New synthetic data products for the SKA precursors

Ed Elson
SKA postdoctoral fellow
drelson.e.c@gmail.com

Astrophysics, Cosmology and Gravity Centre
University of Cape Town

PHISCC 2016 - Cape Town

Motivation

- Synthetic data products for next-gen HI galaxy surveys will be useful in many respects:
  - Source-finding algorithms
  - Visualisation methods
  - Source completeness checks
  - Quantify cosmic variance effects

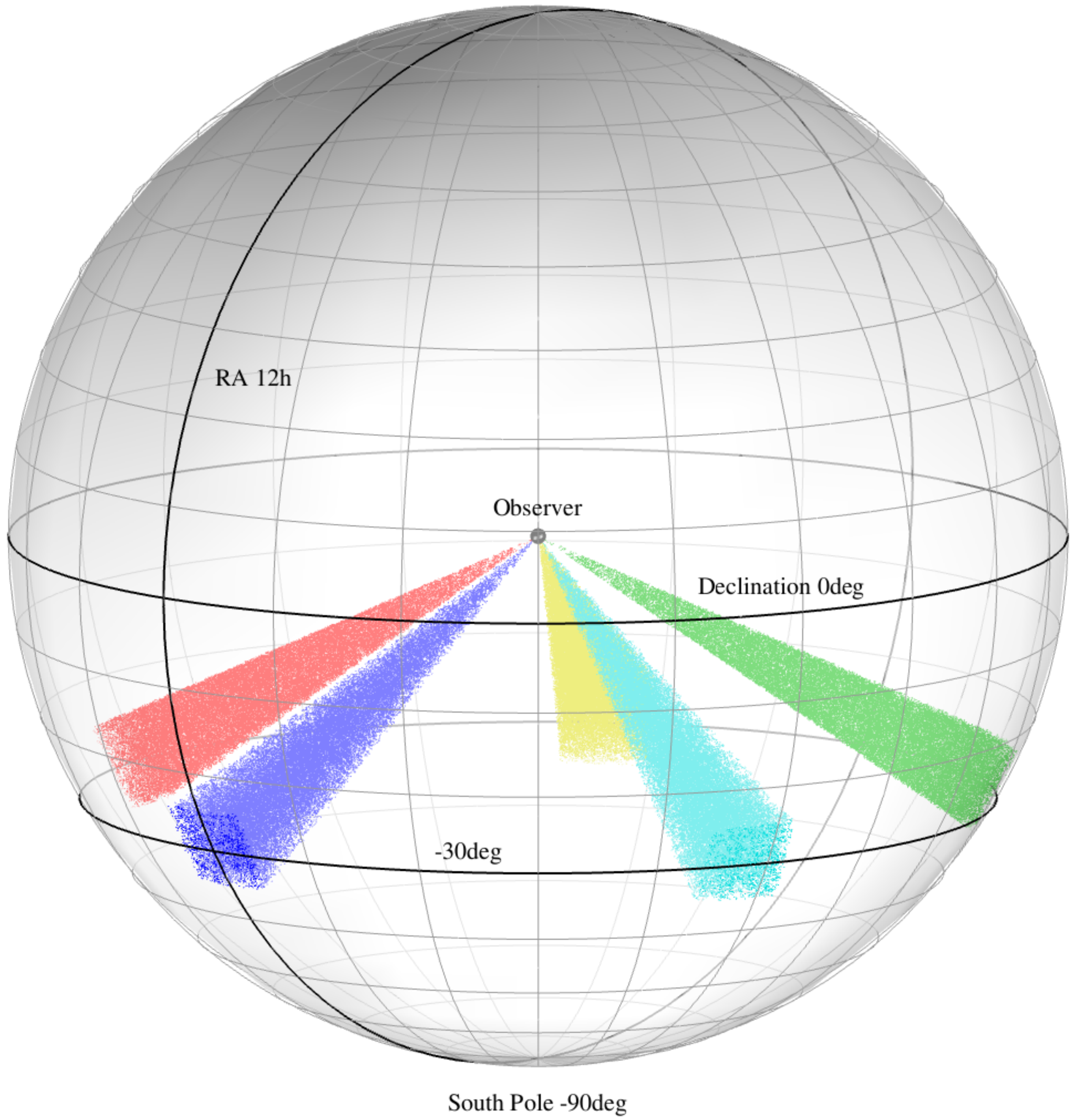
Simulations

- Pipeline converts a catalogue of galaxy parameters into a realistic synthetic HI data cube.
- Obreschkow & Meyer (2014) catalogues:
  - Based on S^3-SAX (Obreschkow+, 2009)
  - Physical models tracing evolution of HI and H_2 in +/- 3e7 galaxies.
  - Used as basis for SKA performance calculations.

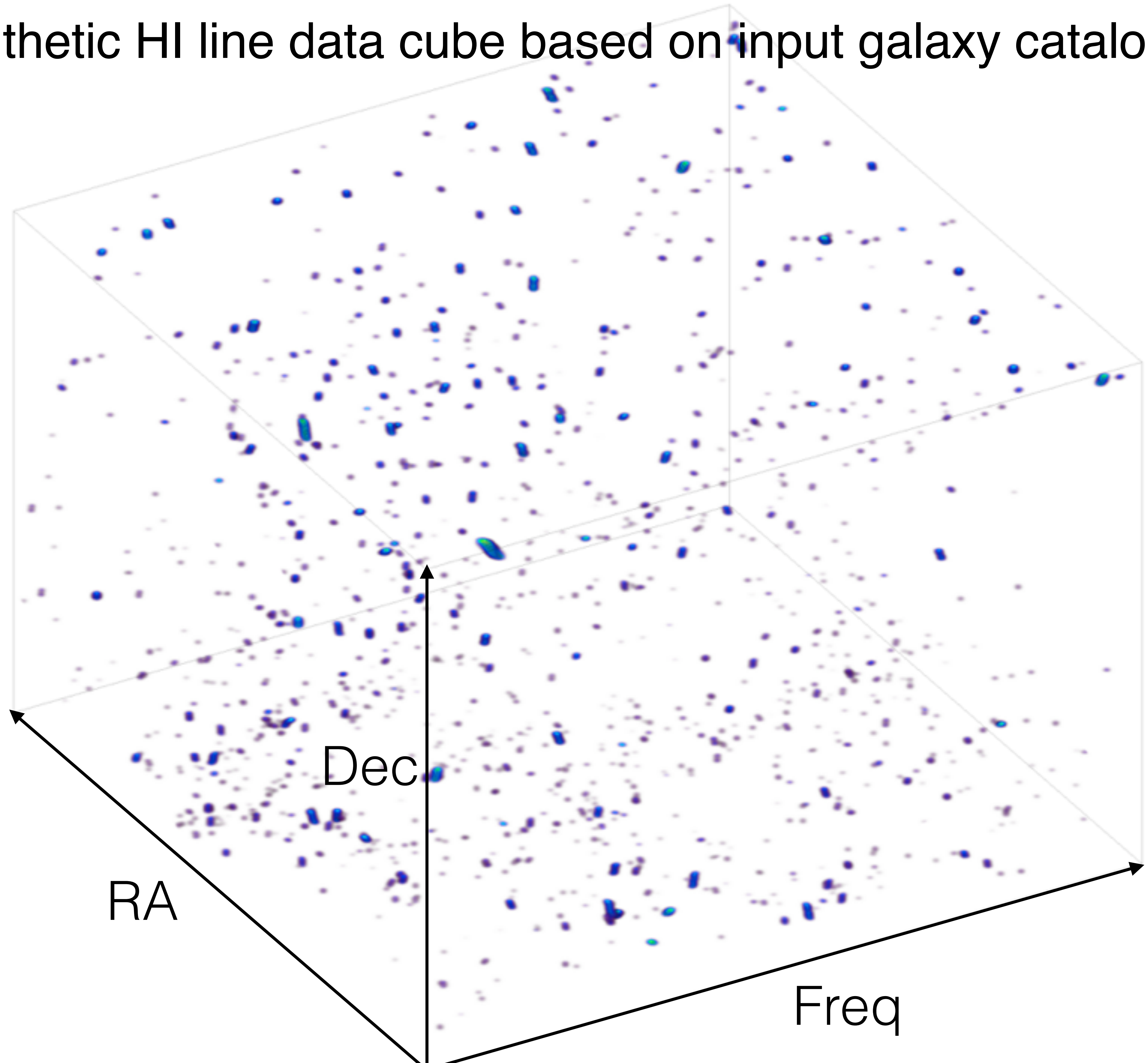
Obreschkow & Meyer (2014) sims:

Symbol	Unit	Description
ID	–	Unique galaxy identifier in the Munich Semi-Analytic Model "DeLucia2006a"
RA	deg	Right ascension of galaxy centre
Dec	deg	Declination of galaxy centre
z	–	Apparent redshift of galaxy centre, including the Doppler component due to peculiar motion relative to the Hubble expansion
i	deg	Galaxy inclination defined as the smaller angle $(0° - 90°)$ between the line-of-sight and the rotational axis of the galaxy
T	–	Numerical Hubble type (−6...0 for ellipticals, 0...10 for spirals, 99 for morphologically unresolved objects, mostly dwarfs)
M_*	$M_\odot$	Stellar mass
$M_{\rm H\,I}$	$M_\odot$	Mass of neutral atomic hydrogen H I, without helium
$M_{\rm H_2}$	$M_\odot$	Mass of molecular hydrogen H_2, without helium
$S^{\rm int}_{\rm H\,I}$	$\rm Jy\,km\,s^{-1}$	Velocity-integrated flux of the redshifted 21 cm H I emission line, with velocity units defined in the galaxy rest-frame
$S^{\rm peak}_{\rm H\,I}$	Jy	Peak flux density of the H I emission line; typically the flux density of the 'horns'
$S^{\rm int}_{\rm CO}$	$\rm Jy\,km\,s^{-1}$	Velocity-integrated flux of the redshifted 115.27 GHz ^{12}CO(1–0) emission line, with velocity units defined in the galaxy rest-frame
$S^{\rm peak}_{\rm CO}$	Jy	Peak flux density of the ^{12}CO(1–0) emission line; typically the flux density of the 'horns'
$W^{50}_{\rm H\,I}$	$\rm km\,s^{-1}$	Width of the H I emission line, in galaxy rest-frame velocity units, measured at 50% of the peak flux density
$W^{20}_{\rm H\,I}$	$\rm km\,s^{-1}$	Width of the H I emission line, in galaxy rest-frame velocity units, measured at 20% of the peak flux density
$r^{\rm edge}_{\rm H\,I}$	arcsec	Apparent H I radius along the major axis out to a H I disk surface density of $1\,M_\odot pc^{-2}$, corresponding to a face-on column density of $1.25 \cdot 10^{20} cm^{-2}$
$r^{\rm half}_{\rm H\,I}$	arcsec	Apparent H I half-mass radius along the major axis
$M_{\rm R}$	mag	Absolute Vega R-band magnitude, corrected for intrinsic dust extinction; 99 if stellar mass and star formation history are insufficiently resolved to compute $M_{\rm R}$
$m_{\rm R}$	mag	Apparent Vega R-band magnitude; value 99 if no absolute magnitudes available
$r_{\rm e}$	arcsec	Effective radius, here approximated as the radius containing half the stellar mass if the galaxy were viewed face-on

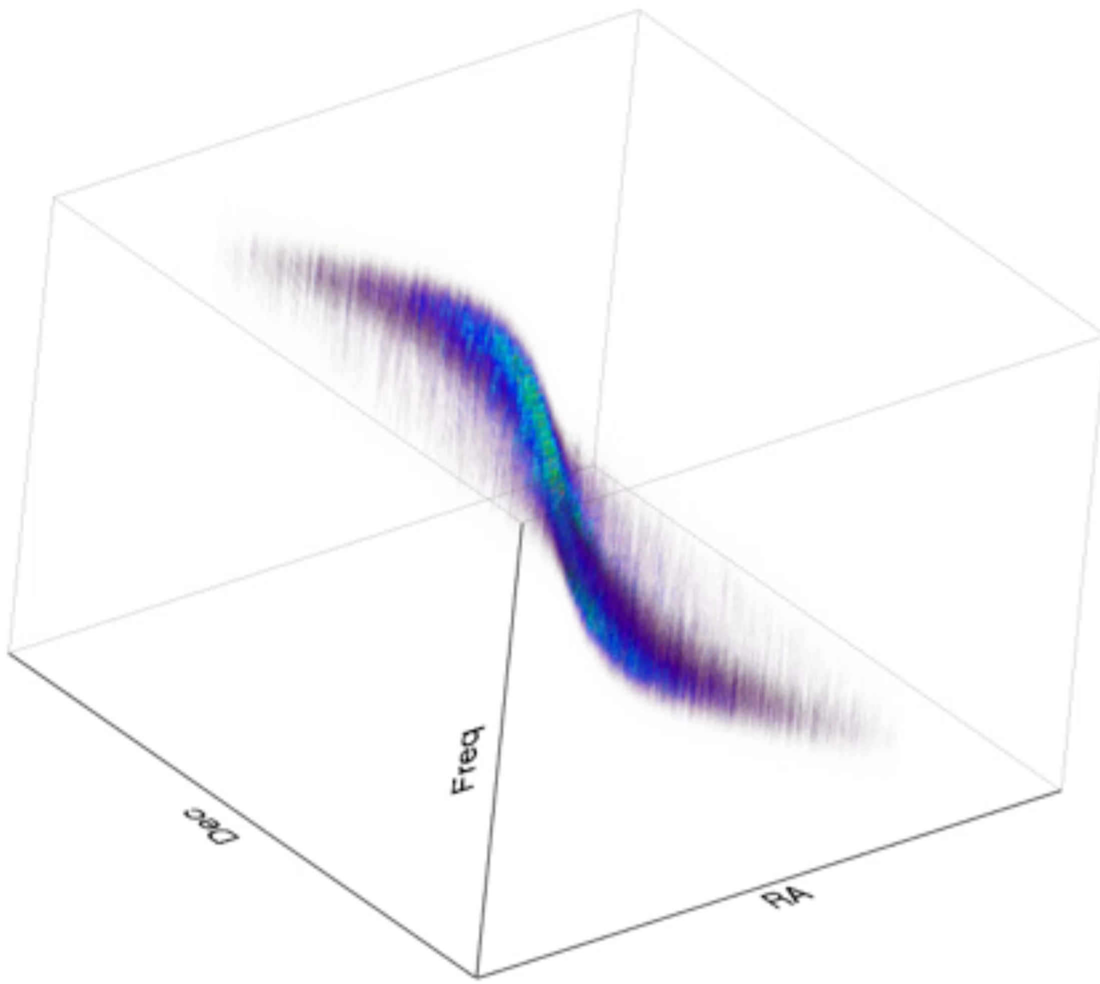
Table 1: Description of the columns of mock catalog in ASCII format.



