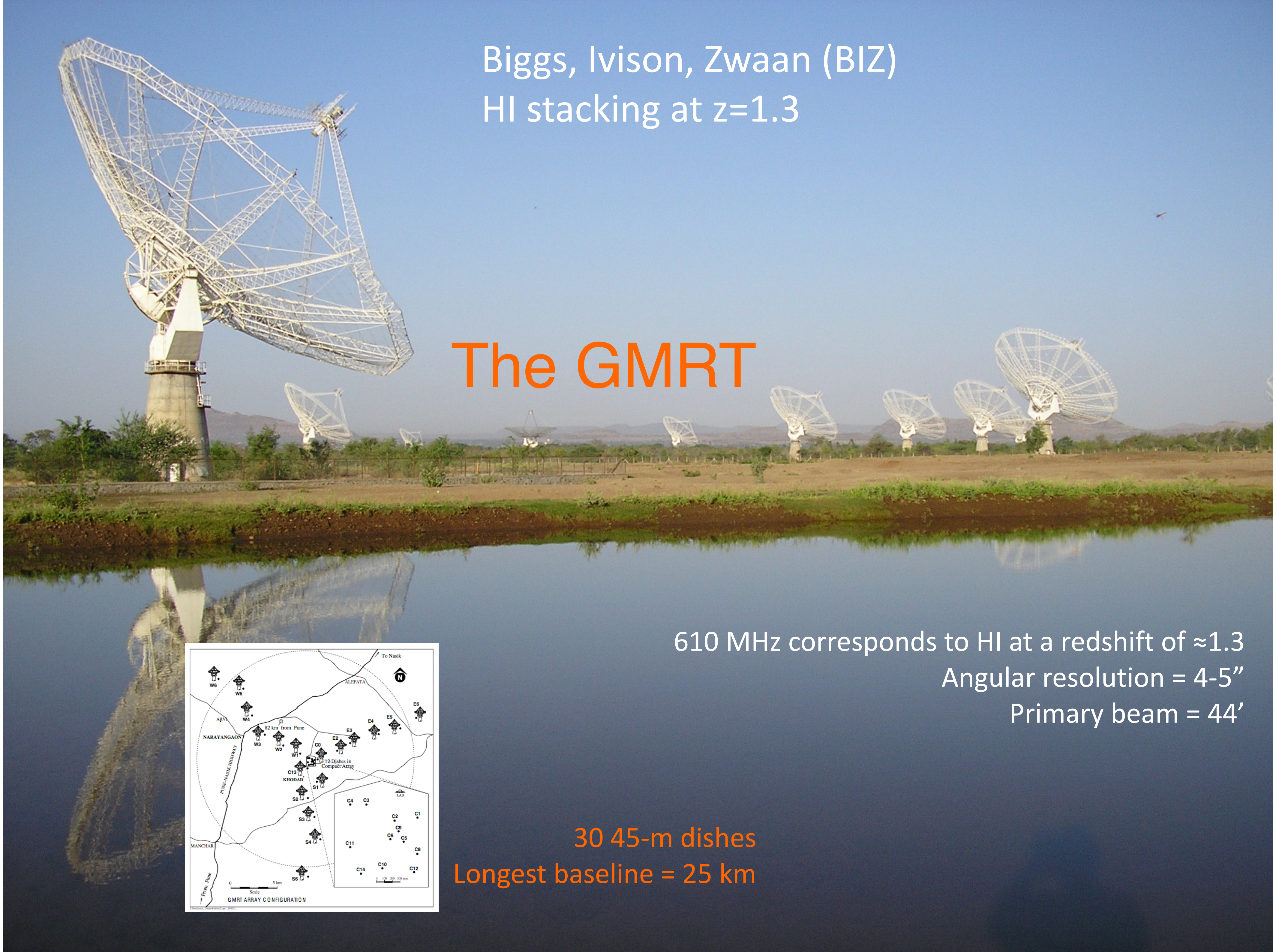
We report a Giant Metrewave Radio Telescope (GMRT) search for HI 21 cm emission from a large sample of star-forming galaxies at $z \approx 1.18 - 1.34$, lying in sub-fields of the DEEP2 Redshift Survey. The search was carried out by co-adding ("stacking") the HI 21 cm emission spectra of 857 galaxies, after shifting each galaxy's HI 21 cm spectrum to its rest frame. We obtain the 3σ upper limit $S_{\rm HI} < 2.5\mu$Jy on the average HI 21 cm flux density of the 857 galaxies, at a velocity resolution of ≈ 315 km s^{-1}. This yields the 3σ constraint $M_{\rm HI} < 2.1 \times 10^{10} \times [\Delta V/315 {\rm km/s}]^{1/2} M_\odot$ on the average HI mass of the 857 stacked galaxies, the first direct constraint on the atomic gas mass of galaxies at $z > 1$. The implied limit on the average atomic gas mass fraction (relative to stars) is $M_{\rm GAS}/M_* < 0.5$, comparable to the cold molecular gas mass fraction in similar star-forming galaxies at these redshifts. We find that the cosmological mass density of neutral atomic gas in star-forming galaxies at $z \approx 1.3$ is $\Omega_{\rm GAS} < 3.7 \times 10^{-4}$, significantly lower than $\Omega_{\rm GAS}$ estimates in both galaxies in the local Universe and damped Lyman-α absorbers at $z \geq 2.2$. Blue star-forming galaxies thus do not appear to dominate the neutral atomic gas content of the Universe at $z \approx 1.3$.

Subject headings: galaxies: high-redshift — galaxies: ISM — galaxies: star formation — radio lines: galaxies

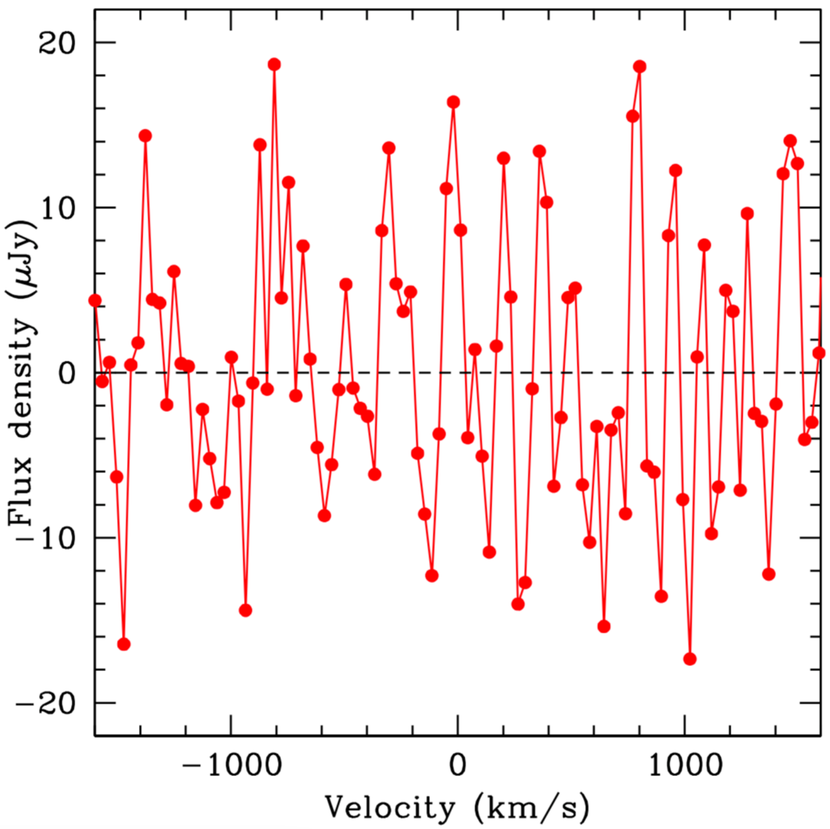
1. INTRODUCTION

Over the last decade, optical and infrared studies of the so-called "deep fields" (e.g. Giavalisco et al. 2004; Scoville et al. 2007; Newman et al. 2013) have revolutionized our under-

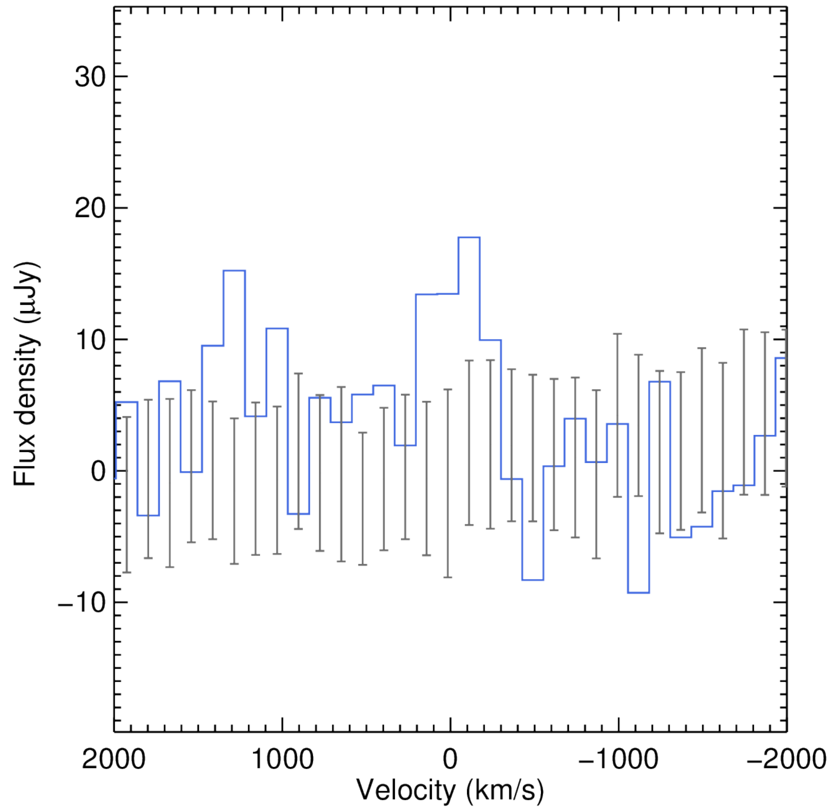
of the molecular component, studies of CO in emission at $z \approx 1.5 - 3$ have indeed found evidence for massive reservoirs of molecular gas ($M_{H_2} \gtrsim 10^{10} M_\odot$, comparable to the stellar mass) in star-forming galaxies (e.g. Daddi et al. 2010; Tacconi et al. 2010, 2013). Unfortunately, CO is only a tracer


Biggs, Ivison, Zwaan (BIZ)
HI stacking at z=1.3
The GMRT
610 MHz corresponds to HI at a redshift of ≈1.3
Angular resolution = 4-5”
Primary beam = 44’
30 45-m dishes
Longest baseline = 25 km
To Nasik
N
ALEFATA
ARVI
NARAYANGAON
82 km from Pune
PUNE-NASIK HIGHWAY
12 Dishes in Compact Array
KHODAD
MANCHAR
From Pune
Scale
5 km
GMRT ARRAY CONFIGURATION

