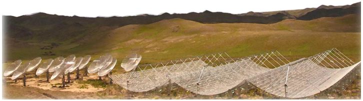
21cm X optical correlation & HI clumps with Tianlai low *z* surveys

O. Perdereau (*on behalf of the Tianlai collaboration*)

21 cm Cosmology Workshop 2023
&Tianlai Collaboration Meeting

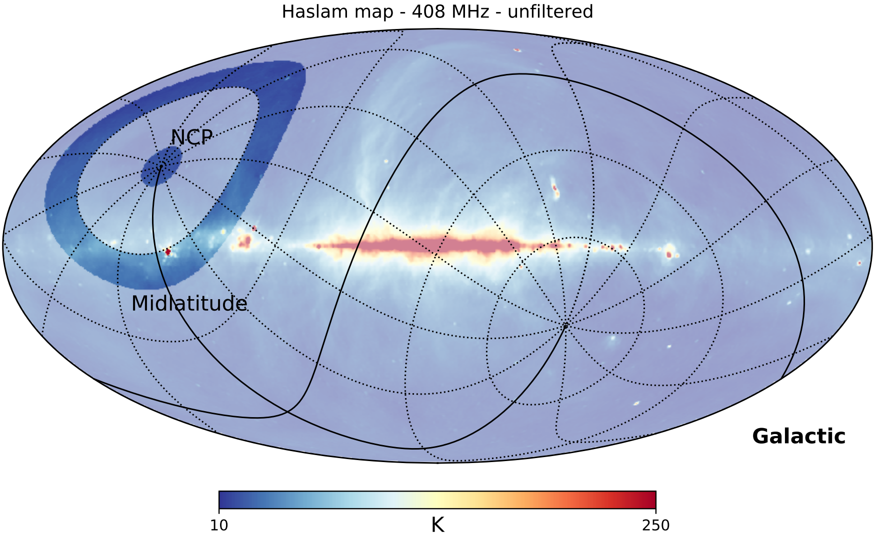
Northeastern University & National Astronomical Observatories CAS
July 16-22,2023. Shenyang, P.R. China



Overview

- Work published in MNRAS 517 (2022) 3, 4637-4655
- Coming soon : low-z observations with Tianlai ($\sim$ 1330-1430 MHz)
- Main purpose : "prove the effectiveness of the dense interferometric dish array and transit mode observations of Tianlai"
- two surveys : a mid-latitude band (6 scans around $\delta \sim 50$ deg) and the North Celestial Polar cap (4 scans around or at NCP)
- Forecasts : nearby HI clump (or galaxies) direct detection and cross-correlation with optical wavelength LSS surveys
- 3 parts :
  1. Simulation tools, efficiencies
  2. direct clumps detection
  3. cross-correlation with optical surveys (SDSS, NCCS)