Synthetic HI line data cube based on input galaxy catalogue:

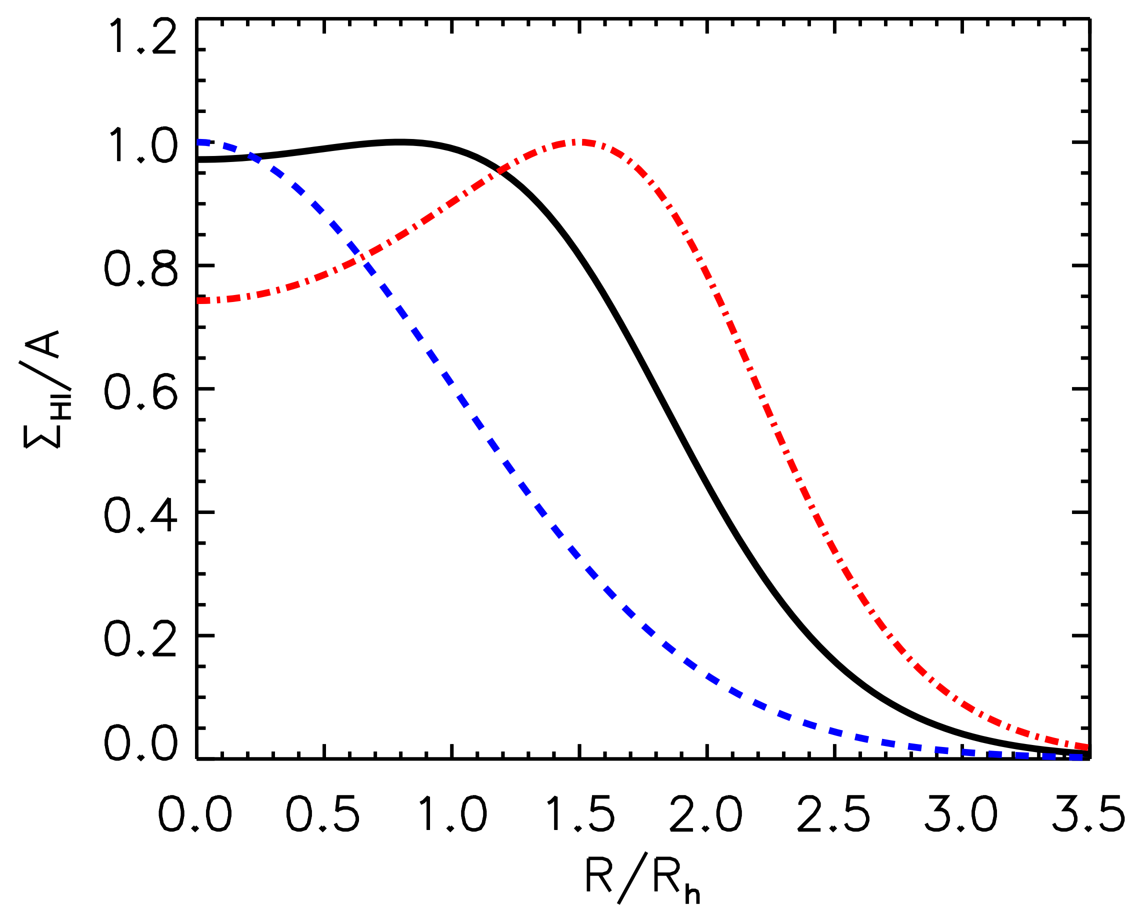


Modelling the HI line emission

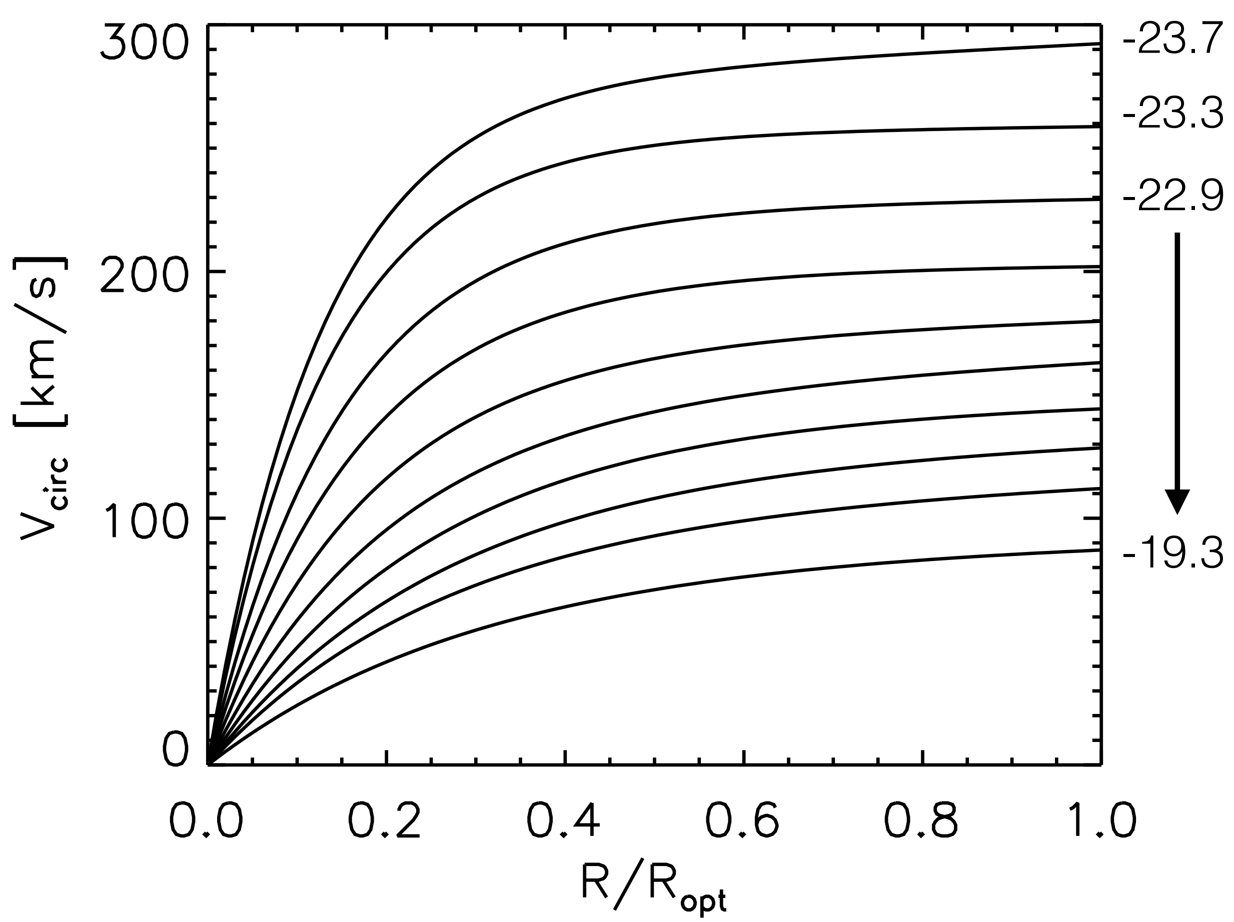


- 3D model generated for each galaxy in galaxy catalogue.
- Unique HI mass distribution, rotation curve, incl, PA, etc.
- Fully automated using custom scripts.

M_{HI} & V_{circ} profiles



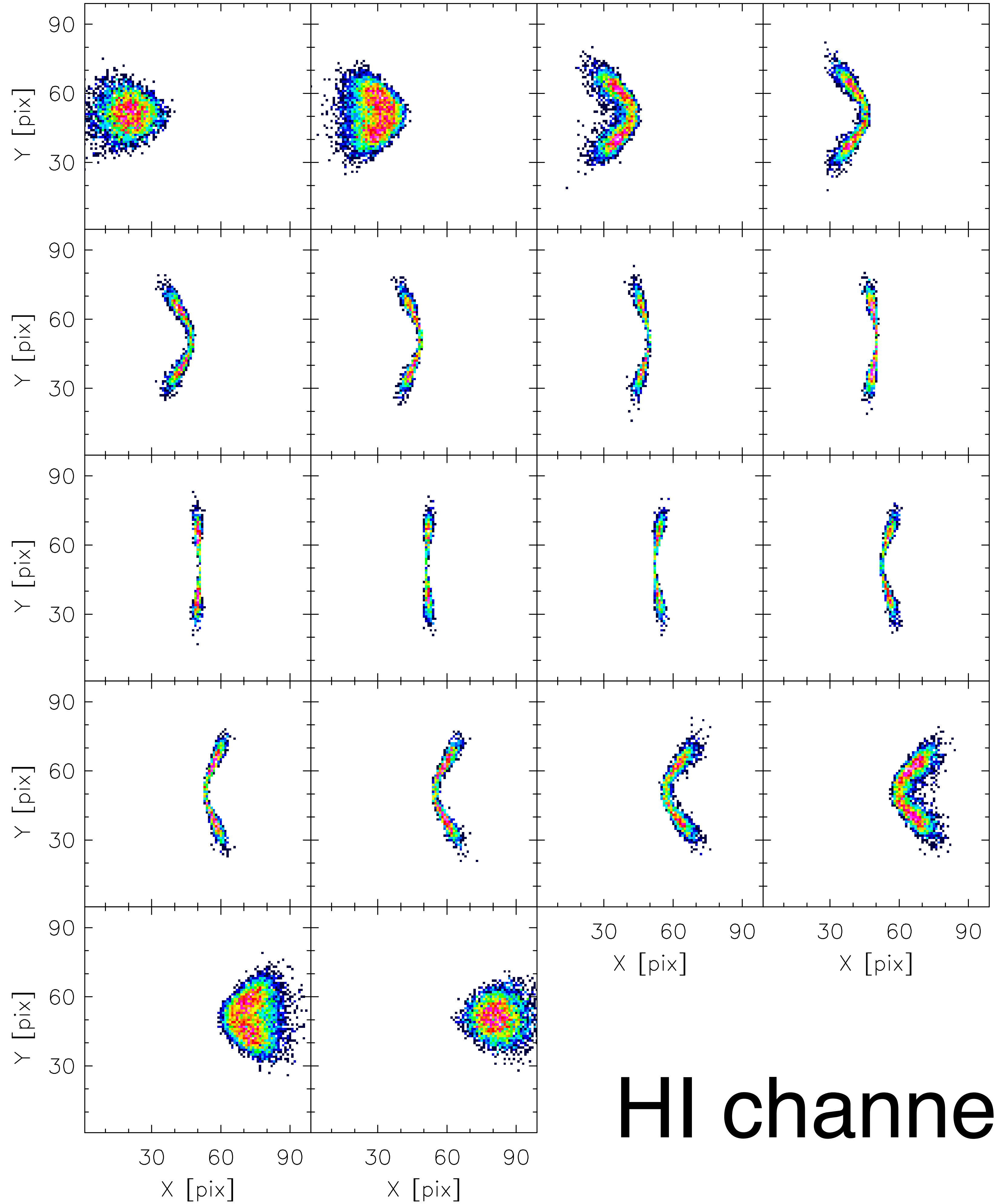
Template rotation curves: Catinella + (2006)