Compare stacked HI spectrum of Kanekar and BIZ

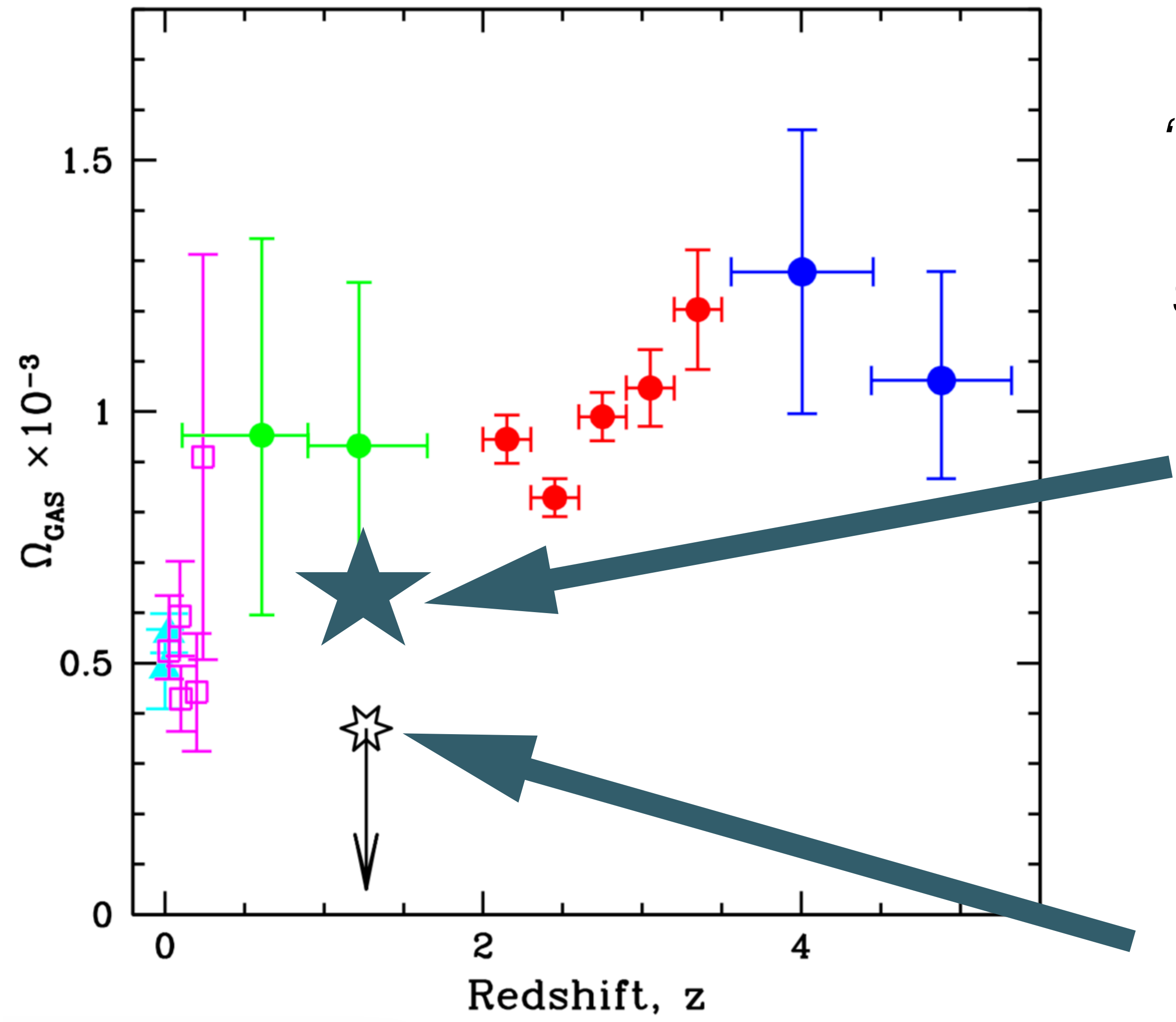


Kanekar et al



BIZ

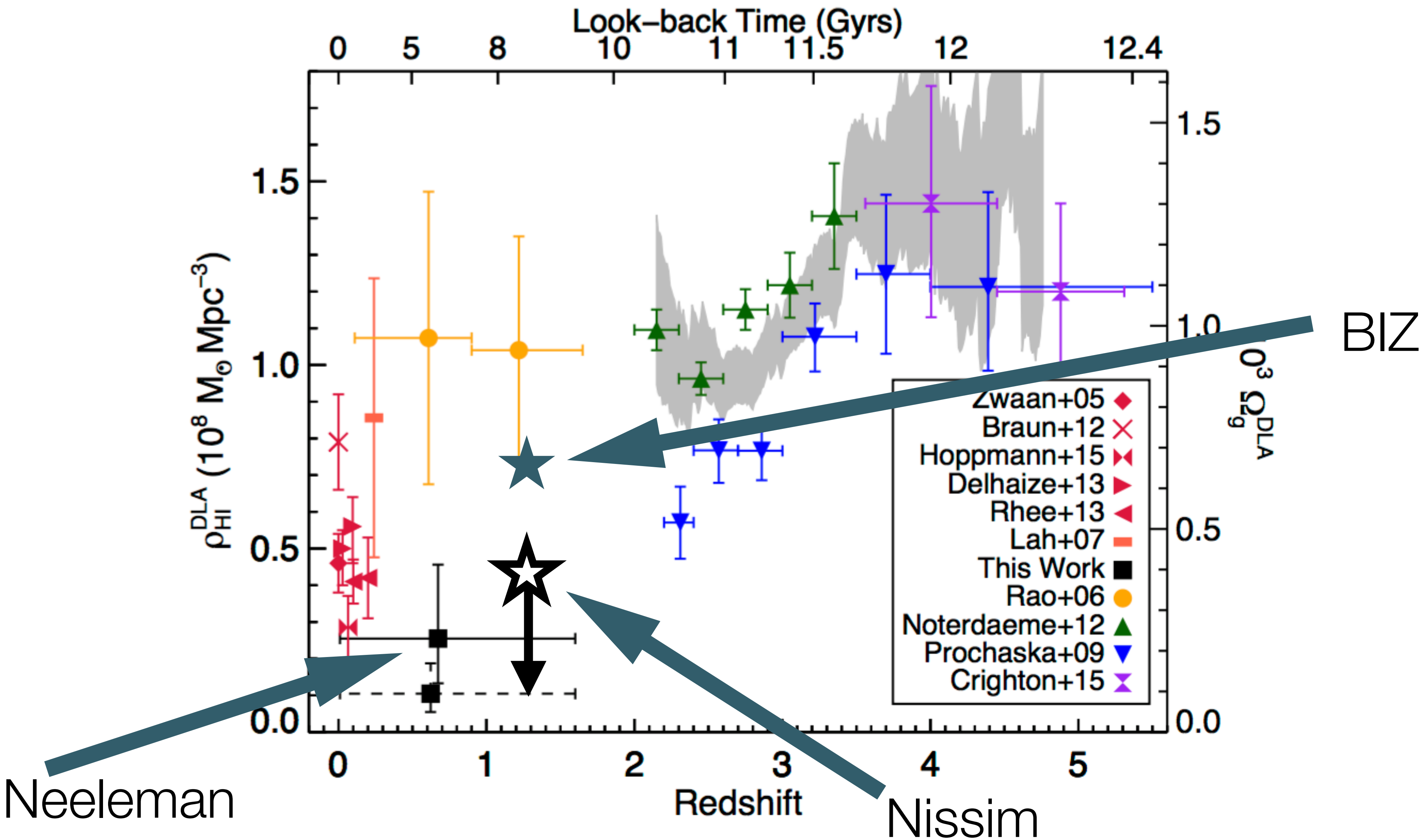
HI mass density



Lagos et al 2014:
'The HI density is predicted to be always dominated by galaxies with SFR< 1M$_\odot$ yr^{-1}'

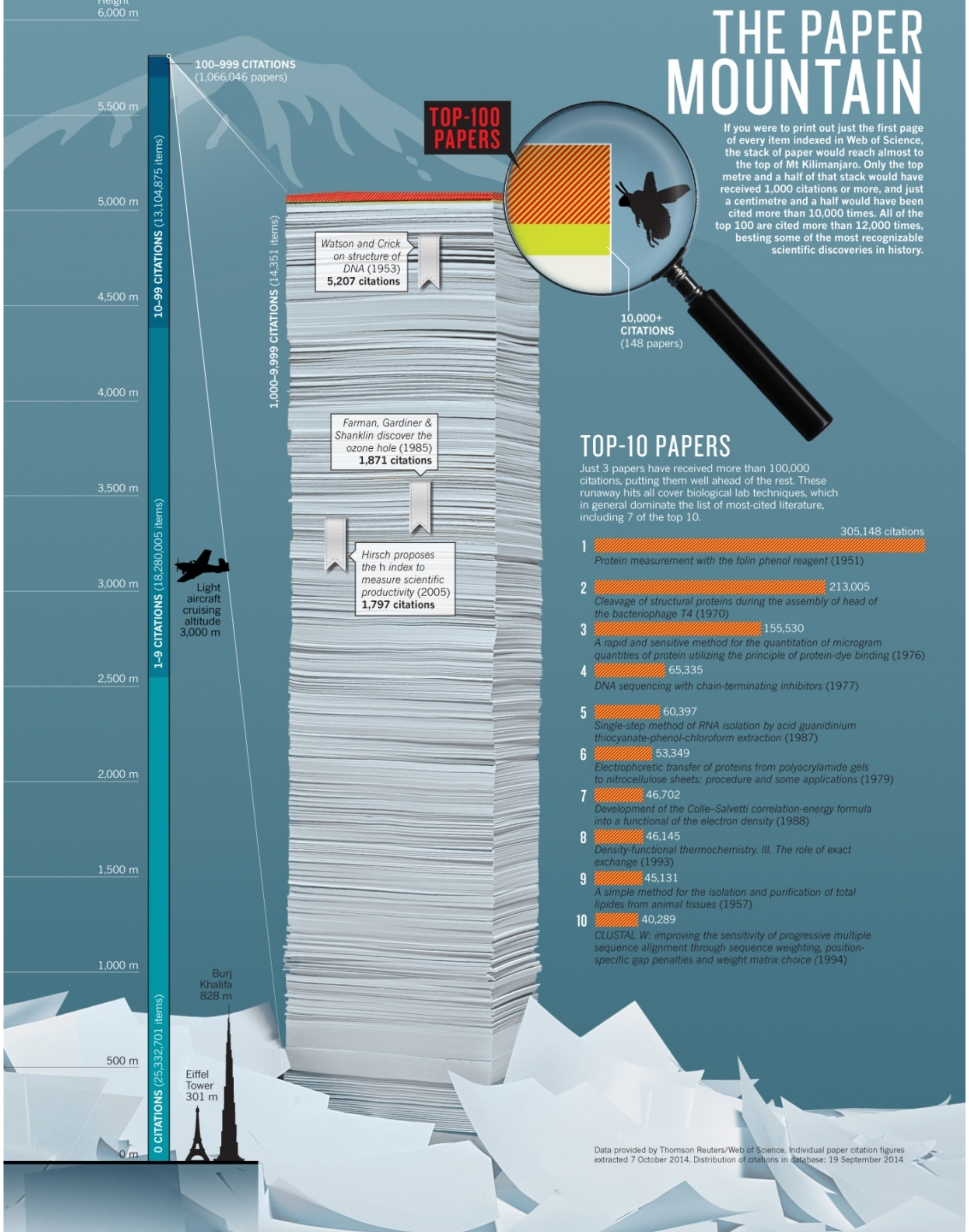
Kanekar et al. 2016

Latest DLA results… (Neeleman et al. 2016)
Four new DLAs at z<1.6



Tools for measuring galaxy space densities from HI surveys

Martin Zwaan - ESO



Top-cited papers present **methods**, not scientific results

source: van Noorden et al, Nature, 2014

Methods for measuring space densities of H I selected galaxies

Martin A. Zwaan,[1⋆] Martin Meyer et al.

[1] *European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching b. München, Germany*

Accepted ... Received ...