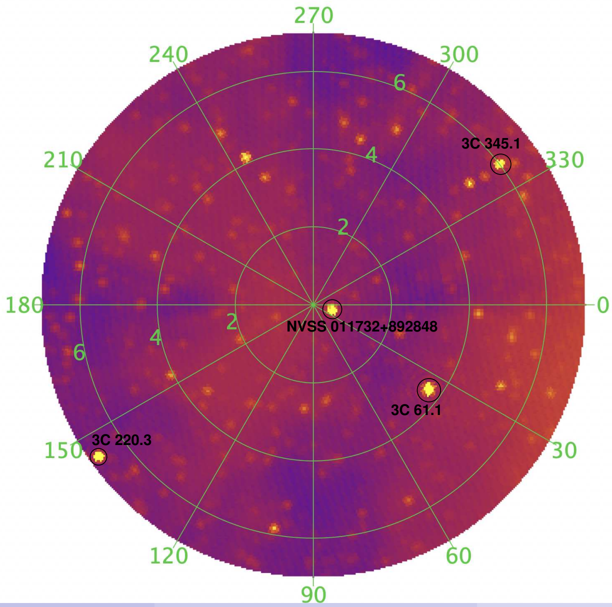
Footprints of surveys



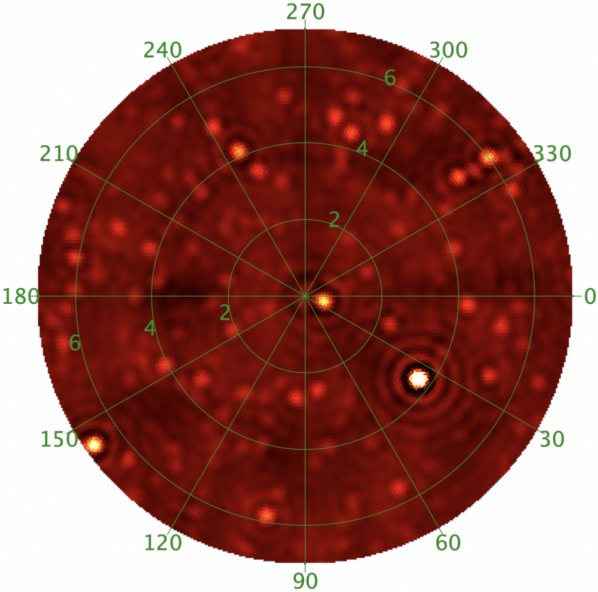
Toolbox

- Available in `JSkyMap` `gitlab.in2p3.fr` repository (R. Ansari et al.)
- `map2vis` = visibility simulator : combines sky components + convolution with beam in (ℓ, m) space given array config, frequency
- `vis2map` = map reconstruction in (ℓ, m) space : noise (photometry, phase) addition 'on the flight' (5mK per sample) ; pseudo-inverse pointing matrix saved
- `filt2map` filtering in (ℓ, m) space
- "post processing" : source detection, (cross-) C_ℓ computation (`spec_dfreq_ecp` module)
- foreground removal methods : local difference or polynom fit (in ν domain)

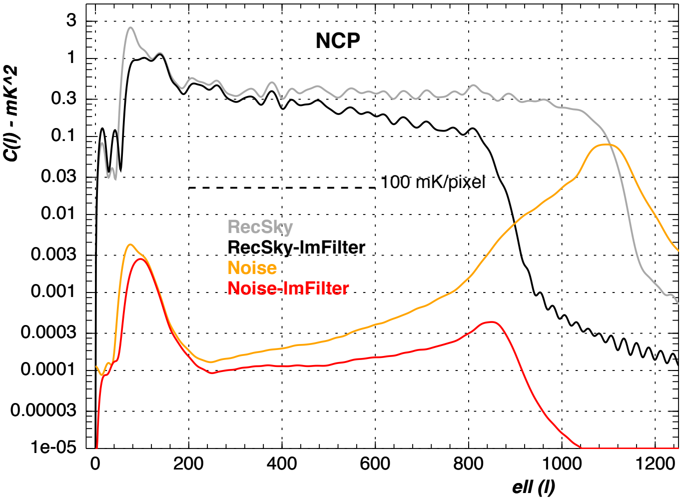
Input map (1350 GHz, no noise, NCP)



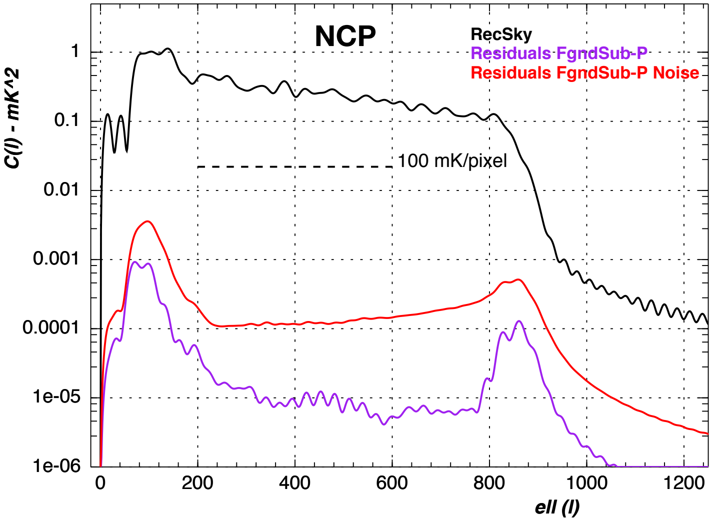
Reconstructed map (1350 GHz, no noise, NCP)