$$\Sigma_{\rm HI}(R) = \frac{A\exp(-R^2/2h^2)}{1+\beta\exp(-1.6R^2/2h^2)}$$

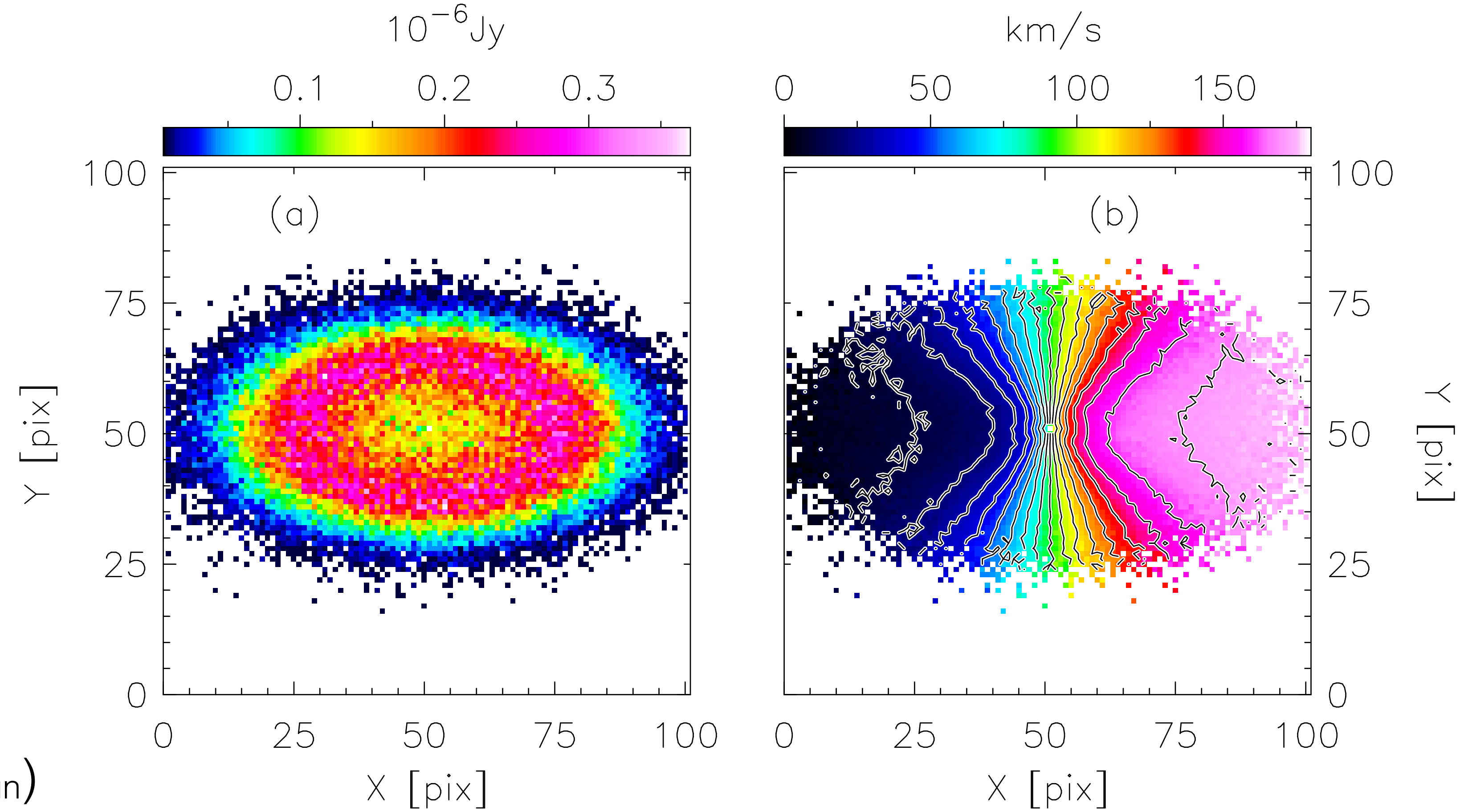
$$V_{\rm PE}(R) = V_0\left(1-e^{-R/R_{\rm PE}}\right)\left(1+\frac{\alpha R}{R_{\rm PE}}\right)$$

Polyex model: Giovanelli & Haynes 2002



HI channel maps

HI data products

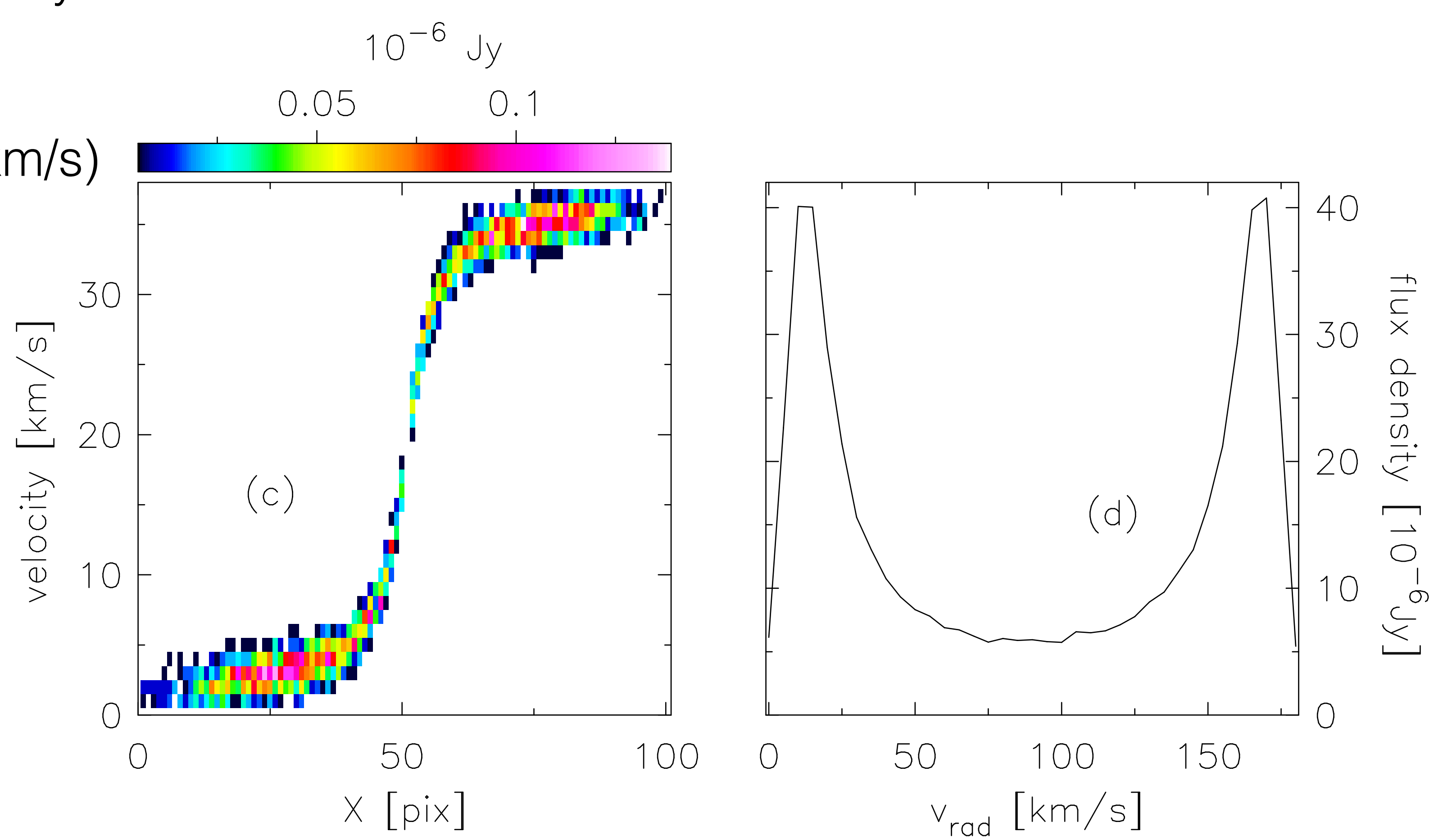


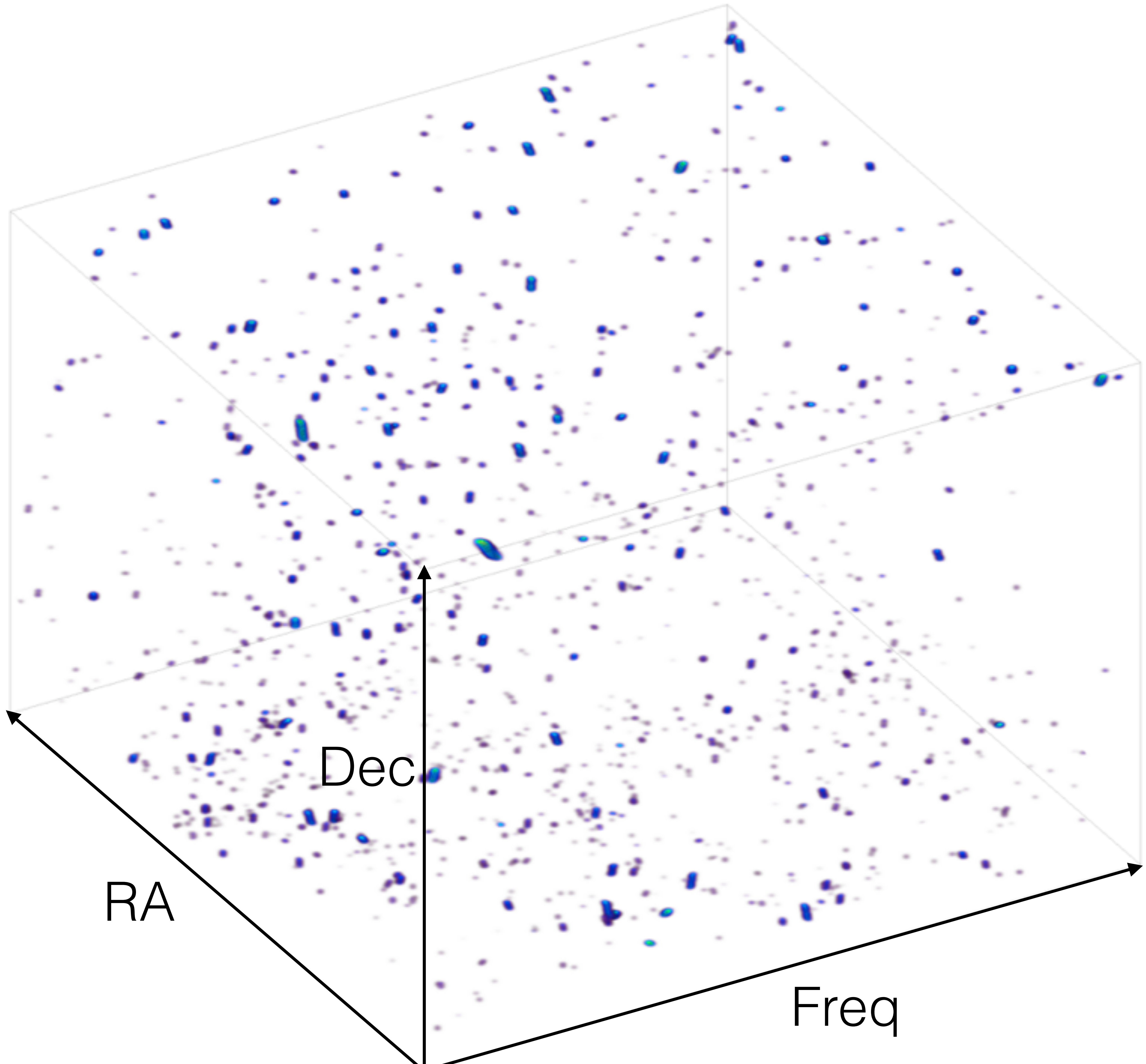
a) HI total intensity ($M_{HI}=10^{8.16}M_{sun}$)