ABSTRACT

Key words:

1 INTRODUCTION

2 THE METHODS

2.1 The classical $1/V_{\rm max}$ method

2.2 The 2DSWML method

On galaxy samples that are other than flux-limited, other parameters have to be included in the maximum likelihood calculation of the space density of objects. For a given H I source at distance D, the peak flux is directly proportional to its H I mass $M_{\rm HI}$, and inversely proportional to the velocity width over which the flux is distributed. Therefore, the velocity width is the second parameter we have to include in the maximum likelihood analysis. A 2-dimensional maximum likelihood algorithm can then be employed to find the true galaxy distribution. In the following we use w_{20}, the profile width measured at 20% of peak intensity, as a measurement of the velocity width.

... ility of detecting galaxy i with H I mass $M^i_{\rm HI}$ and ...

where

$$j = 1, \ldots, N_M \;\; \text{and} \;\; k = 1, \ldots, N_W, \tag{4}$$

and N_M and N_W are the number of bins in M and W, respectively, and we define

$$M = \log(M_{\rm HI}) \text{ and } W = \log(w_{20}). \tag{5}$$

The logarithm of the likelihood of detecting all galaxies in the sample can now be expressed as

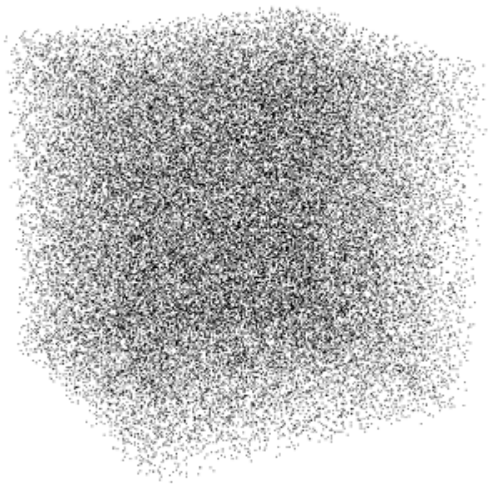
$$\ln \mathcal{L} = \sum_{i=1}^{N_{\rm g}} \sum_{j=1}^{N_{\rm M}} \sum_{k=1}^{N_{\rm W}} V(M_i - M_j, W_i - W_k) \ln \theta_{jk} - \sum_{i=1}^{N_{\rm g}} \ln \left(\sum_{j=1}^{N_{\rm M}} \sum_{k=1}^{N_{\rm W}} H_{ijk} \theta_{jk} \right) + c, \tag{6}$$

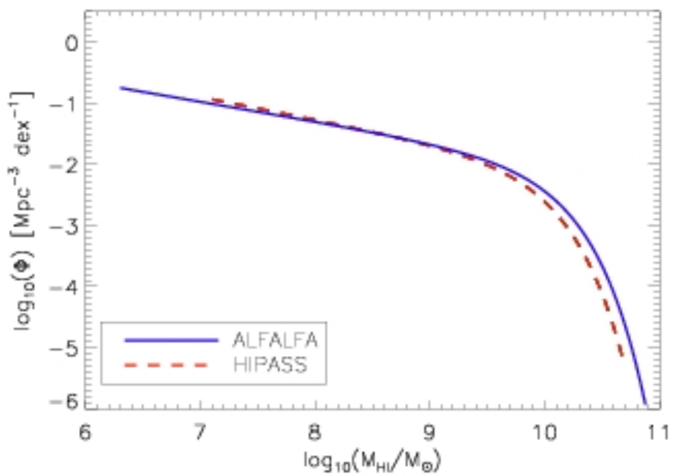
where c is a constant and V is a function defined by

$$V(x,y) = \begin{cases} 1, & |x| \leqslant \Delta M/2 \text{ and } |y| \leqslant \Delta W/2 \\ 0, & \text{otherwise} \end{cases}, \tag{7}$$

... over galaxies in bin $\Delta M \Delta W$. ...

The problem…





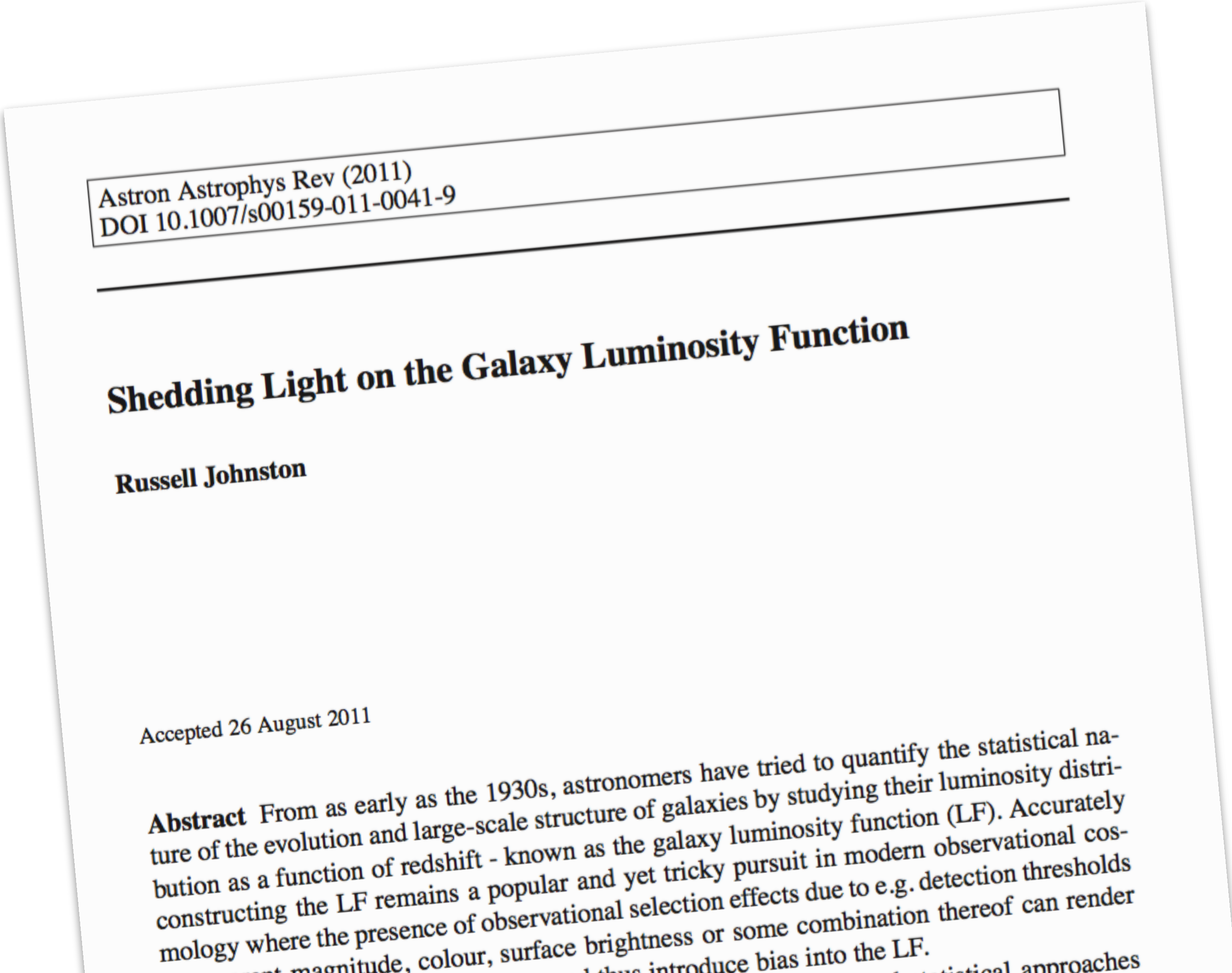
- Complications:
  - Complicated completeness limit (S_{peak}, W, profile shape, freq)

$\otimes$

  - Large scale structure

Isn’t this a solved problem?

- Yes, it is for surveys with a well-defined and uni-dimensional selection function



Astron Astrophys Rev (2011)
DOI 10.1007/s00159-011-0041-9

Shedding Light on the Galaxy Luminosity Function

Russell Johnston

Accepted 26 August 2011

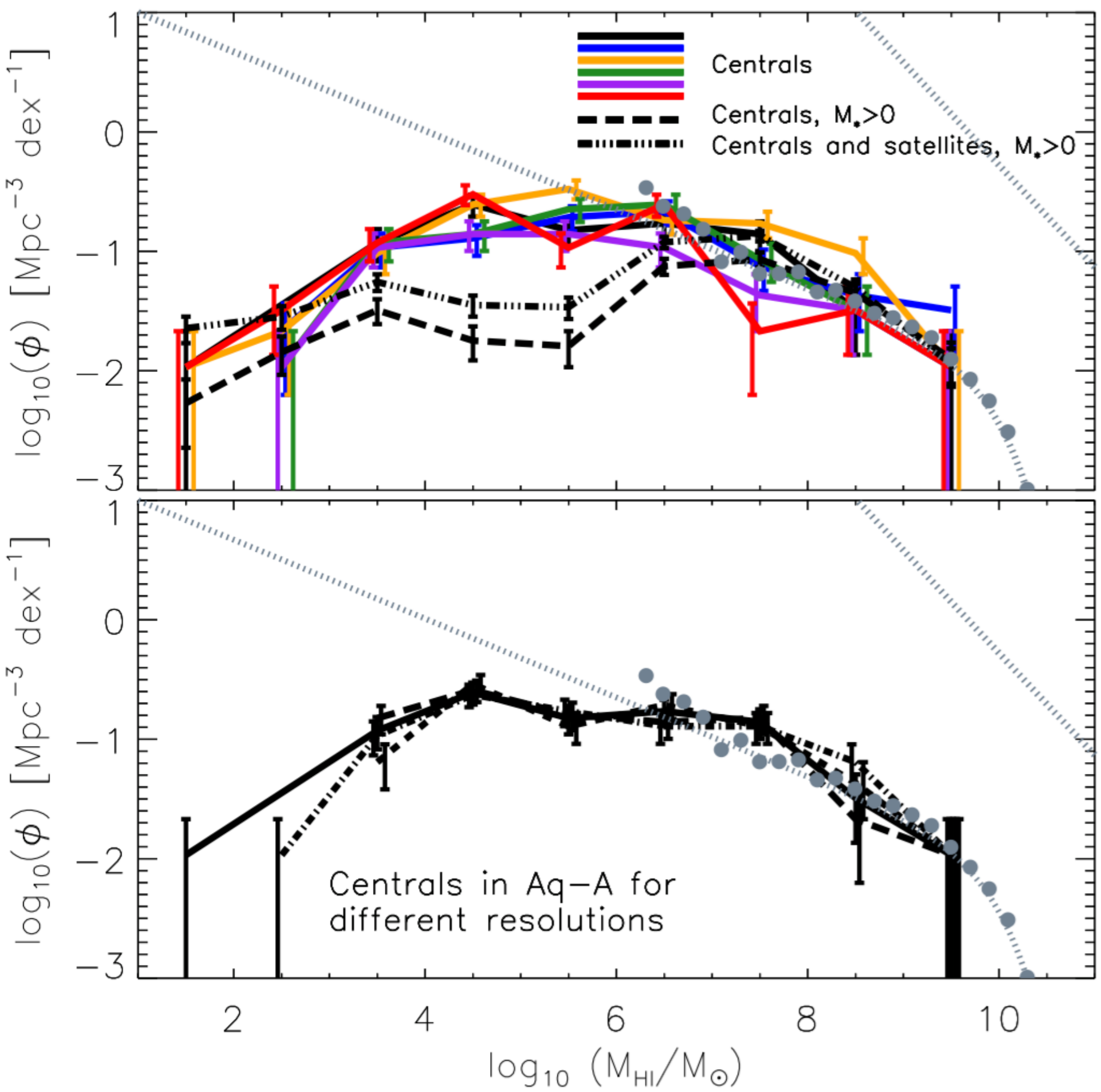
Abstract From as early as the 1930s, astronomers have tried to quantify the statistical nature of the evolution and large-scale structure of galaxies by studying their luminosity distribution as a function of redshift - known as the galaxy luminosity function (LF). Accurately constructing the LF remains a popular and yet tricky pursuit in modern observational cosmology where the presence of observational selection effects due to e.g. detection thresholds ... magnitude, colour, surface brightness or some combination thereof can render ... introduce bias into the LF.