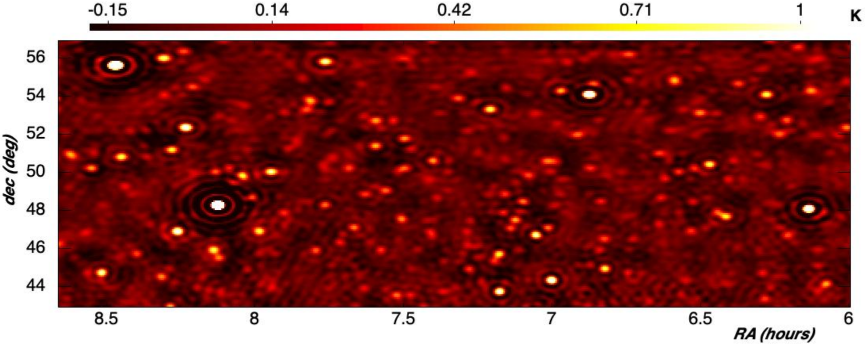
Power spectra (NCP, 1300-1400 MHz averaged) : filtering



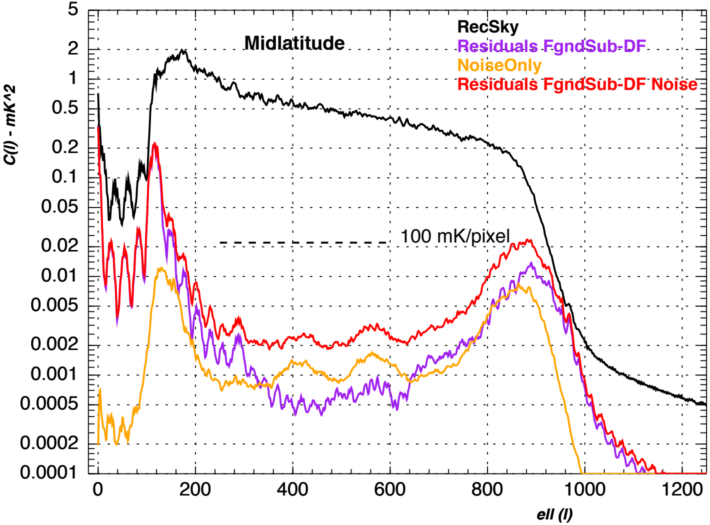
Power spectra (NCP, 1300-1400 MHz averaged) : foreground subtraction



Reconstructed maps (1350 MHz, mid-latitude)



Power spectra (Mid-latitude, 1350 MHz)



Conclusions

- map reconstruction leaves artefacts (noise + mode mixing)
- NCP case "noise dominated" : reconstruction artefacts lower than noise (thanks to more orientations)
- larger integration time per pixel on NCP ⇒ lower noise per pixel
- midlatitude case covers larger area
- ℓ interval selected by fitering and sky coverage
- other conclusions (not shown) :
  - 'polynom' fg subtraction slightly better
  - phase noise can be damaging

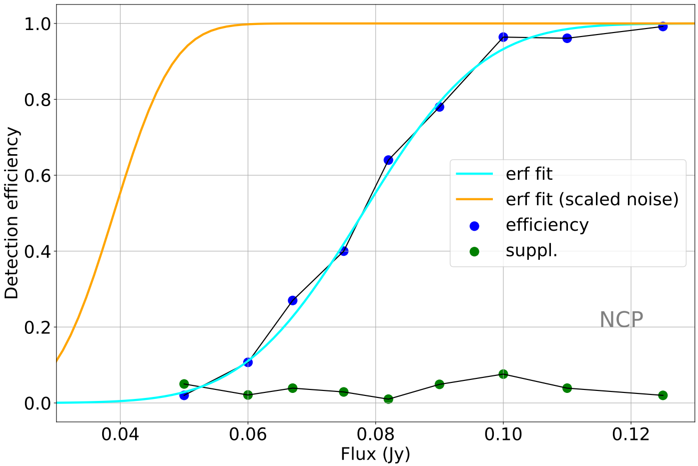
Direct clump detection forecast

- two step process
  1. flux threshold determination with simulation
  2. Use HI mass function (from ALFALFA) to get number of sources vs redshift