b) HI intensity-weighted mean velocity field

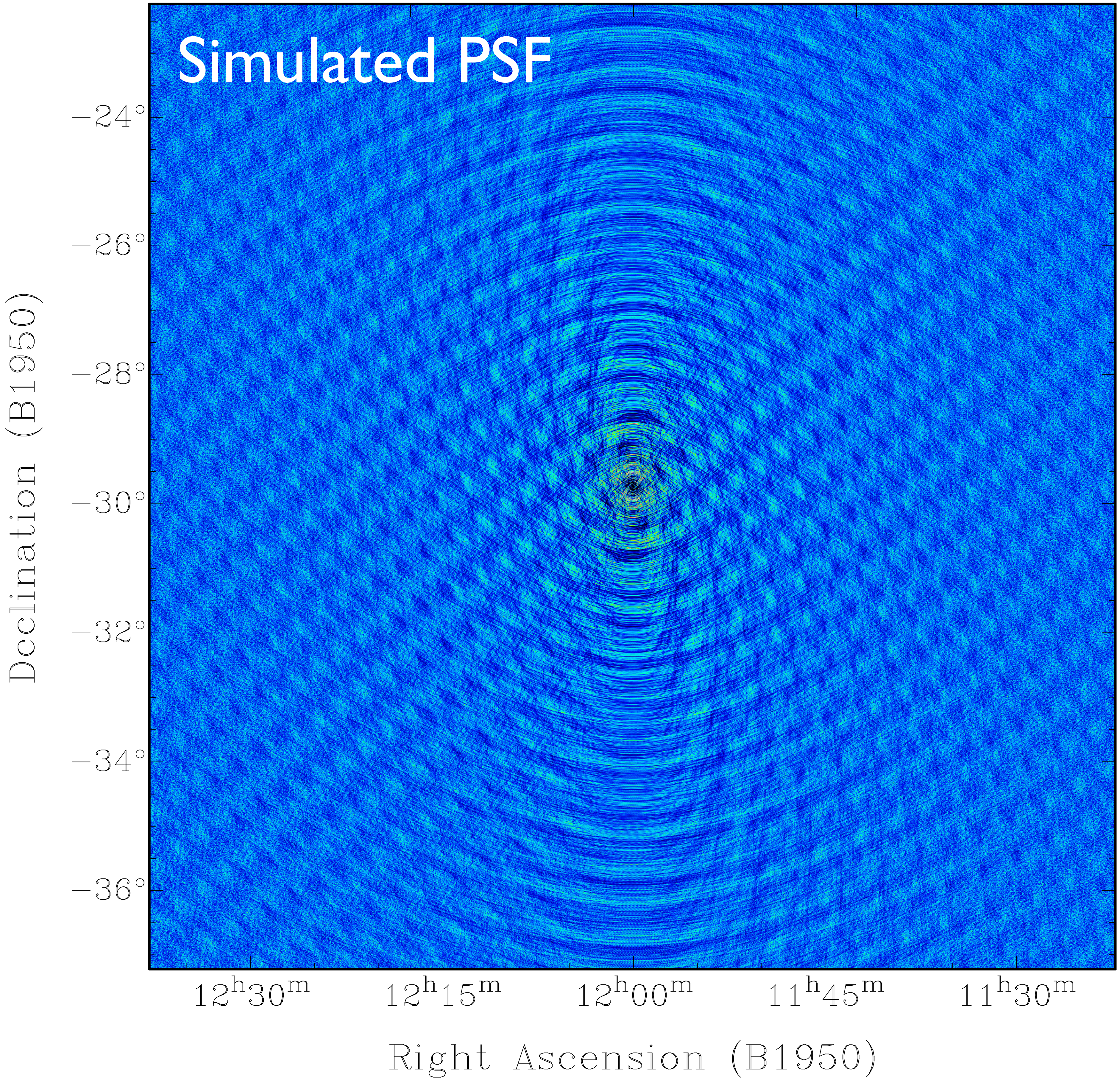
c) Major axis pv slice

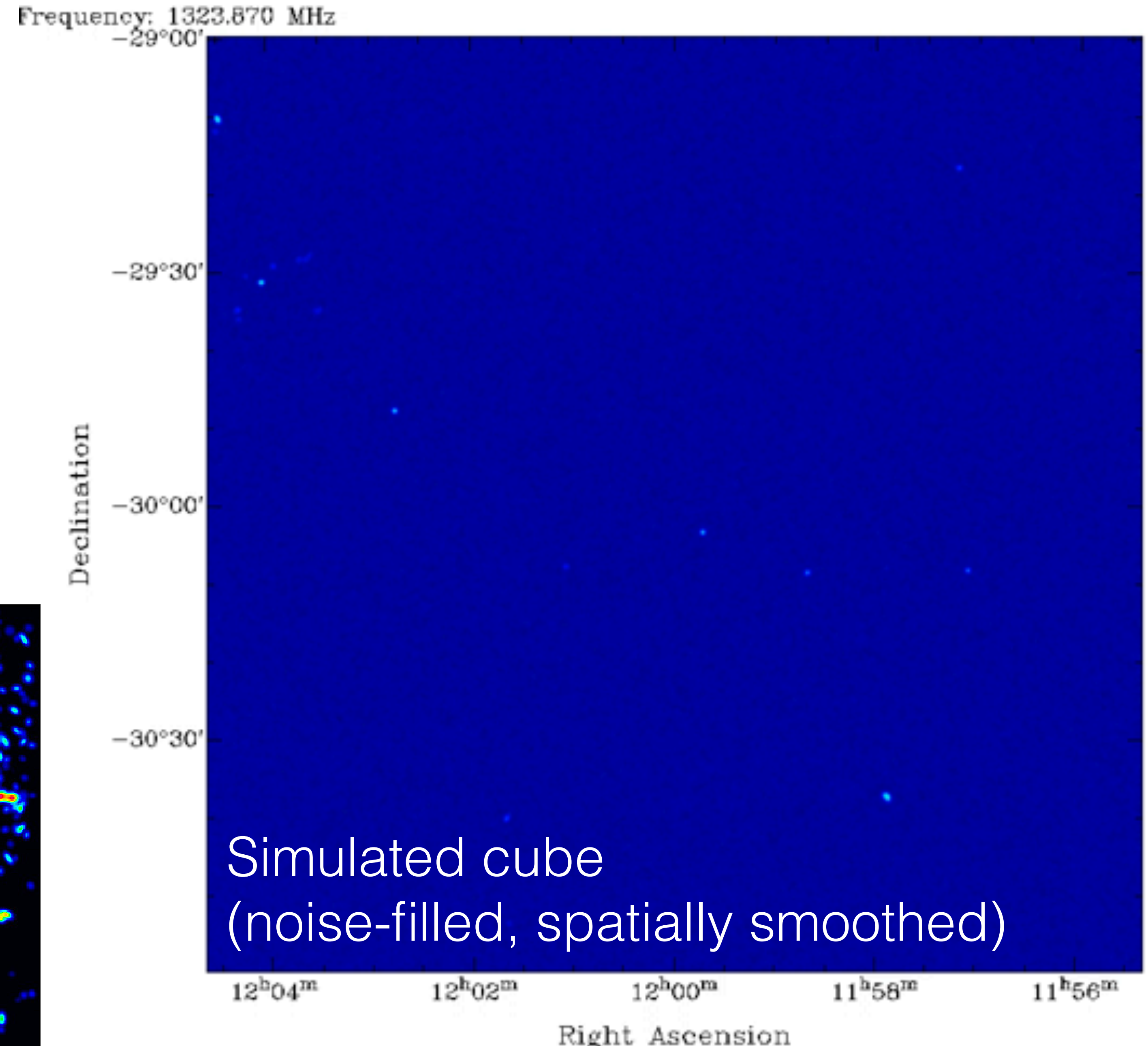
d) HI global profile (S_{int}=2.66 mJy km/s)

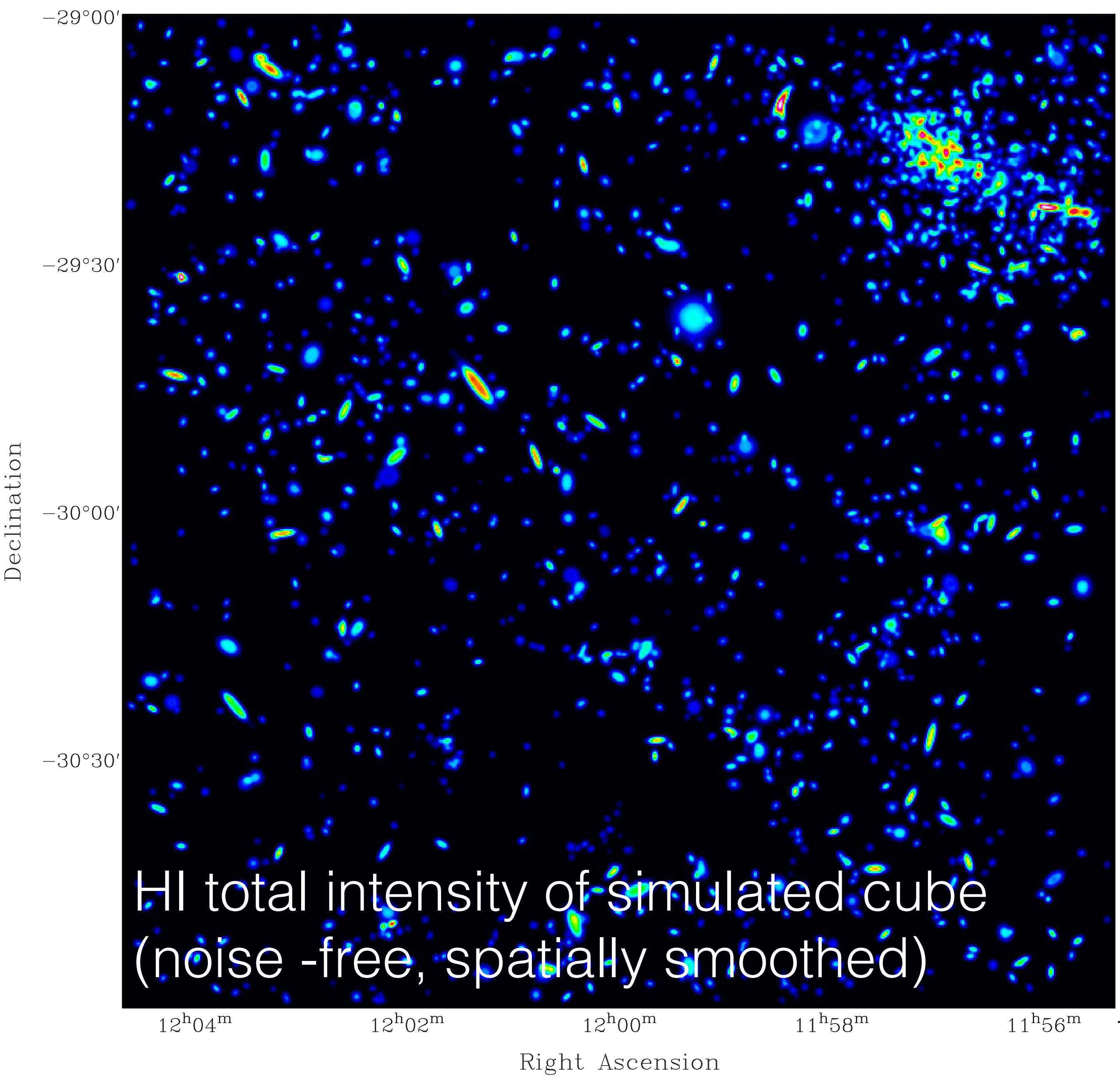




Dec
RA
Freq


Simulated PSF
−24°
−26°
−28°
−30°
−32°
−34°
−36°
Declination (B1950)
12h30m
12h15m
12h00m
11h45m
11h30m
Right Ascension (B1950)


Frequency: 1323.870 MHz
−29°00′
−29°30′
−30°00′
−30°30′
Declination
12h04m
12h02m
12h00m
11h58m
11h56m
Right Ascension
Simulated cube
(noise-filled, spatially smoothed)


−29°00′
−29°30′
−30°00′
−30°30′
Declination
12h04m
12h02m
12h00m
11h58m
11h56m
Right Ascension
HI total intensity of simulated cube
(noise -free, spatially smoothed)

User-specified inputs

Item	Example
RA range	5 deg.
Dec range	5 deg.
z range	0.2 - 0.7
cosmology	$h=0.7$, $\Omega_m=0.3$, $\Omega_{\Lambda}=0.7$
channel width	100 kHz
pixel scale	1 arcmin
BMAJ	4 arcmin
BMIN	3 arcmin
RMS noise	0.1 mJy/bm

\+ optical magnitude cut

\+ S_{int} cut

- HI line cube (un-smoothed)
- HI line cube, (smoothed)
- HI + noise cube
- Catalogue of true galaxy properties
- Means of easily identifying each galaxy in a cube

Application: HI stacking

Unappreciated shortcoming of HI stacking method:

source confusion

Application: HI stacking

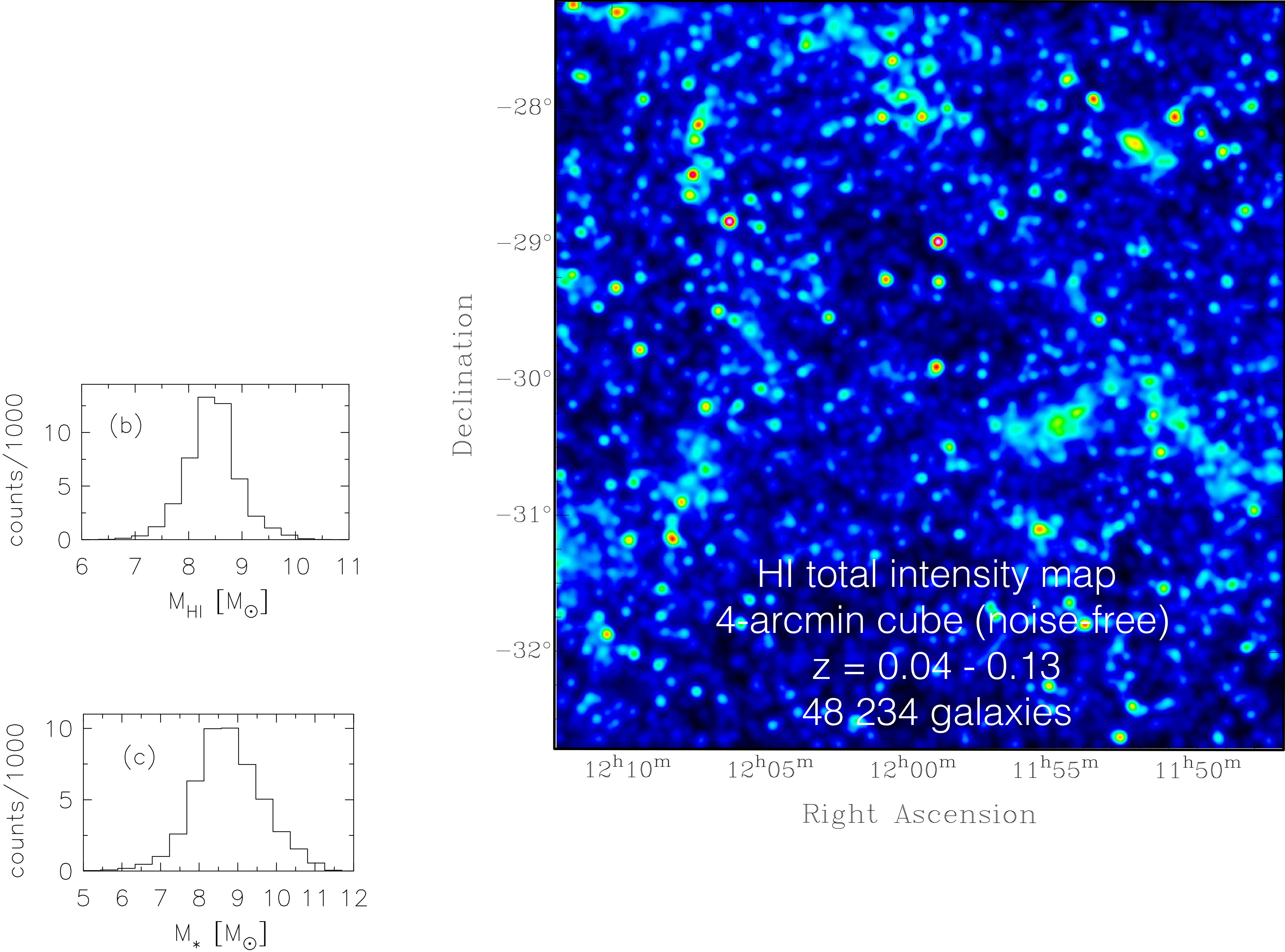
Unappreciated shortcoming of HI stacking method:

source confusion

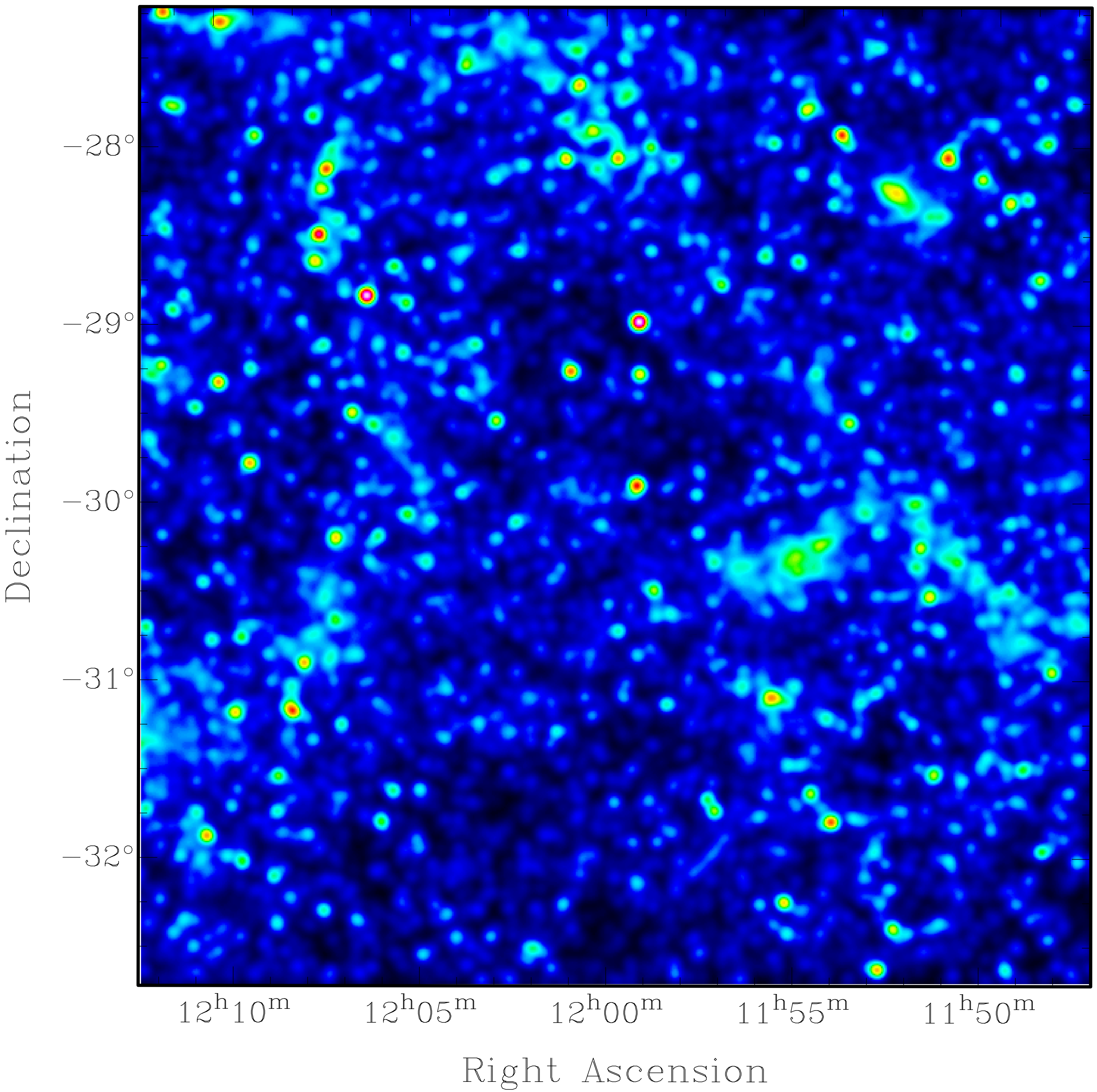
- Galaxy volume density
- Spatial resolution of observations

Application: HI stacking

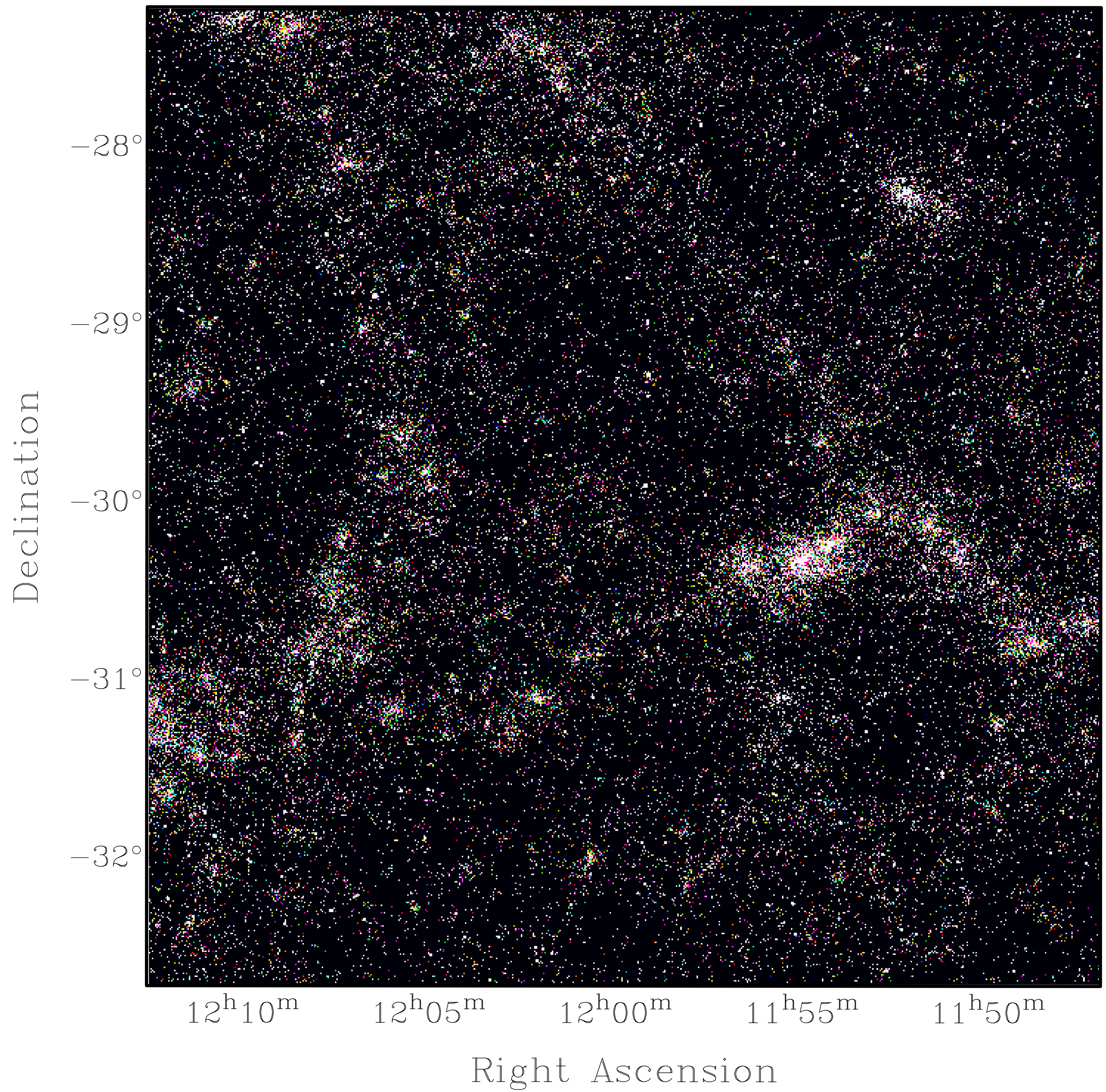
- Generate noise-free cube with characteristics:
  - sky area: 30 sq deg
  - $z = 0.04 - 0.13$
  - 48 234 galaxies ($M_{HI} = 2.81e13\ M_{\odot}$)
  - $\vartheta = [0, 4, 15]$ arcmin.
- Extract spectra at (3583) galaxies with $M_{*} > 10^{10}\ M_{\odot}$
- **How badly are our stacks affected by source confusion?**



(b)
counts/1000
0
5
10
6
7
8
9
10
11
M_HI [M_⊙]
(c)
counts/1000
0
5
10
5
6
7
8
9
10
11
12
M_* [M_⊙]
Declination
−28°
−29°
−30°
−31°
−32°
12h10m
12h05m
12h00m
11h55m
11h50m
Right Ascension
HI total intensity map
4-arcmin cube (noise-free)
z = 0.04 - 0.13
48 234 galaxies

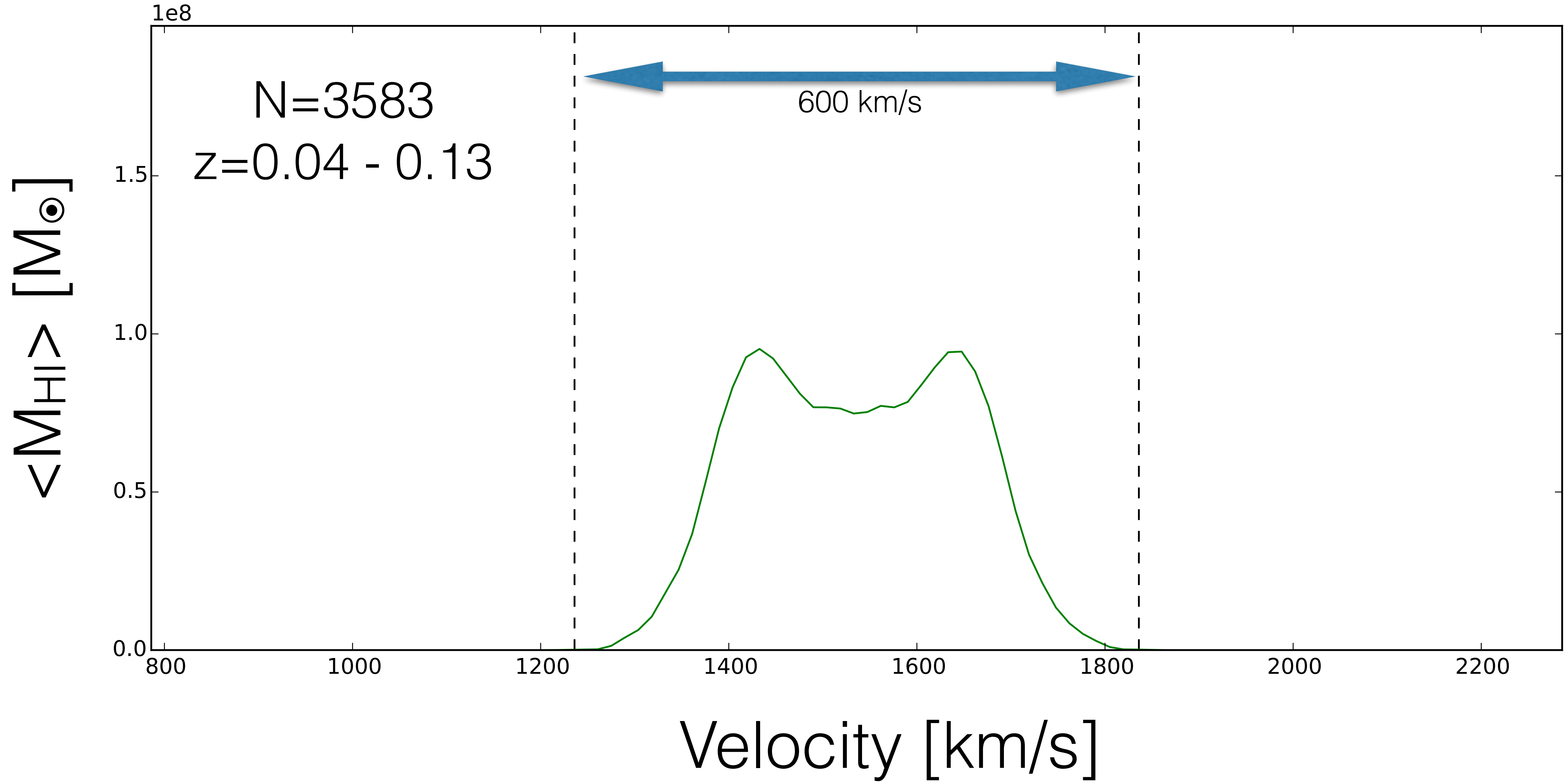


−28°
−29°
−30°
−31°
−32°
Declination
12h10m
12h05m
12h00m
11h55m
11h50m
Right Ascension