HI mass function in the late 2010s…

- **Faint end**

HI mass function in the late 2010s…

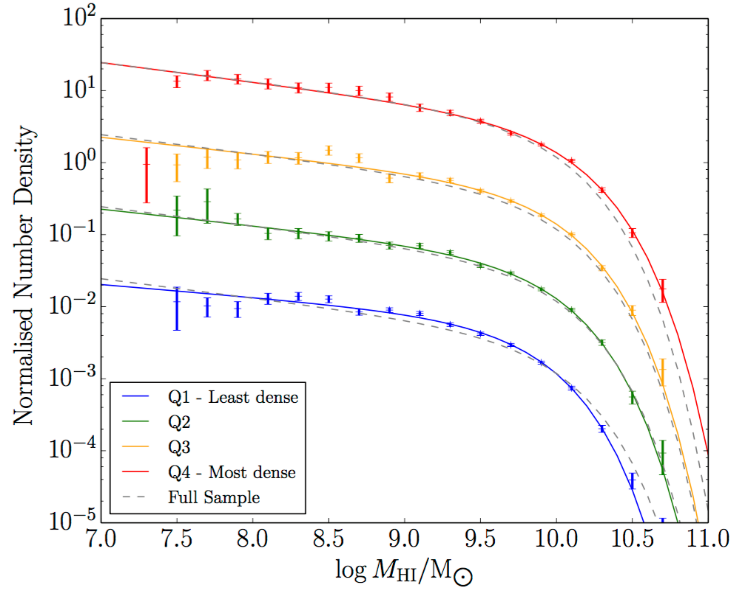
- **Faint end**



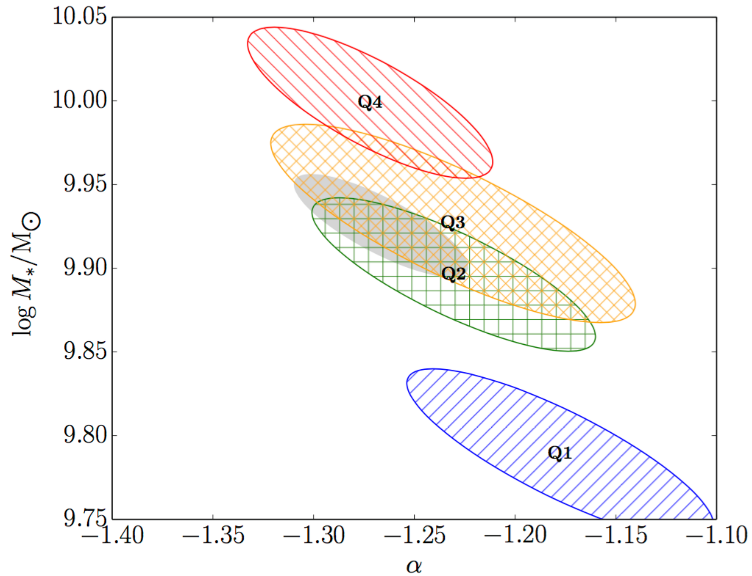
Yaryura et al. 2016: cosmological dark matter simulations coupled to the semi–analytic model, compared with measured HI mass function (and velocity function)

HI mass function in the late 2010s…

- **Faint end**
- **Environment**

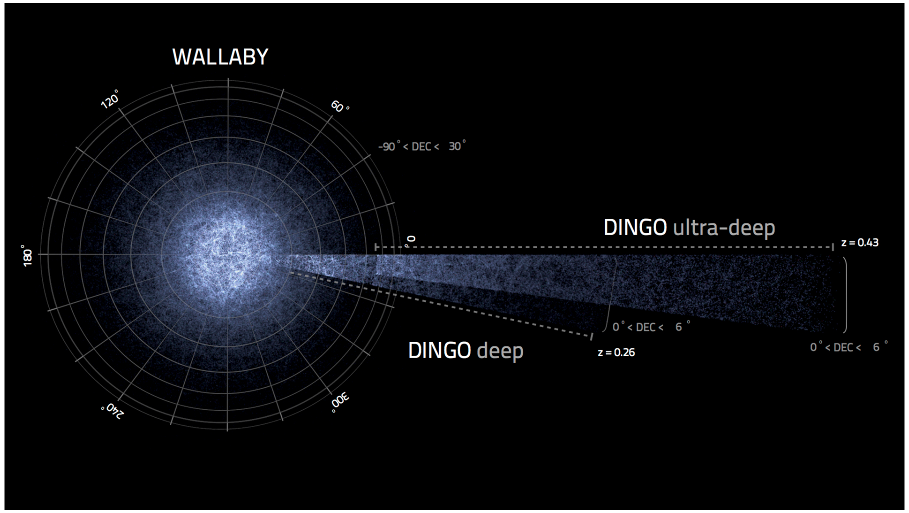


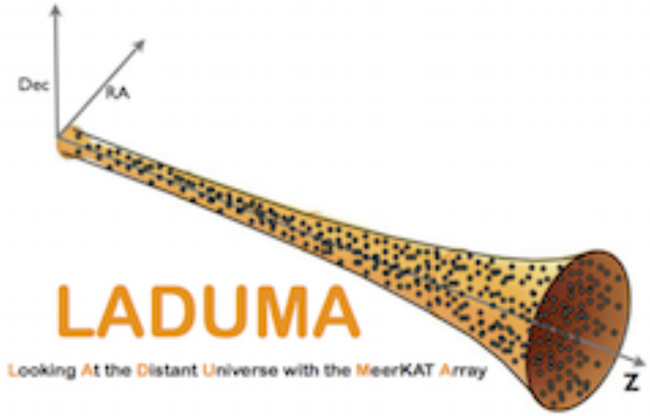
Jones et al. 2016:
HI mass function 'knee' is
dependent on
environment



HI mass function in the late 2010s…

- **Faint end**
- **Environment**
- **Evolution**

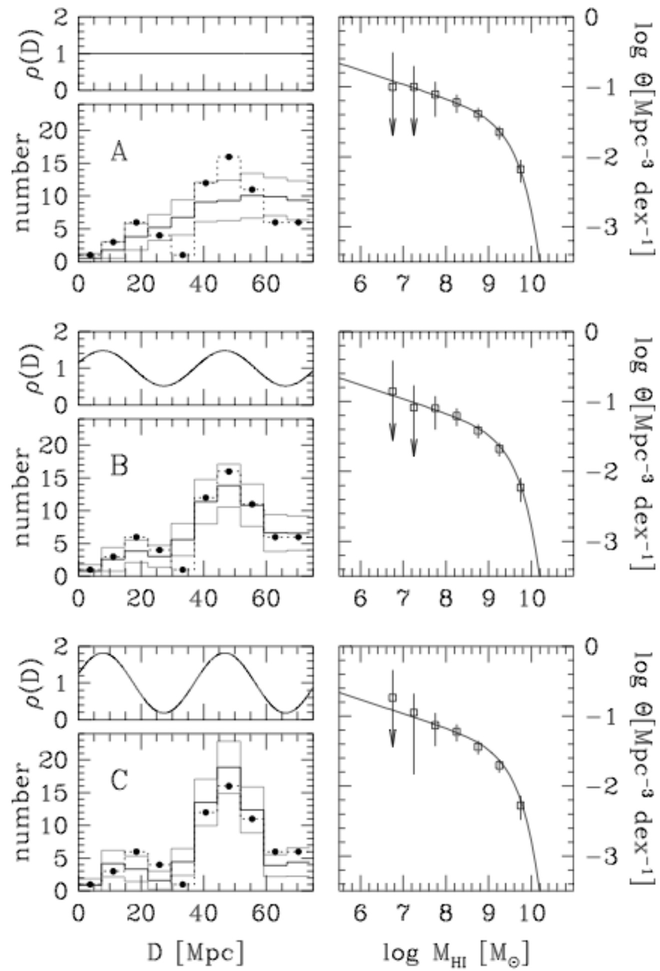




CHILES

The 1/V_{max} method

- The ‘classical’ Schmidt (1968) method
- Calculate maximum distance D_{max} out to which the galaxies can be detected
- Convert D_{max} into a V_{max}
- Used for early Arecibo surveys
- Advantages: simple and automatically normalised ✔
- Disadvantage: sensitive to large scale structure ✗



Zwaan et al 1997

Maximum likelihood methods

- Defined by Efstathiou et al 1988, Sandage et al 1979
- Find θ that yields maximal joint probability of detecting all sources in sample

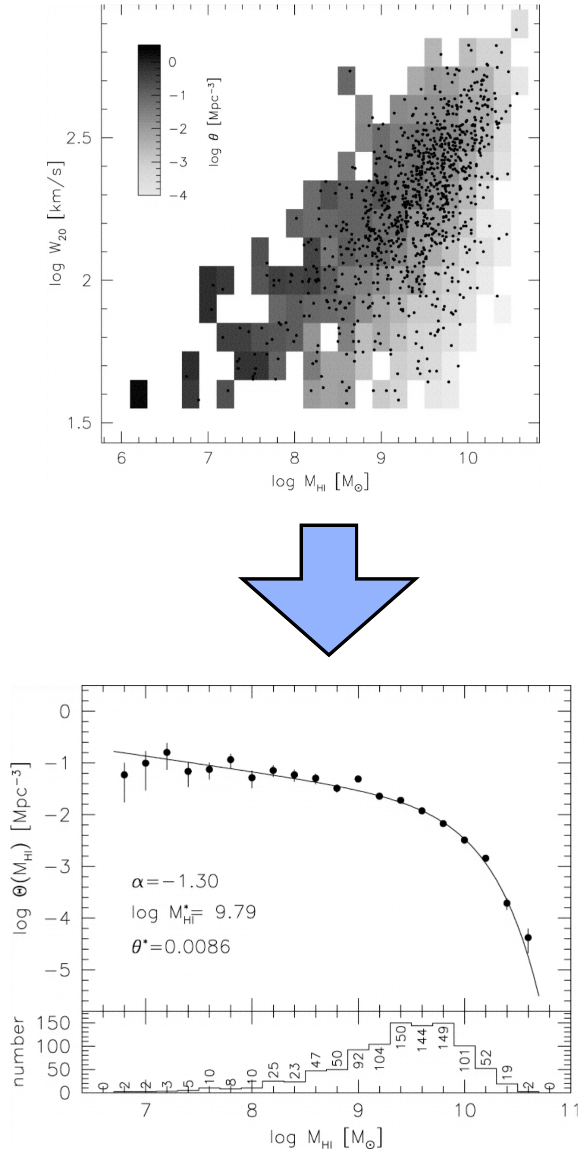
$$p(M_{\mathrm{HI},i}|D_i) = \frac{\theta(M_{\mathrm{HI},i})}{\int_{M_{\mathrm{HI,min}(D_i)}}^{\infty} \theta(M_{\mathrm{HI}})dM_{\mathrm{HI}}}$$

minimal detectable HI mass at distance D_i

generally **not defined** for HI selected samples

2D Stepwise Maximum likelihood method

- Solution: multi-dimensional stepwise maximum likelihood methods
- Find θ(M_{HI},W)
- Collapse to find HIMF
- **Used for HIPASS and ALFALFA**
- Advantage: robust against LSS ✔
- Disadvantage: slow ✗