Detection efficiency simulation

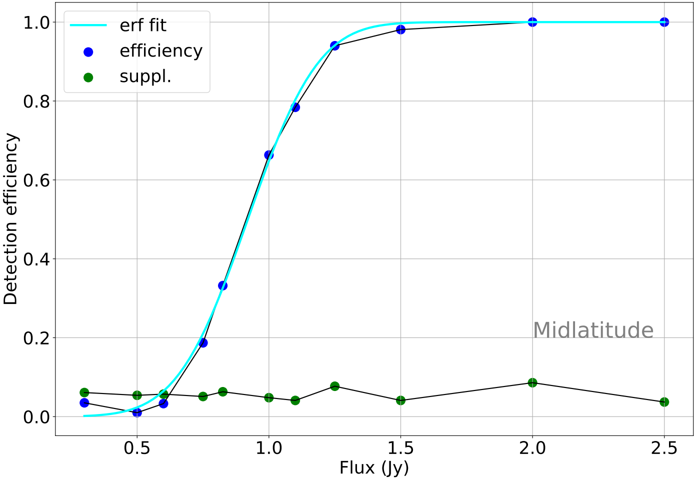
- simulate observations with Tianlai 16 dish setup ; 5mK noise per visibility sample (3mK could be achieved)
- 4 scans (NCP) ; 6 scans (43-57 deg.)
- foreground subtraction : simulate observations at 3 freqs (e.g. 1348, 1350 and 1352 MHz), subtract average of outer frequencies
- add sources for central frequency only
- detection algo : `DAOStarFinder` from `photutils` python package - several parameters, used mainly threshold (unit = image stddev) to avoid spurious "detections"
- compute detection efficiency vs flux (Jy)

Detection efficiency towards NCP



efficiency well described by erf function ; lower threshold with 3mK noise/sample (orange)

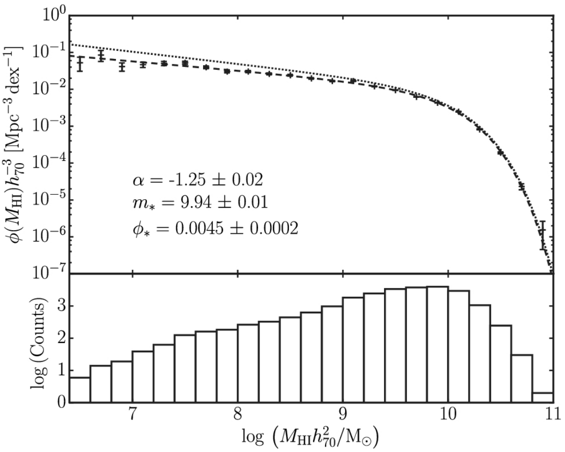
Detection efficiency at midlatitude



Number counts

- HI clumps mass function measured by ALFALFA (arXiv :1802.00053) parametrized as Schechter law (parameters seem to vary with position, distance)
- same reference : $\frac{M_{HI}}{M_\odot} = 2.356x10^5 D^2_{Mpc} S_{21}$
  S_{21} : integrated HI flux (used $\sim$ 200 km/s velocity dispersion here to get peak flux)
- Flux (Jy) $\Rightarrow M_{HI}(z)$
- integration of the MF above threshold / with efficiency at each $z \mapsto$ n(z)
- sum over $z \mapsto$ expected numbers of detections per unit area

Figure 2. The ALFALFA 100 per cent HIMF. The lower panel shows the number counts in logarithmic H i mass bins, and the ...

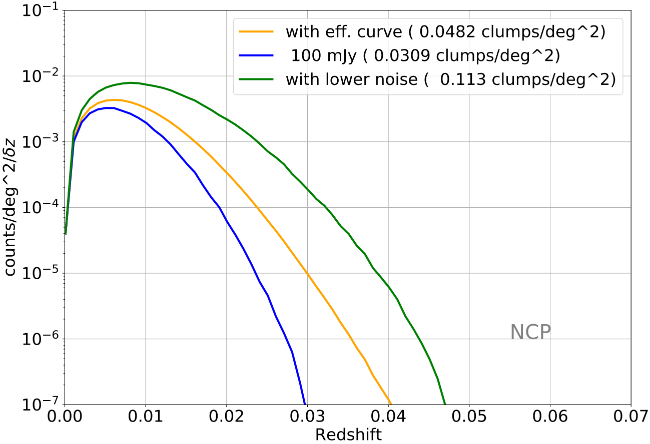


Mon Not R Astron Soc, Volume 477, Issue 1, June 2018, Pages 2–17, https://doi.org/10.1093/mnras/sty521

The content of this slide may be subject to copyright: please see the slide notes for details.