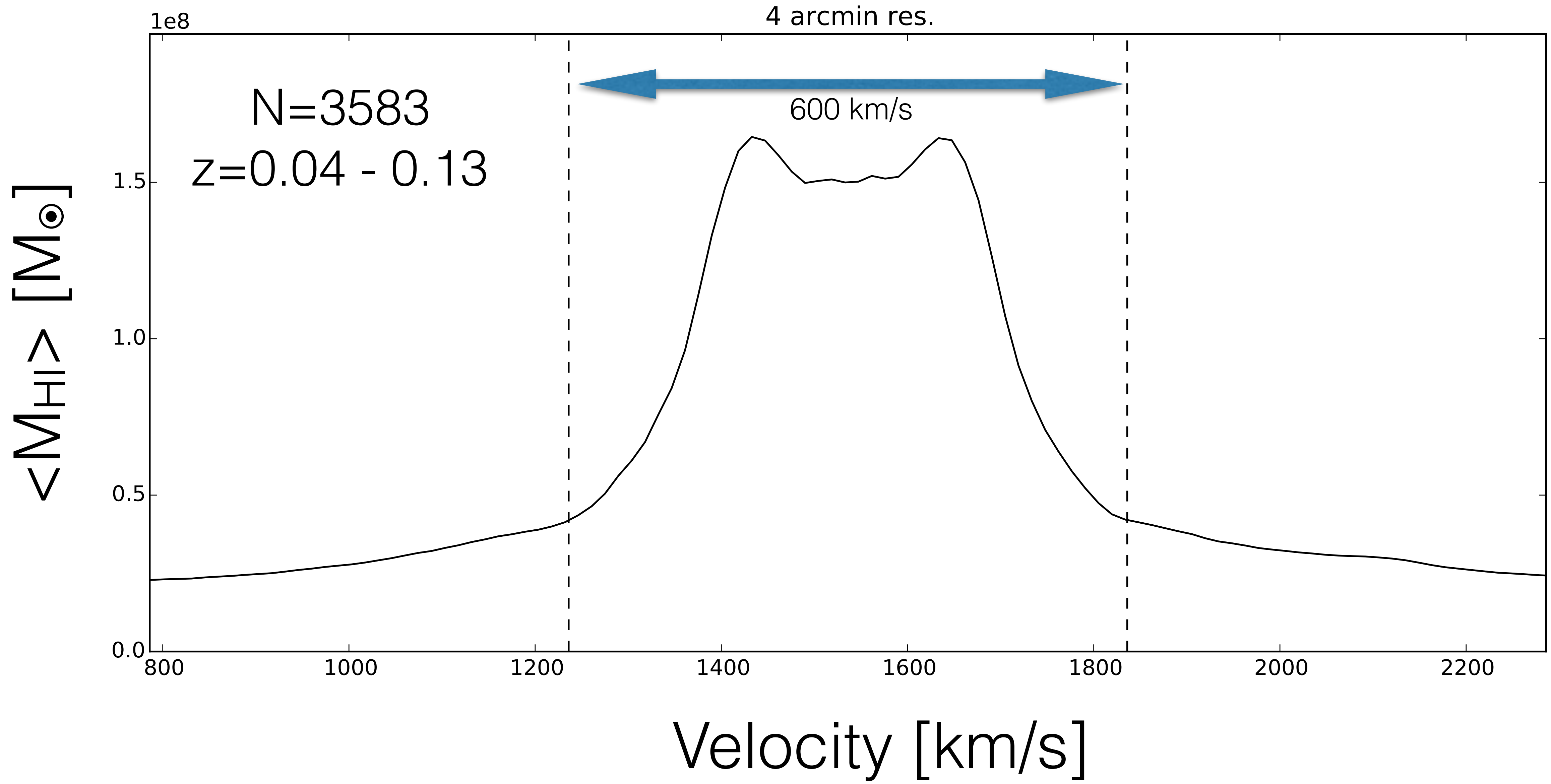
−28°
−29°
−30°
−31°
−32°
Declination
12h10m
12h05m
12h00m
11h55m
11h50m
Right Ascension