The Turner or φ/Φ-method

- Introduced by Turner (1979) for 3C and 4C quasar catalogues
- Calculate the ratio of number of galaxies in interval dM_{HI} and number of galaxies brighter than M_{HI}

$$
\begin{aligned}
Y(M'_{\rm HI})dM_{\rm HI} &= \frac{N(dM'_{\rm HI})}{N(\geqslant M'_{\rm HI})} \\
&= \frac{\theta(M'_{\rm HI})\rho(D)dM_{\rm HI}dV}{\int_{M'_{\rm HI}}^{\infty}\theta(M_{\rm HI})\rho(D)dM_{\rm HI}dV}
\end{aligned}
$$

$$
Y(M'_{\rm HI})dM_{\rm HI} = \frac{\theta(M'_{\rm HI})dM_{\rm HI}}{\Theta(M'_{\rm HI})} = \frac{d\Theta(M'_{\rm HI})}{\Theta(M'_{\rm HI})}
$$

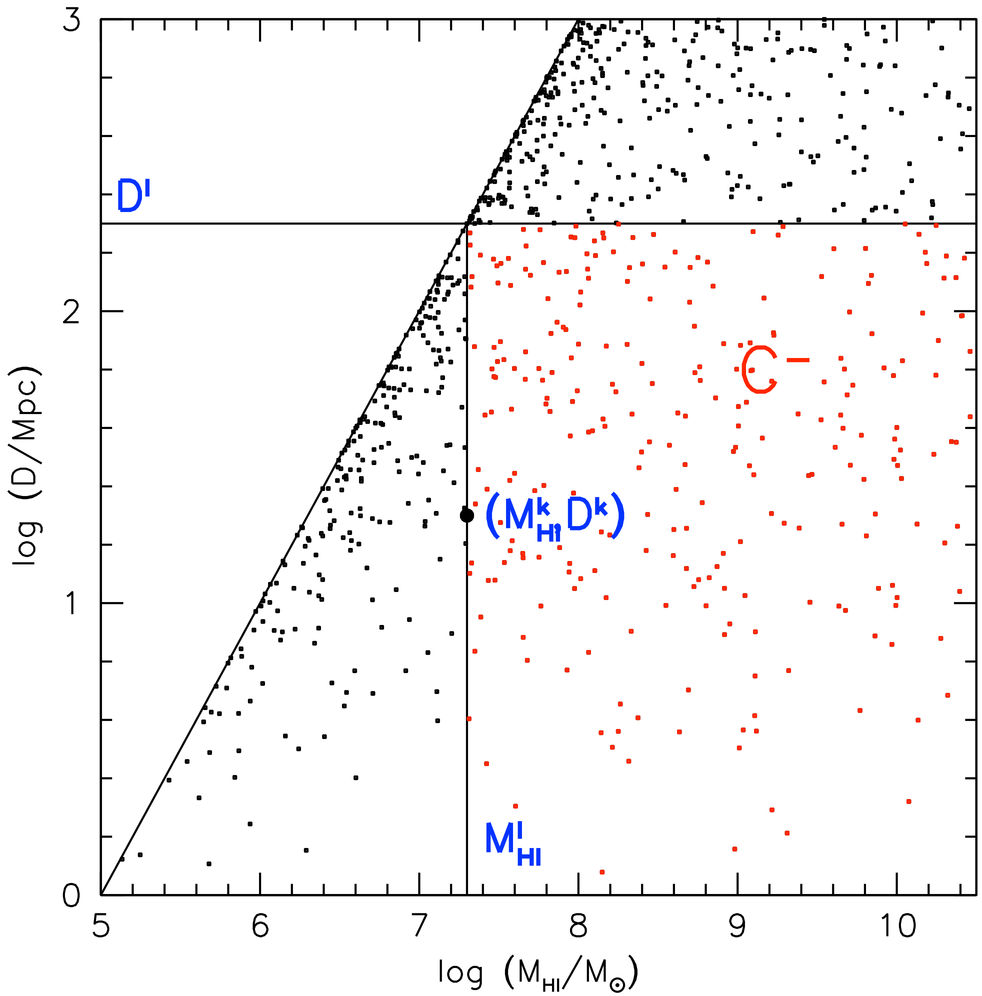
- Advantage: fast and robust against LSS ✔
- Disadvantage: correlated errors ✗

The C^- method

- Developed by Lynden-Bell (1971) for quasars.
- Does not require any binning.
- Does not require any assumptions about the distribution of objects within the data-set.
- Estimates the cumulative luminosity function (CLF).

The C⁻ method

$$\Theta(M^i_{\rm HI}) = \int_{M_{\rm HI}}^{\infty} \theta(M_{\rm HI})dM_{\rm HI}$$
$$= A \prod_k^{M^k_{\rm HI} > M^i_{\rm HI}} \frac{C^-(M^k_{\rm HI}) + 1}{C^-(M^k_{\rm HI})}.$$

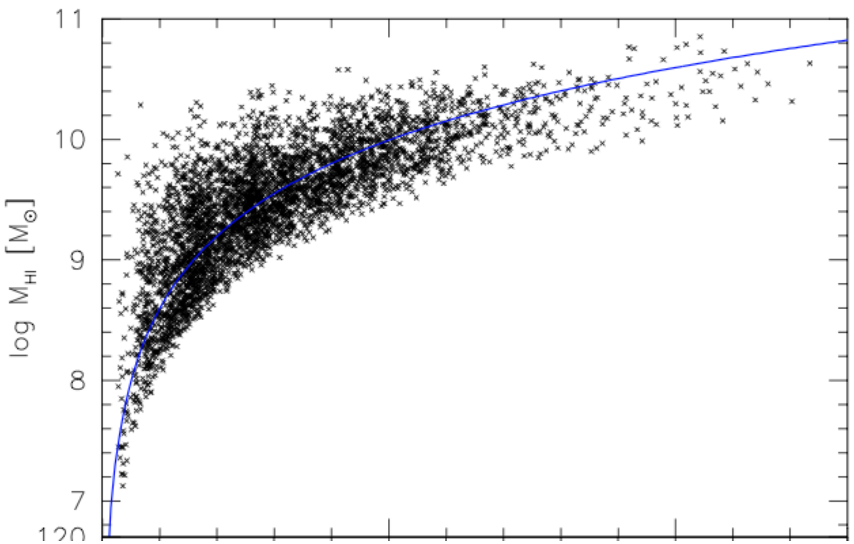


- Then differentiate to obtain $\theta(M_{\rm HI})$
- Variation of C⁻ is the C⁺ method (Zucca et al. 1998)
- Advantage: independent of clustering effects ✔ and fast ✔

Dealing with gradual drop off in completeness (as opposed to sharp flux limits…)

- All these methods are designed for optical galaxy samples with sharp magnitude limits (m_{lim})

- The 2DSWML, Turner, and C- method are easily adaptable to work for complicated completeness limits



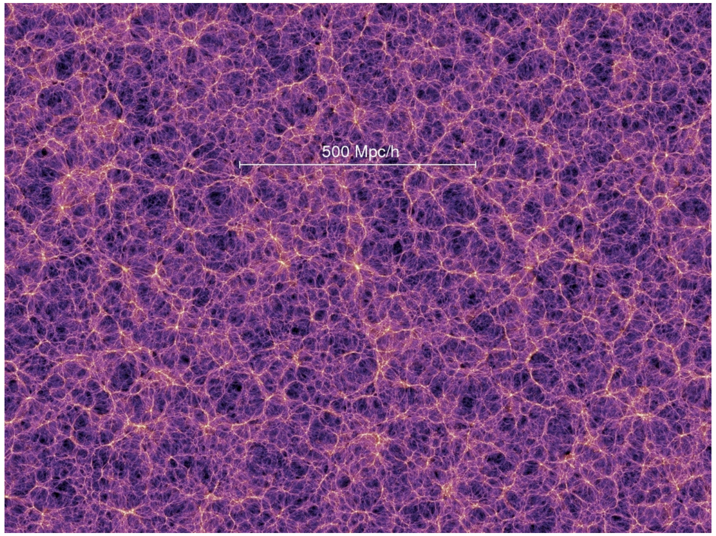
$$C_i^- = C^-(M_{\rm HI}^i) = \sum_j^{M_{\rm HI}^j > M_{\rm HI}^i} c,$$

Effective volumes

- We always calculate $V_{\mathbf{eff}}^{i}$ per individual galaxy
- These need to be summed per HI mass bin to get the HIMF
- These values can be binned as a function of local density, morphological type, etc.

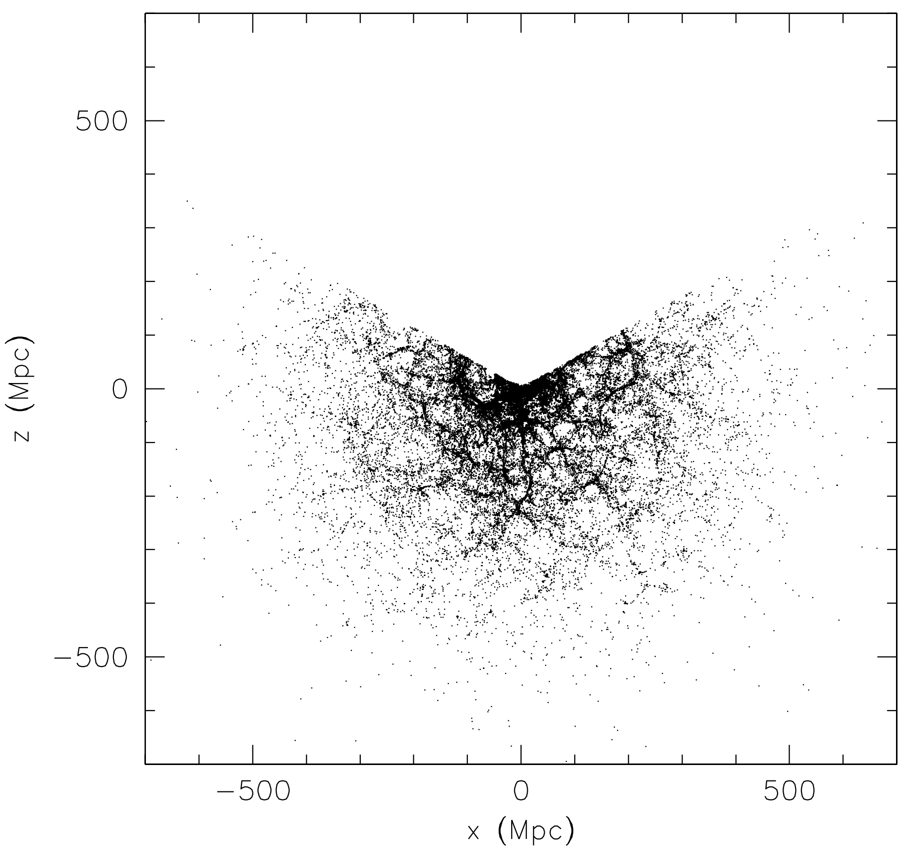
Simulations to test HIMF recovery

- Millennium Simulation (Springel et al 2005)
- Assume a HIPASS HI mass function
- Low mass (log M_{HI}<8.5) cluster around larger ones
- Realistic scatter on all parameters
- Select galaxies from simulated boxes, assuming ‘optimal smoothing’