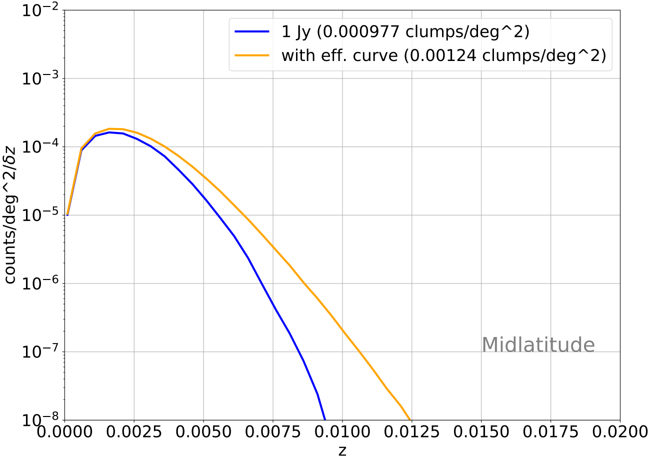
OXFORD UNIVERSITY PRESS

Clumps towards NCP



With $\sim 100deg^2$: ~ 5 HI clumps at $z \lesssim 0.02$ (~ 12 HI clumps at $z \lesssim 0.03$ for lower noise)

Clumps - midlatitude case



With $\sim 2400 deg^2$: $\sim 1 - 2$ HI clumps at $z \lesssim 0.005$

Caveats

- noise is the dominant factor towards NCP : number of clumps would be 2 to 3 times higher with 3mK ($\sim$ no impact for mid-lat case)
- HMF parameter variations seen by ALFALFA (near case less favorable ?)
- test different observation strategy (e.g. cover larger area towards NCP) ?

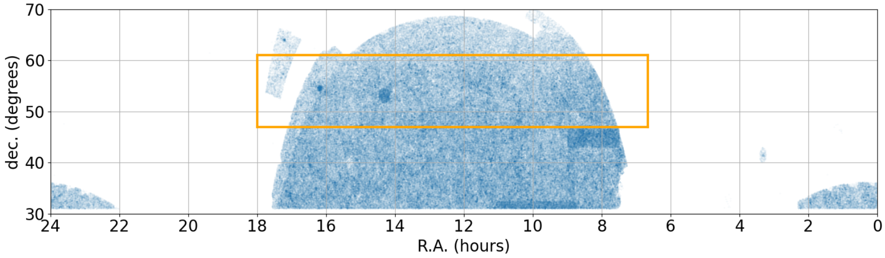
Cross-correlation with optical surveys

- optical surveys :
  - midlatitude : SDSS
  - NCP : ideal case with rotated SDSS
    more realistic : estimate incompleteness of NCCS (on going observations with WYIN) wrt SDSS
- estimate HI properties from optical photometry
- cross-correlation significance estimated at maps' power spectra level
  - midlatitude : C_ℓ cross-spectra
  - NCP : smaller area ⇒ flat approximation

SDSS catalogue

- online queries to 'photo', 'spectro' and 'photo-z' SDSS catalogues
- $\sim$ 900000 objects with $\delta > 31\,\mathrm{deg}$
- basic information : ID, ra,dec, u,g,r,i,z, type, field, specid, class,redshift,error (can be complemented with other catalogues)
- empiric selection (to avoid "crazy" HI parameters) :
  - galaxy (type=3)
  - $.005 < \textit{redshift} < 1$ (avoids e.g. very nearby objects)
  - $-0.5 < r - i < 2.5$ ("standard galaxies ?")
  - $r - i < .8$ or $\textit{redshift} > 0.1$ ("standard galaxies ?")
- $\sim$ 420000 objects

Footprint



The path to HI flux and parameters

1. get the stellar mass from photometry
2. infer the HI mass from stellar mass & photometry
3. compute the HI integrated flux from HI mass
4. get the velocity dispersion from HI mass ($\Rightarrow$ frequency width)
5. pass HI flux, frequency (center+width) params list to simulation