co-add: noise-free, un-smoothed



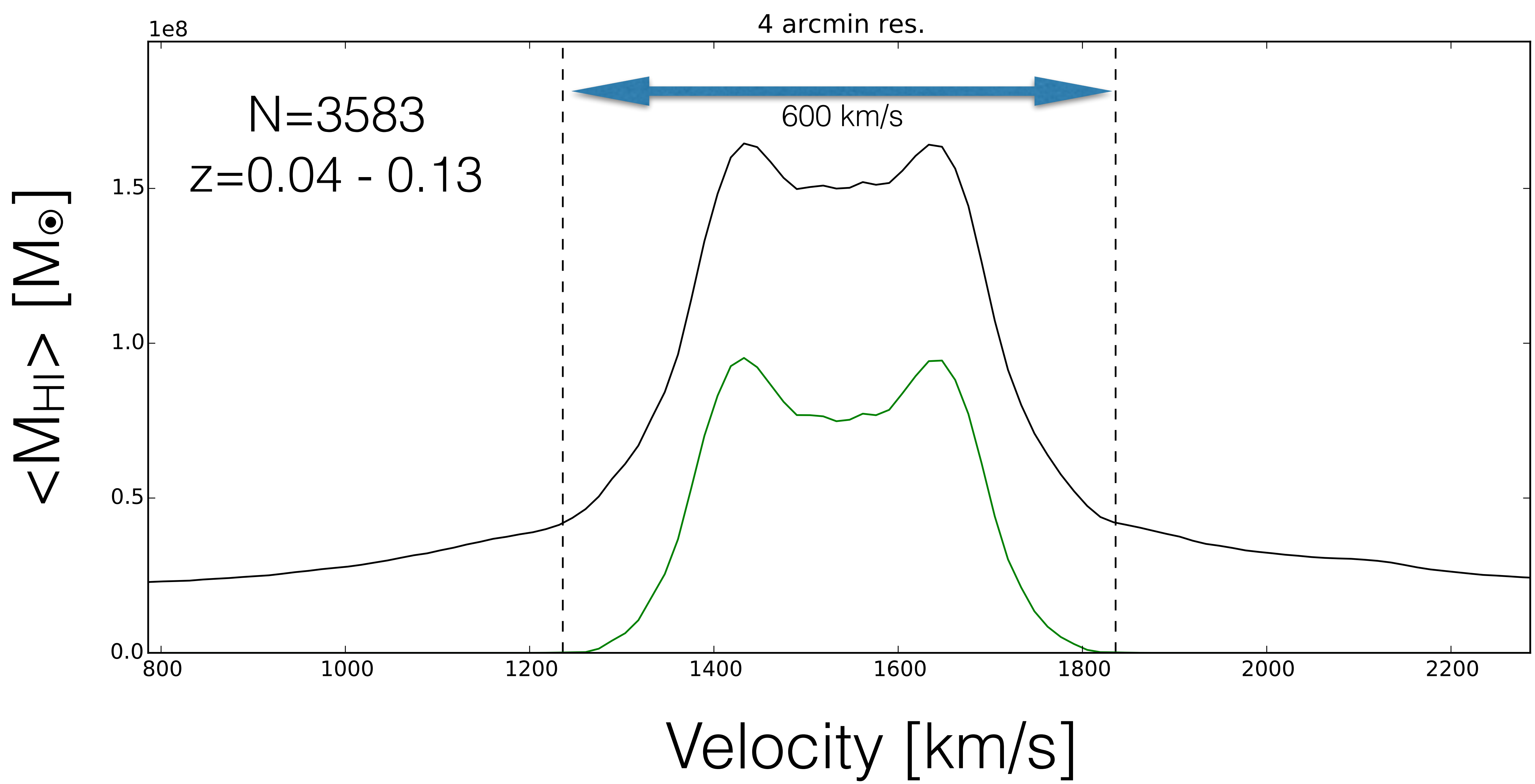
<M$_{HI}$> = 2.08 x 10^9 M$_\odot$ ✓

co-add: noise-free, **4 arcmin**



$\langle M_{HI} \rangle = 8.52 \times 10^9\ M_\odot$ ❌

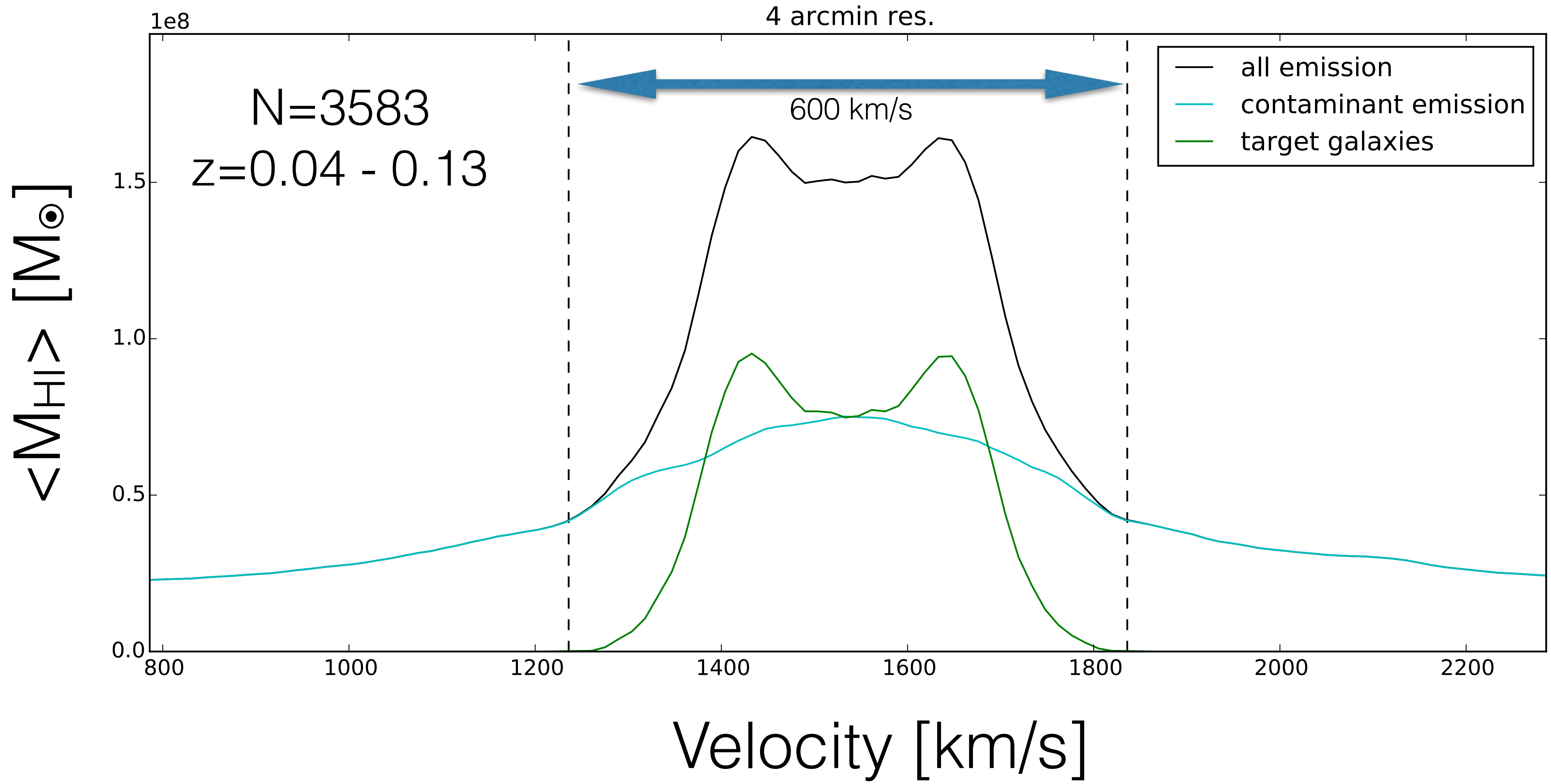
co-add: noise-free, **smoothed**



$\langle M_{HI} \rangle = 8.52 \times 10^9 M_\odot$ ❌

$\langle M_{HI} \rangle = 2.08 \times 10^9 M_\odot$ ✔

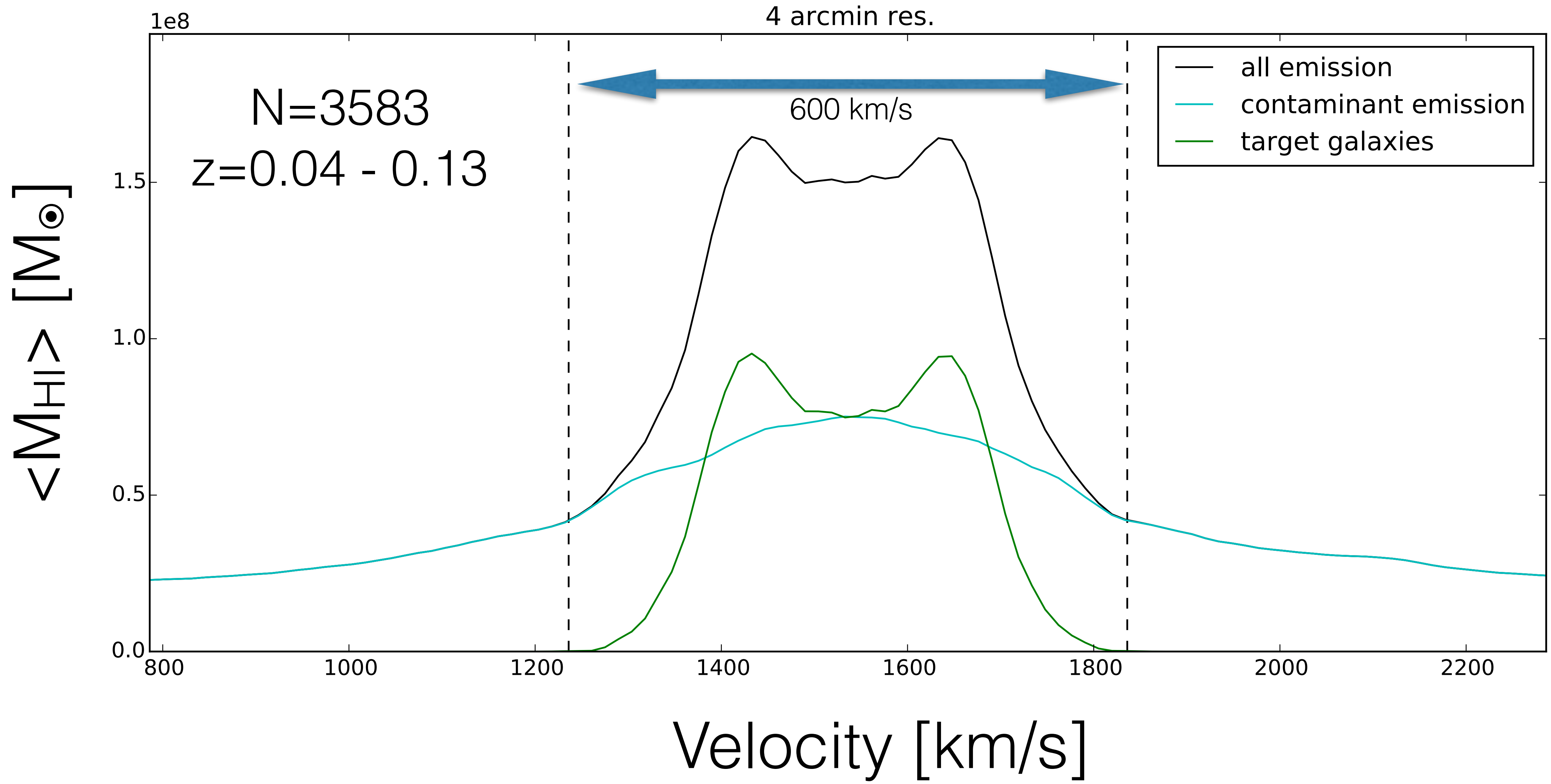
co-add: noise-free, **smoothed**



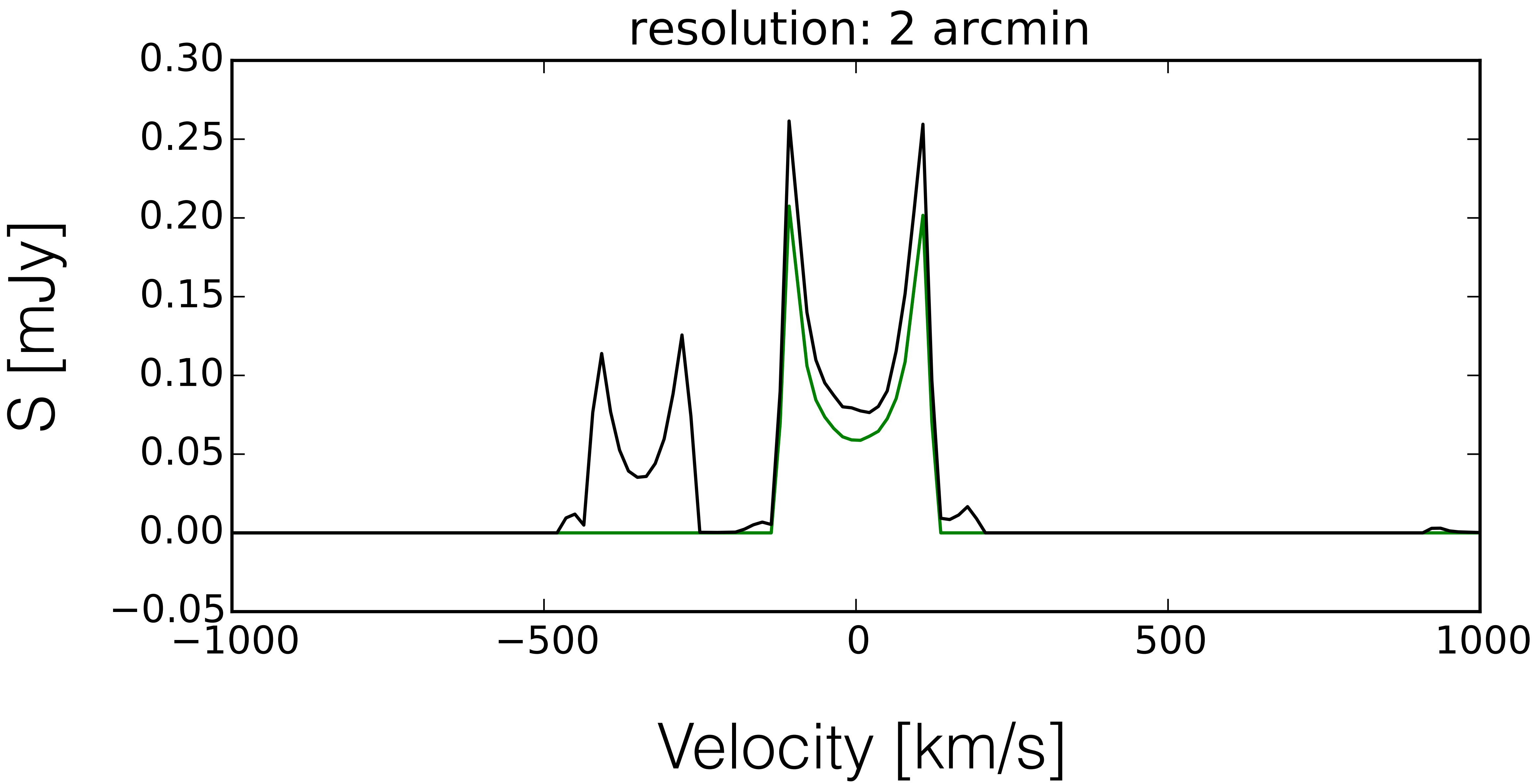
$\langle M_{HI} \rangle = 8.52 \times 10^9\ M_\odot$ ❌

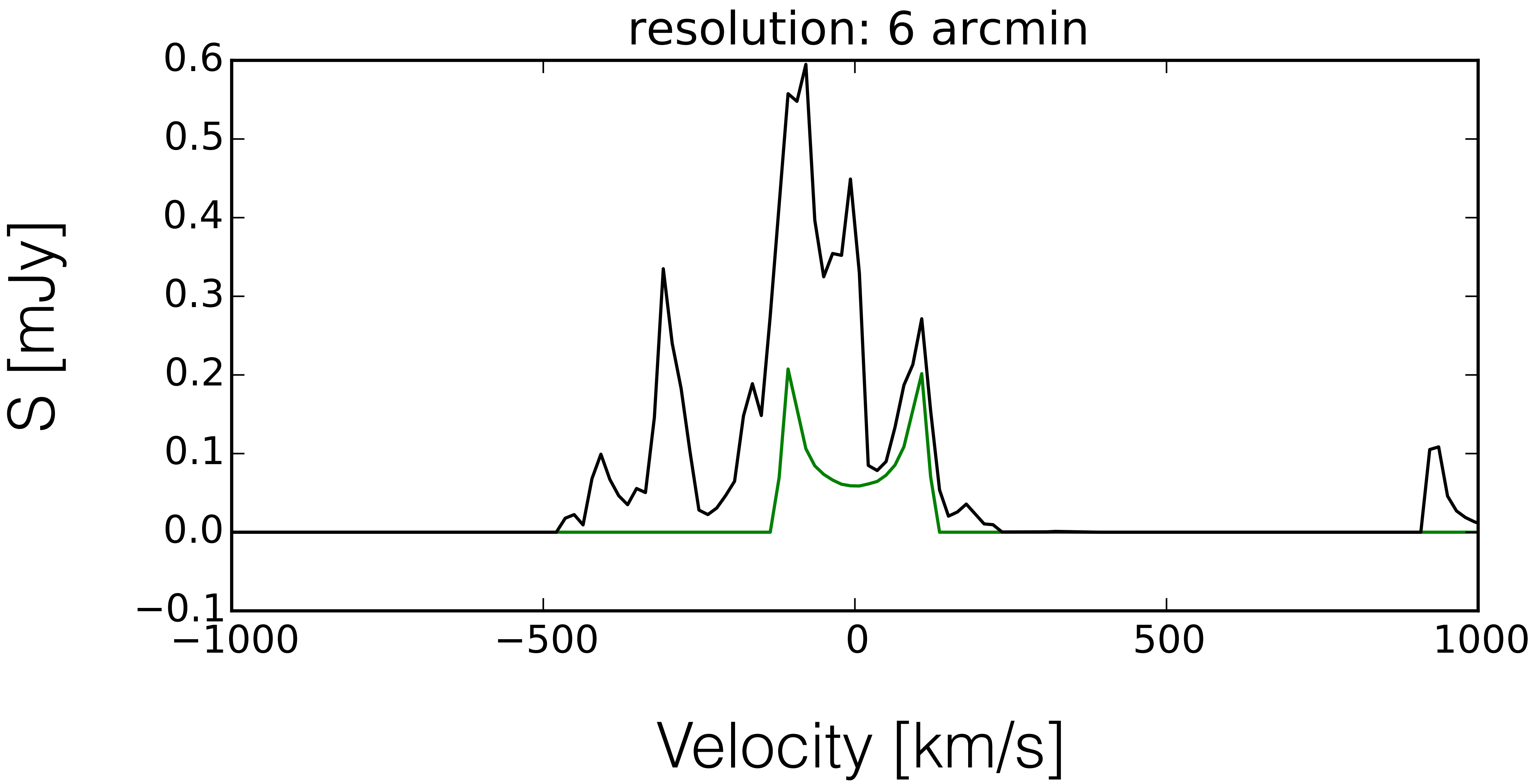
$\langle M_{HI} \rangle = 2.08 \times 10^9\ M_\odot$ ✔