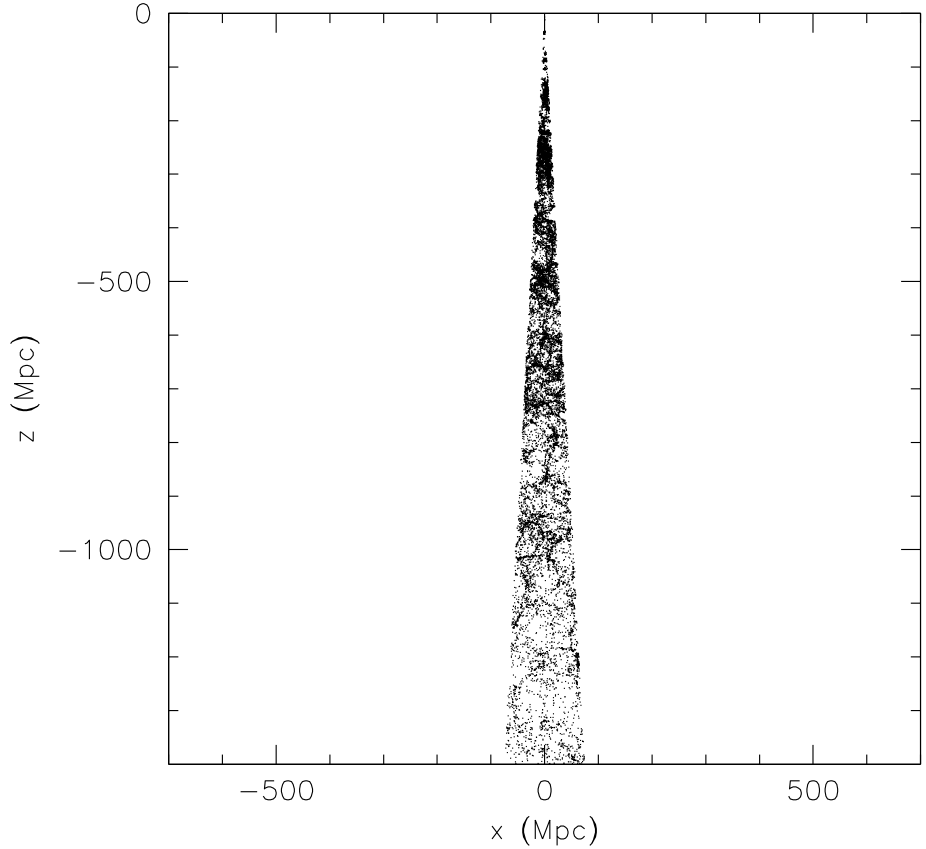


Simulated HI skies

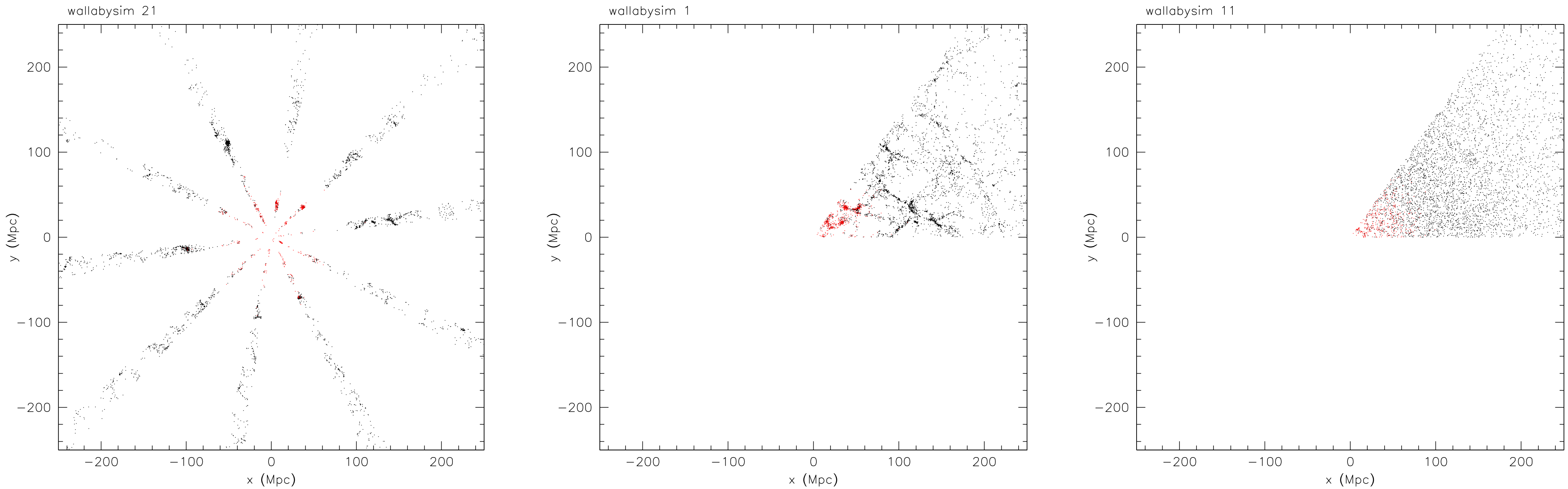
“Wallaby”
all sky shallow



“Dingo”
30deg^2 deep



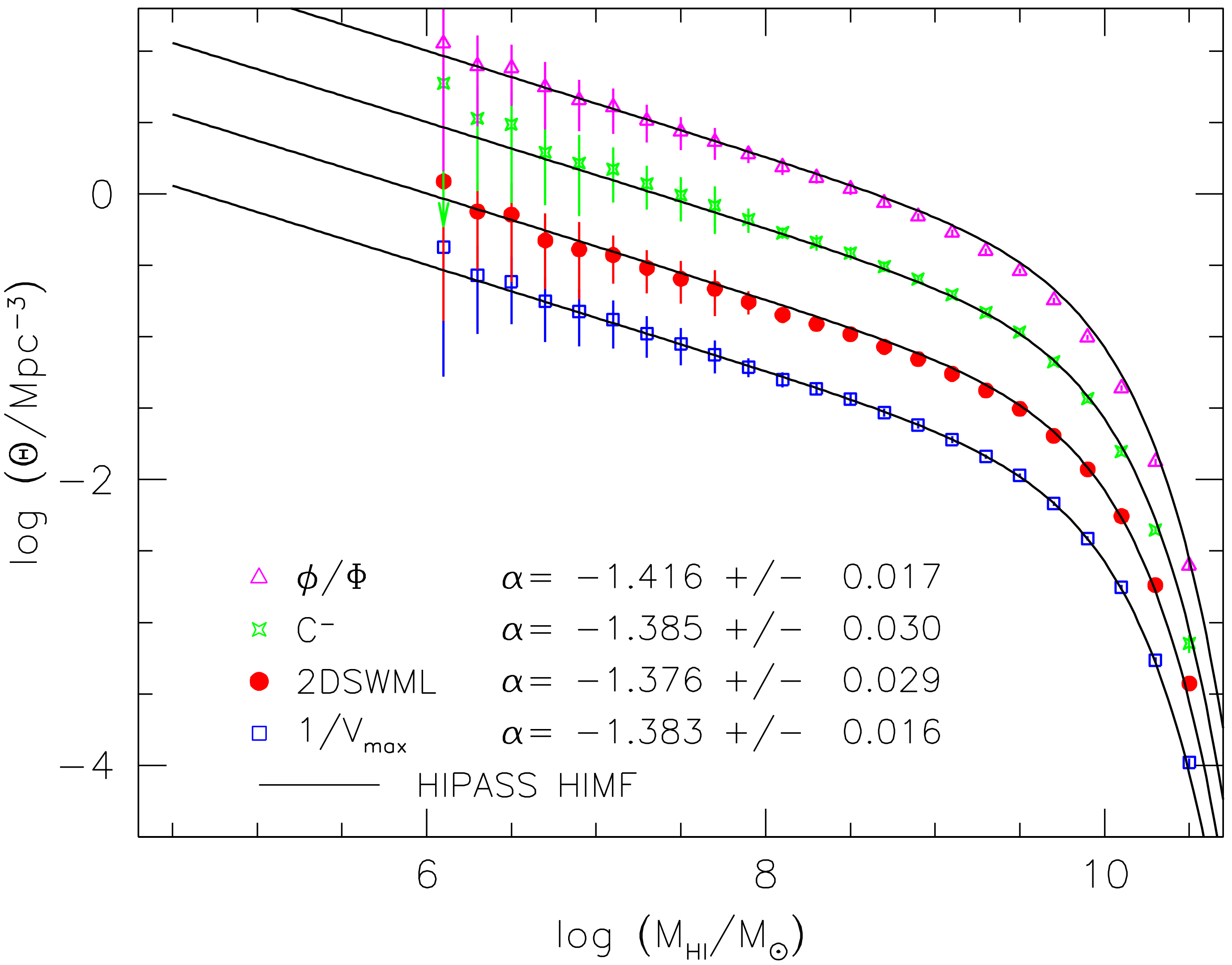
Testing the methods - 10 ASKAP pointings