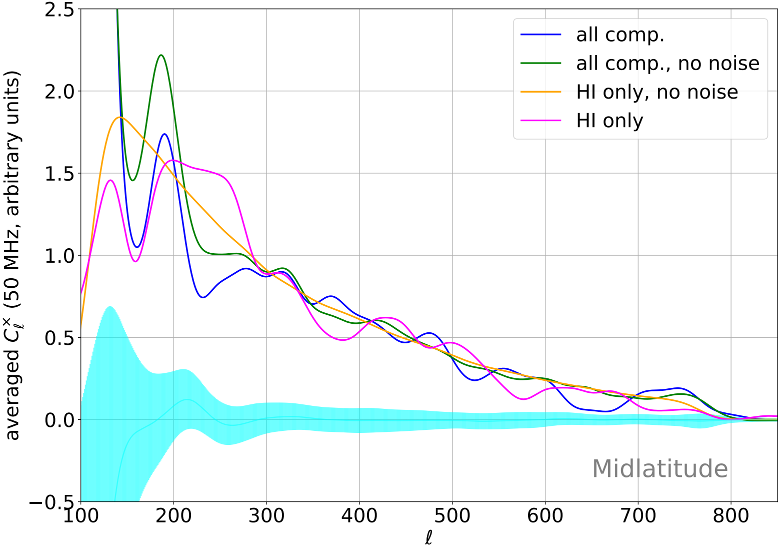
Simulations

- pipelined JSkyMap modules : `map2vis` , `vis2map`, `filt2map`
- Tianlai set-up w/o auto-correlations
- without or with 5 mK noise per visi sample
- maps with all components, HI sources only, foregrounds only, noise only
- planes between 1250 and 1400 MHZ, 1 MHz spacing in 6 'blocks' (data cubes)

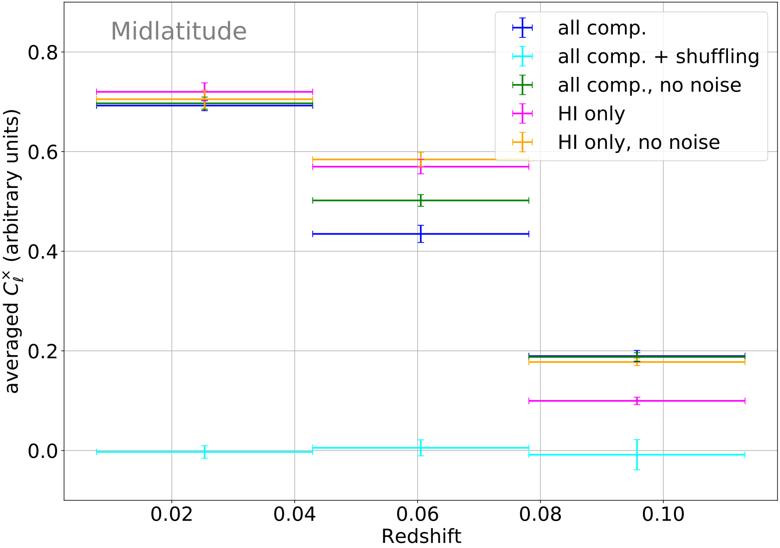
Cross-correlation with input catalog

- build a data cube from input catalog :
  - use only sources' positions (no flux) : normalization of X-correlation arbitrary
  - add a spread in redshift (frequency) according to HI width (gaussian)
  - same frequencies as simulated maps
  - optional shuffling of sources' positions (for cross-checks, error estimation)
- compute the (cross) C_ℓ (harmonic space) of each plane (planes pair) and average them with `spec_dfreq_ecp` (JSkyMap package)
- foreground subtraction with polynomial fit (or running difference)