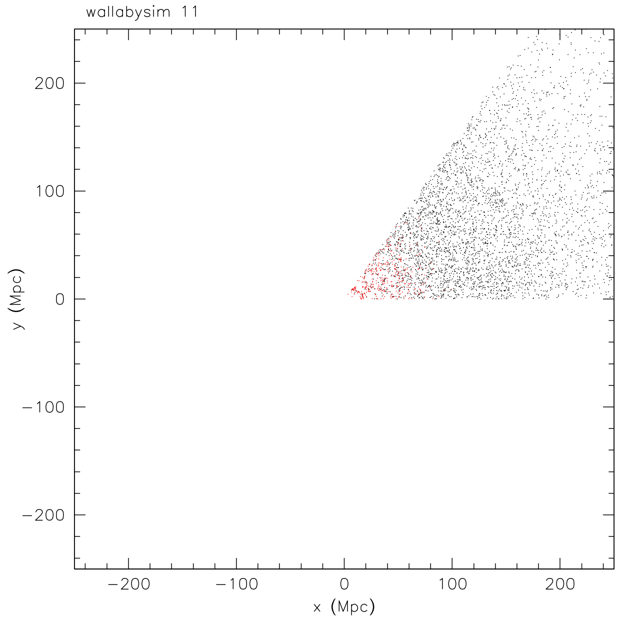


Simulate 10 times

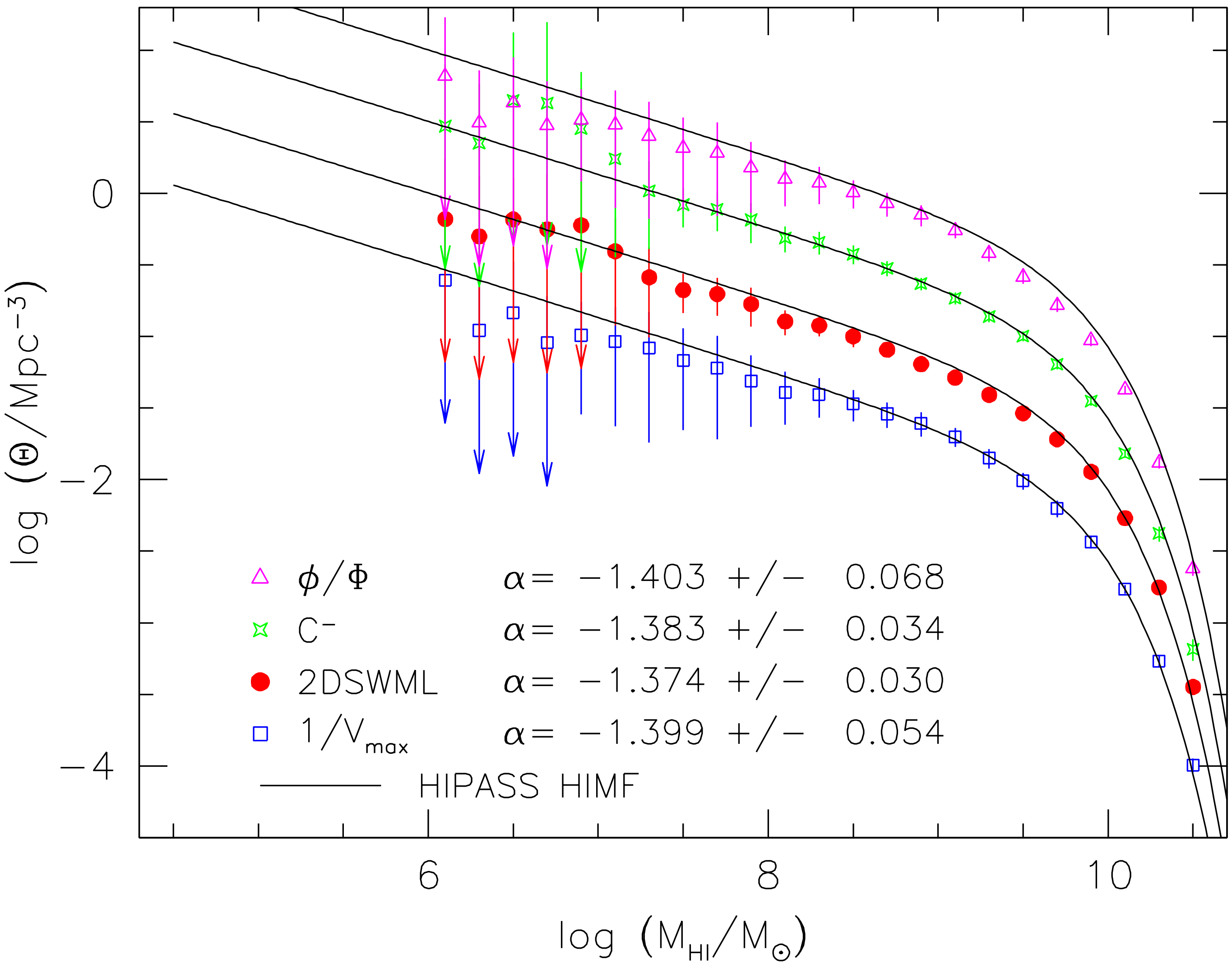
$1/V_{max}$	ML
φ/Φ	C^-

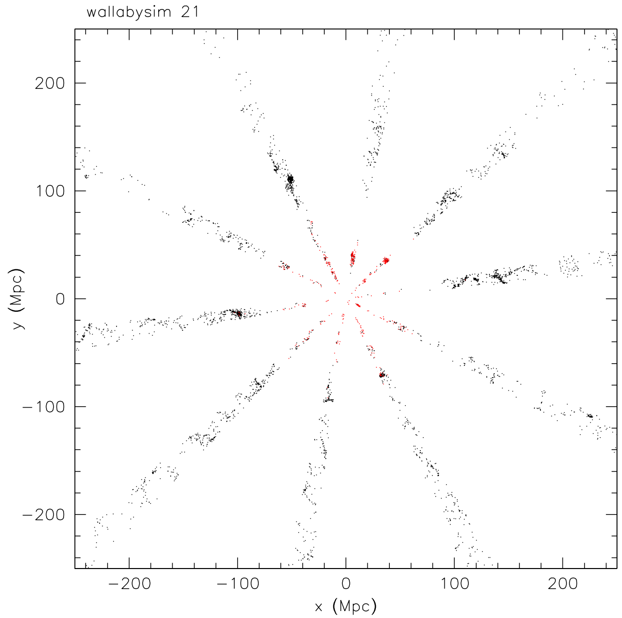
10 ASKAP pointings - No large scale structure



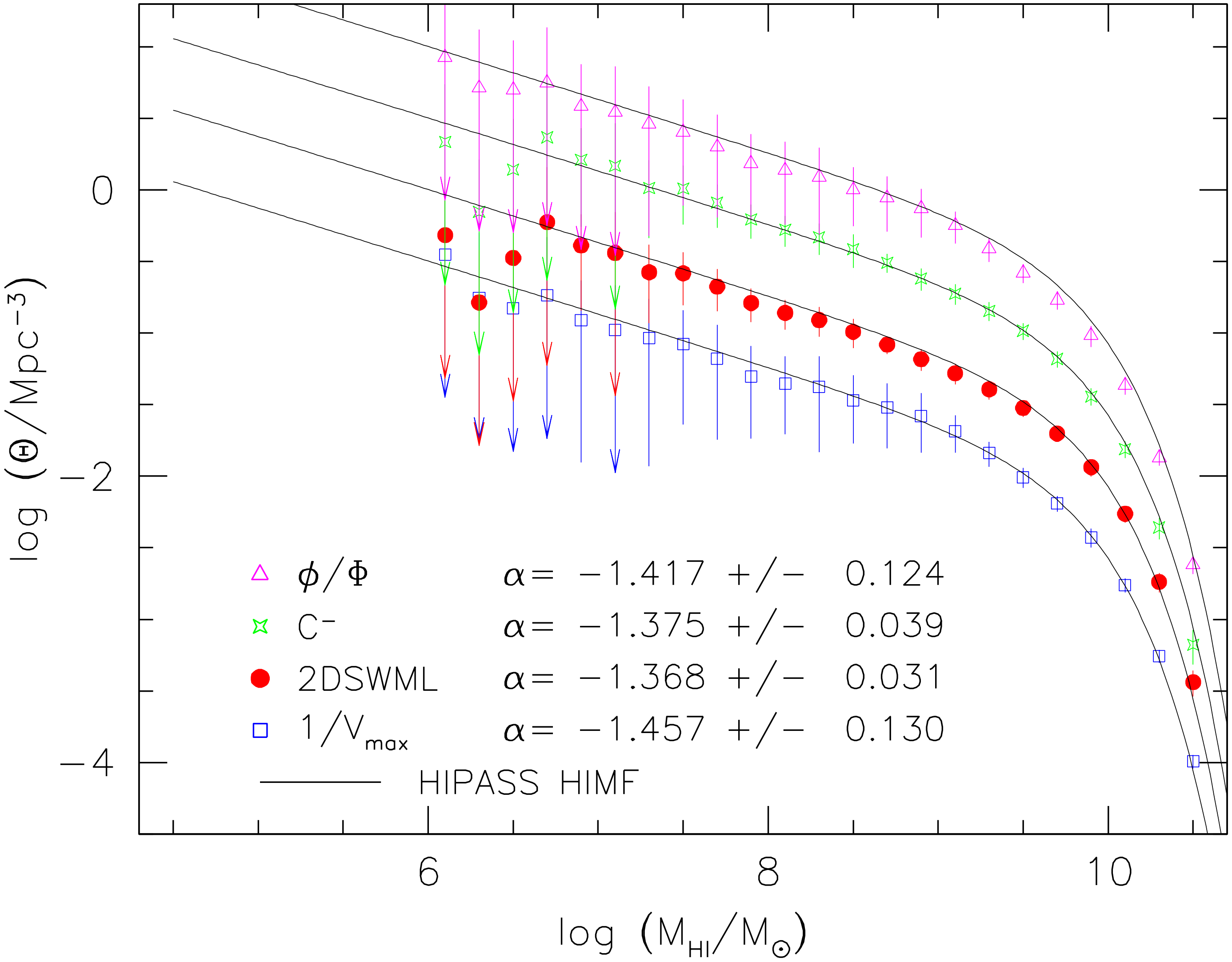


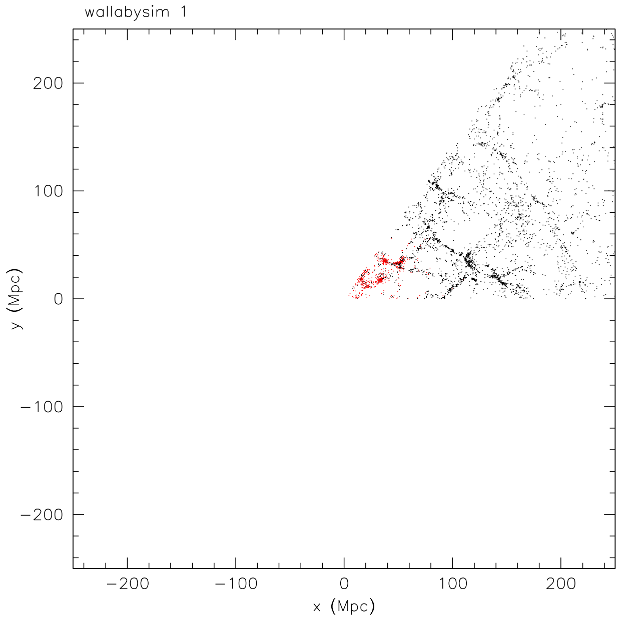
10 ASKAP pointings, with LSS - widely spaced





10 ASKAP pointings, with LSS- contiguous





Results of testing methods on shallow HI survey

As expected. Sensitive to LSS

Performs well. Robust against LSS

$1/V_{max}$	ML
φ/Φ	C^-

Not better than $1/V_{max}$

As good as 2DSWML. Faster!