Midlatitude : averaged spectra 1360-1410 MHz

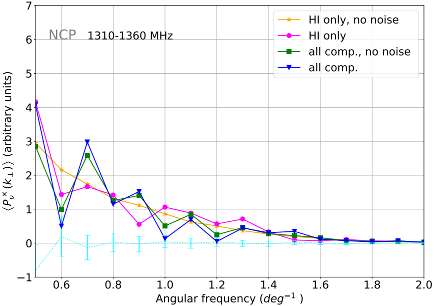


Midlatitude : summary

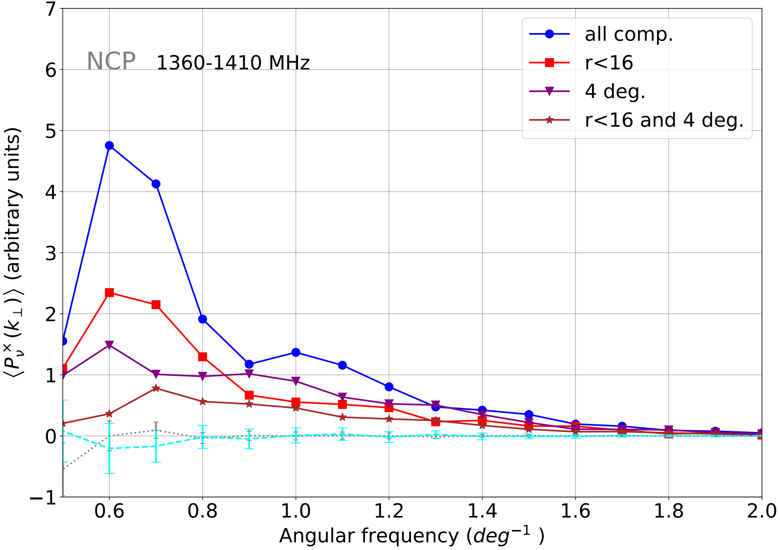


NCP spectra (rotated SDSS)

Work with flat sky projection approximation ⇒ Fourier spectra



NCP spectra : incompleteness effect

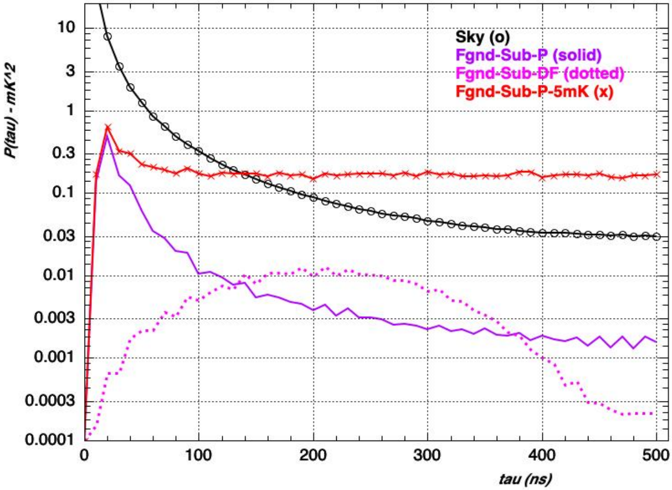


Outlook

- preparation for Tianlai dish array low-z observations
- 2 foreseen surveys : NCP (7 deg diameter) and midlatitude band, 3-5 mK per pixel noise (2-6 months)
- HI clumps/galaxies direct detection : hard at midlatitude, easier towards NCP (O(10) objects)
- cross-correlation with optical surveys :
  - midlatitude : larger area, very significant signal (several possible methods)
  - NCP : smaller area, on-going optical spectroscopic survey but signal also significant

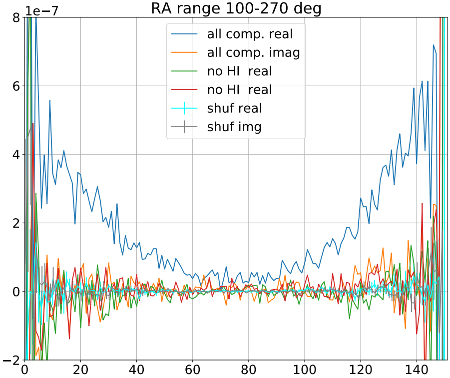
BACK-UPs

Component separation methods



Midlatitude alternative method (I)

- compute visibilities from catalog data cube ⇒ 3D arrays $V_b(RA, desc, \nu)$ for each baseline B
- 1D FFT in ν axis for each baseline and coords
- average cross-power spectra over baselines and coords
- NB : no explicit component separation



Midlatitude alternative method (II)

- build maps from sources' cube visibilities from method (I) (same params & filtering as in published paper)
- option : subtract average of neighbouring freq. maps
- cross-FFTs in ν axis for each pixel + average