Results of testing methods on shallow HI survey

$1/V_{max}$	ML
φ/Φ	C^-

As expected. Sensitive to LSS

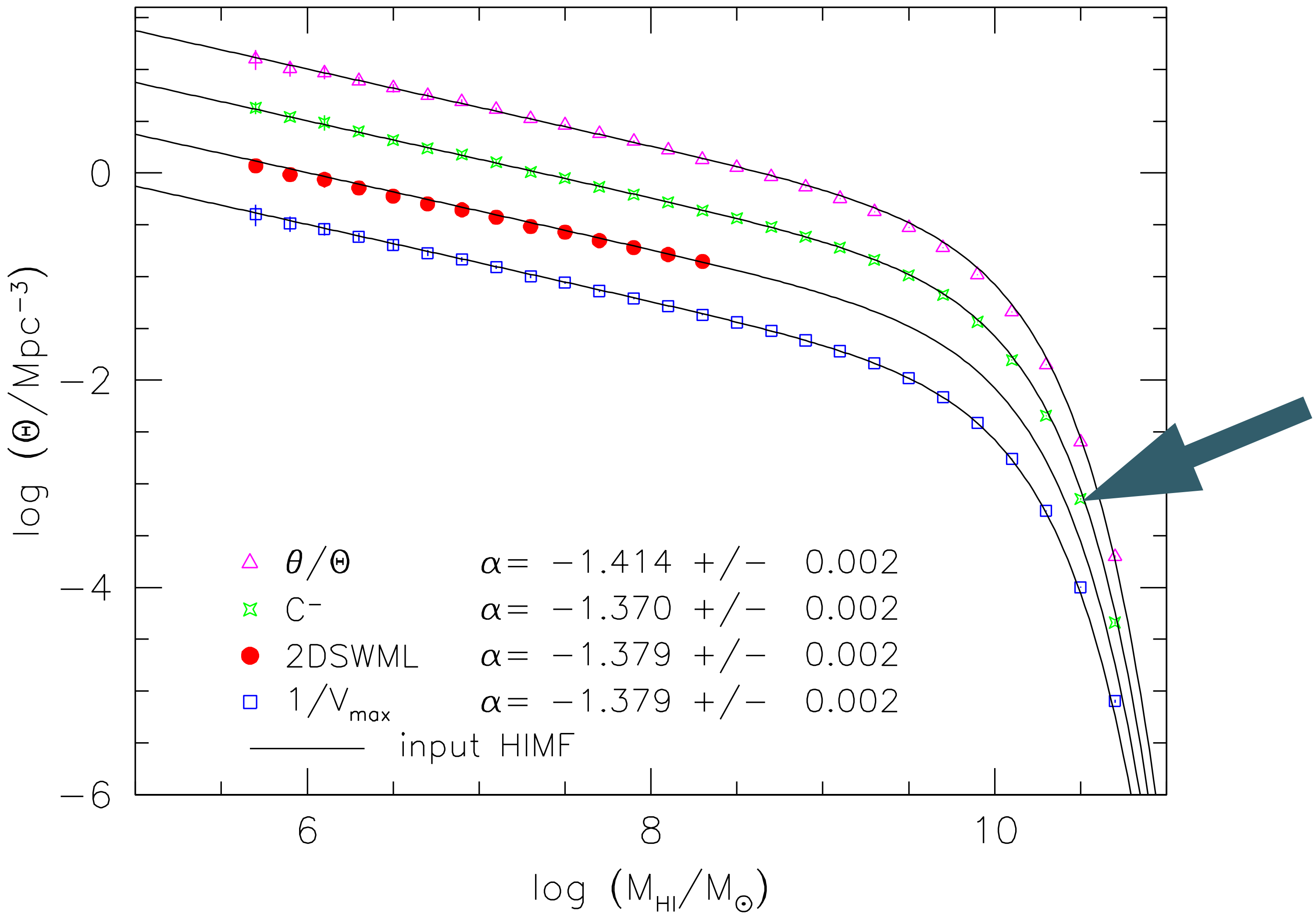
Performs well. Robust against LSS

Not better than $1/V_{max}$

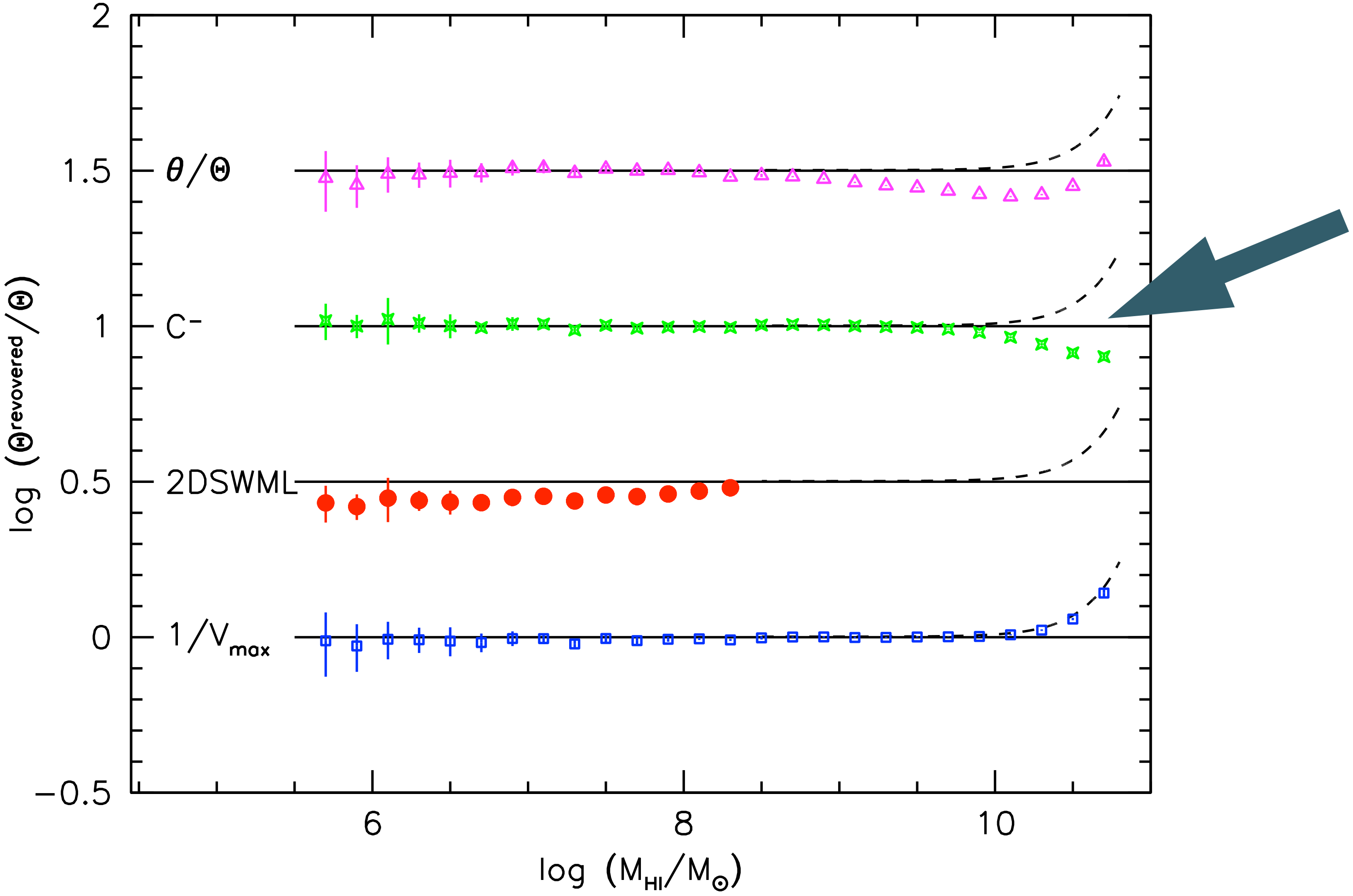
As good as 2DSWML. Faster!

But…

A full Wallaby-type survey, no large scale structure



C⁻ method underestimating space densities at high mass end



Normalising the space densities

- Most methods (apart from $1/V_{max}$) lose the normalisation of the HIMF

- See Davis & Huchra (1982), Willmer (1997) and Johnston (2011)

- First need selection function:

$$S(D) = \frac{\int_{W_{\rm low}}^{W_{\rm high}} \int_{M_{\rm lim}(D,W)}^{M_{\rm high}} \theta(M,W)dMdW}{\int_{W_{\rm low}}^{W_{\rm high}} \int_{M_{\rm low}}^{M_{\rm high}} \theta(M,W)dMdW},$$

- Then normalise:

$$\int_{M_{\rm low}}^{M_{\rm high}} \theta(M)dM = n.$$

Normalising the HIMF

- Various methods for recovering the normalisation
  - **n_3**: integral over selection function
  - **n_1**: calculating number of galaxies in redshift shells
  - **$\bar{n}$**: “minimum-variance” - weighting by selection function and second moment of correlation function
  - **counts**: compare real and expected number of galaxies

$$n_3 = \frac{N_T}{\int_0^{z_{max}} s(z)dz}$$

$$n_1 = \frac{\int_0^{zmax} \frac{N(z)}{s(z)} dz}{\int_0^{zmax} \frac{dV}{dz} dz}.$$

$$n = \frac{\sum_{i=1}^{N_g} N_i(z_i)w(z_i)}{\int_0^{z_{max}} s(z)w(z)\frac{dV}{dz}dz}$$

Normalising the HIMF

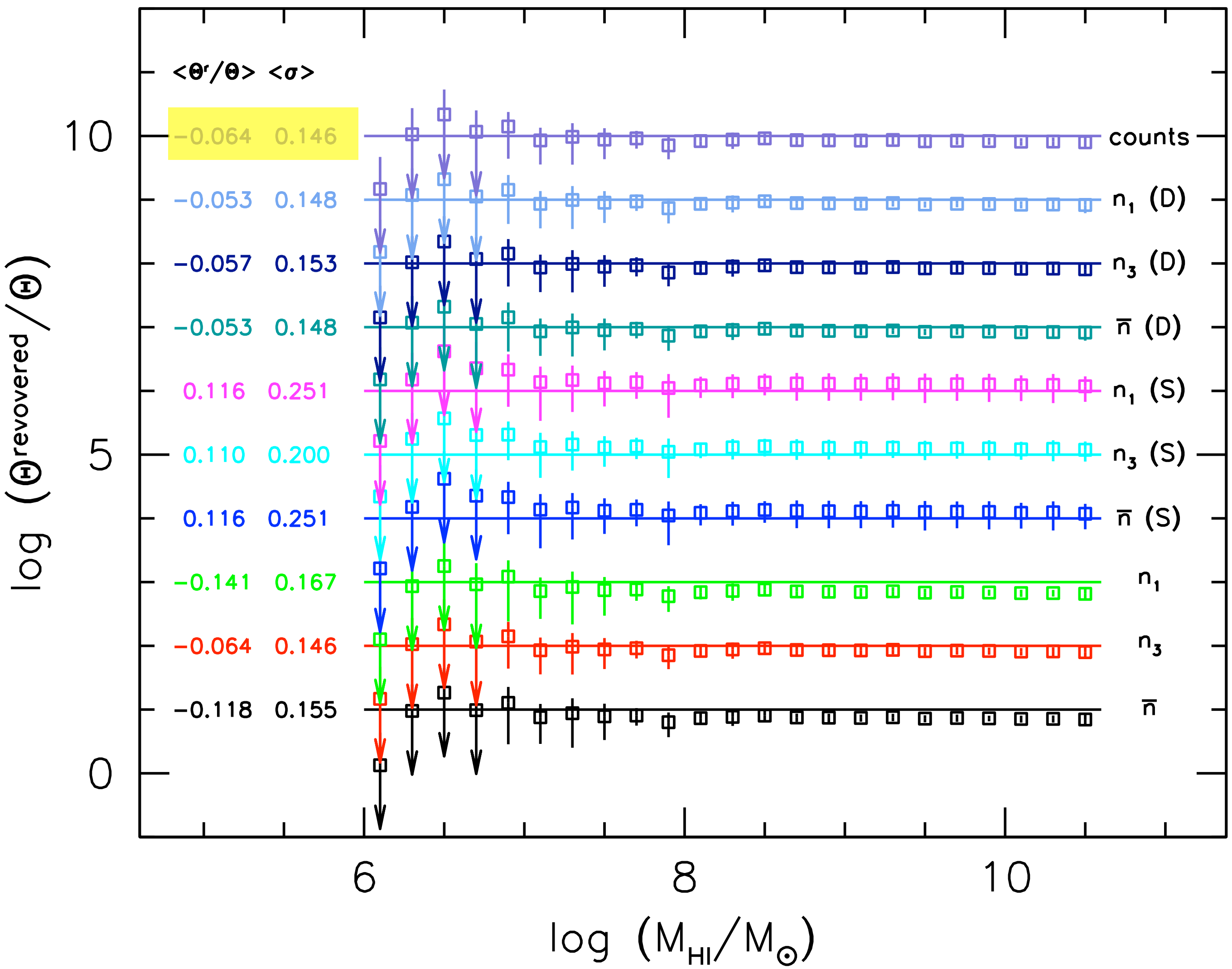
- Uncertainty of normalisation:

$$\sigma_{\bar{n}} = \left\{ \frac{\bar{n}}{\int dV\, S(D) w} \right\}^{1/2}$$

- For example, for HIPASS, the relative uncertainty on normalisation is at least ~5%.

Testing normalisations

- Normalising using 'counts' is very reliable



HIMFs from next generation HI surveys

- (Probably) use the C^- method
  - Robust against LSS
  - Works with 'soft' completeness limits
  - Fast
  - Can be used for HIMF evolution, and environment
  - But it has problem at high HI masses…
- Normalisation: use counts