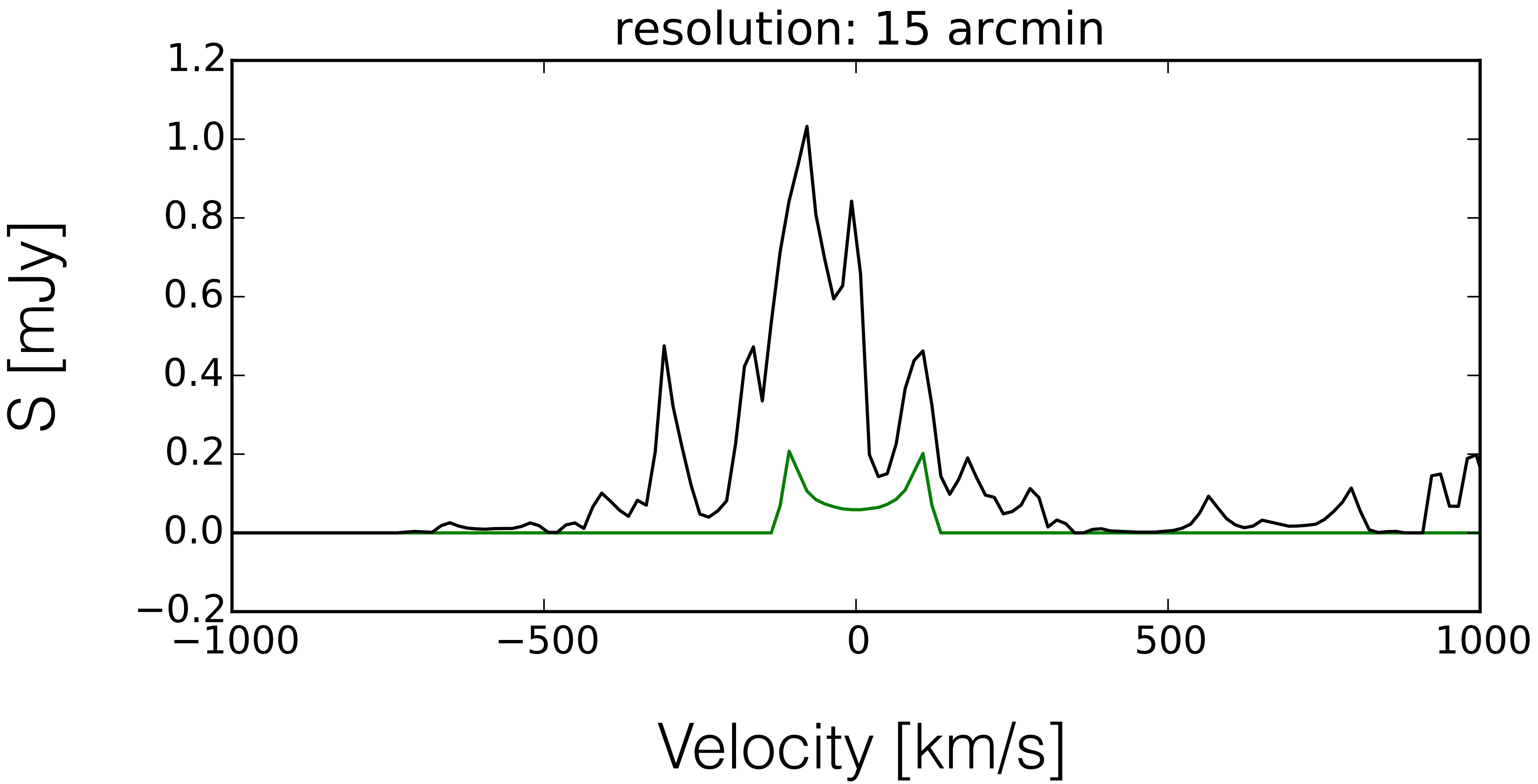
co-add: noise-free, **smoothed**



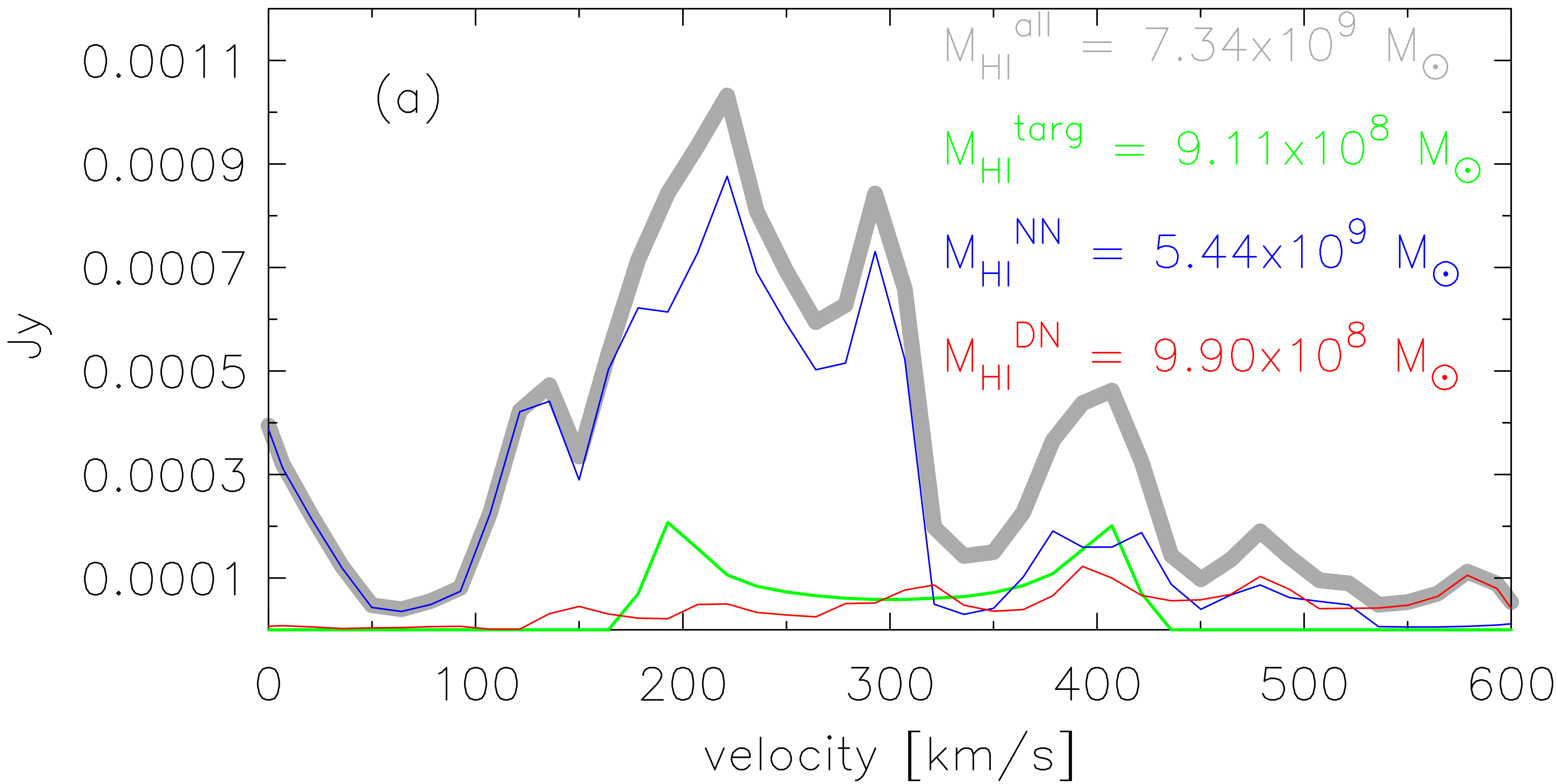
~ 6.44 x 10^9 $M_\odot$ contaminant mass per galaxy


resolution: 2 arcmin
S [mJy]
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
−1000
−500
0
500
1000
Velocity [km/s]

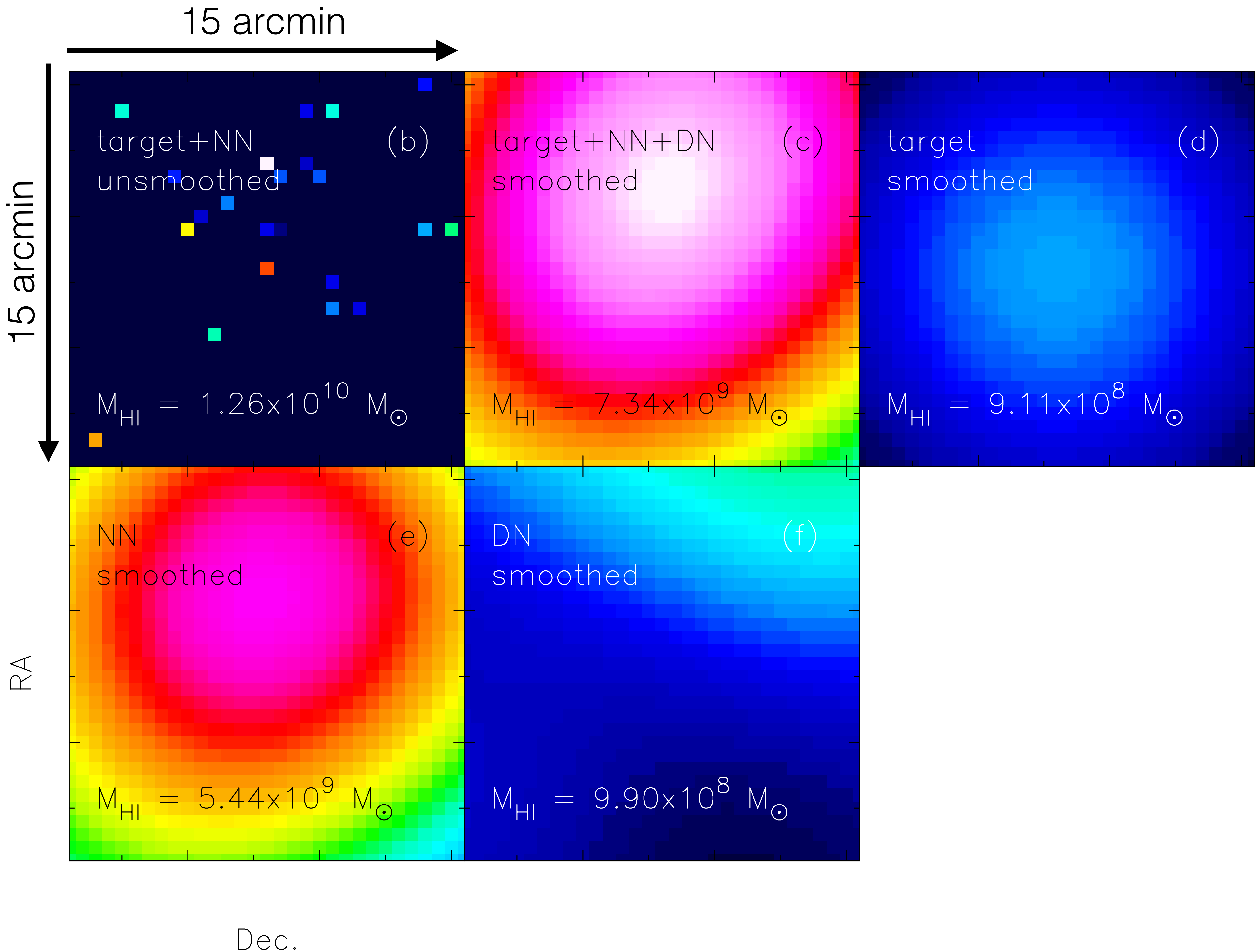

resolution: 6 arcmin
S [mJy]
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−1000
−500
0
500
1000
Velocity [km/s]



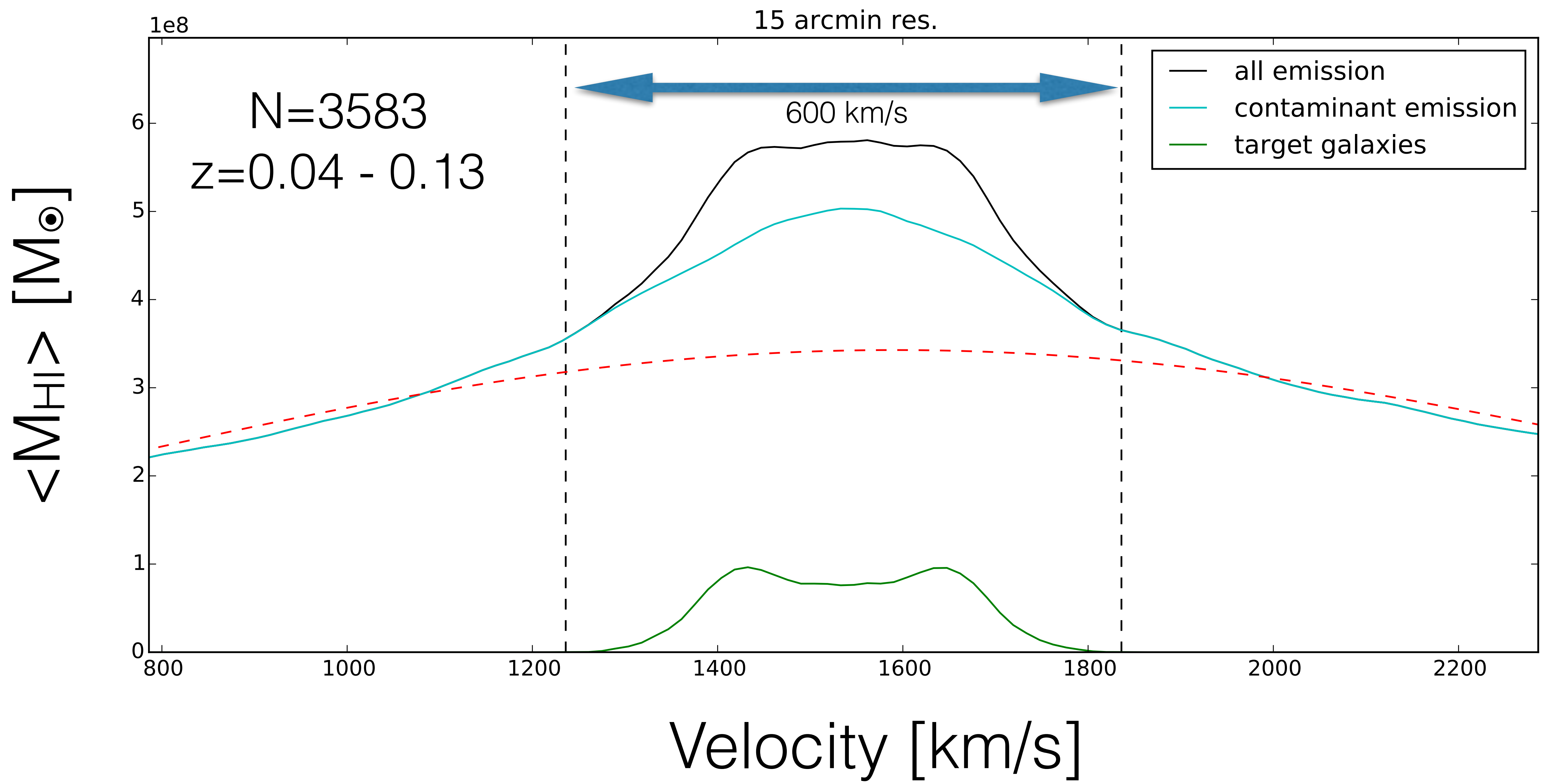
resolution: 15 arcmin
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
S [mJy]
−1000
−500
0
500
1000
Velocity [km/s]



(a)
M_HI^all = 7.34x10^9 M⊙
M_HI^targ = 9.11x10^8 M⊙
M_HI^NN = 5.44x10^9 M⊙
M_HI^DN = 9.90x10^8 M⊙
Jy
0.0001
0.0003
0.0005
0.0007
0.0009
0.0011
0
100
200
300
400
500
600
velocity [km/s]


15 arcmin
15 arcmin
target+NN
unsmoothed
(b)
M_{HI} = 1.26x10^{10} $M_\odot$
target+NN+DN
smoothed
(c)
M_{HI} = 7.34x10^{9} $M_\odot$
target
smoothed
(d)
M_{HI} = 9.11x10^{8} $M_\odot$
NN
smoothed
(e)
M_{HI} = 5.44x10^{9} $M_\odot$
DN
smoothed
(f)
M_{HI} = 9.90x10^{8} $M_\odot$
RA
Dec.

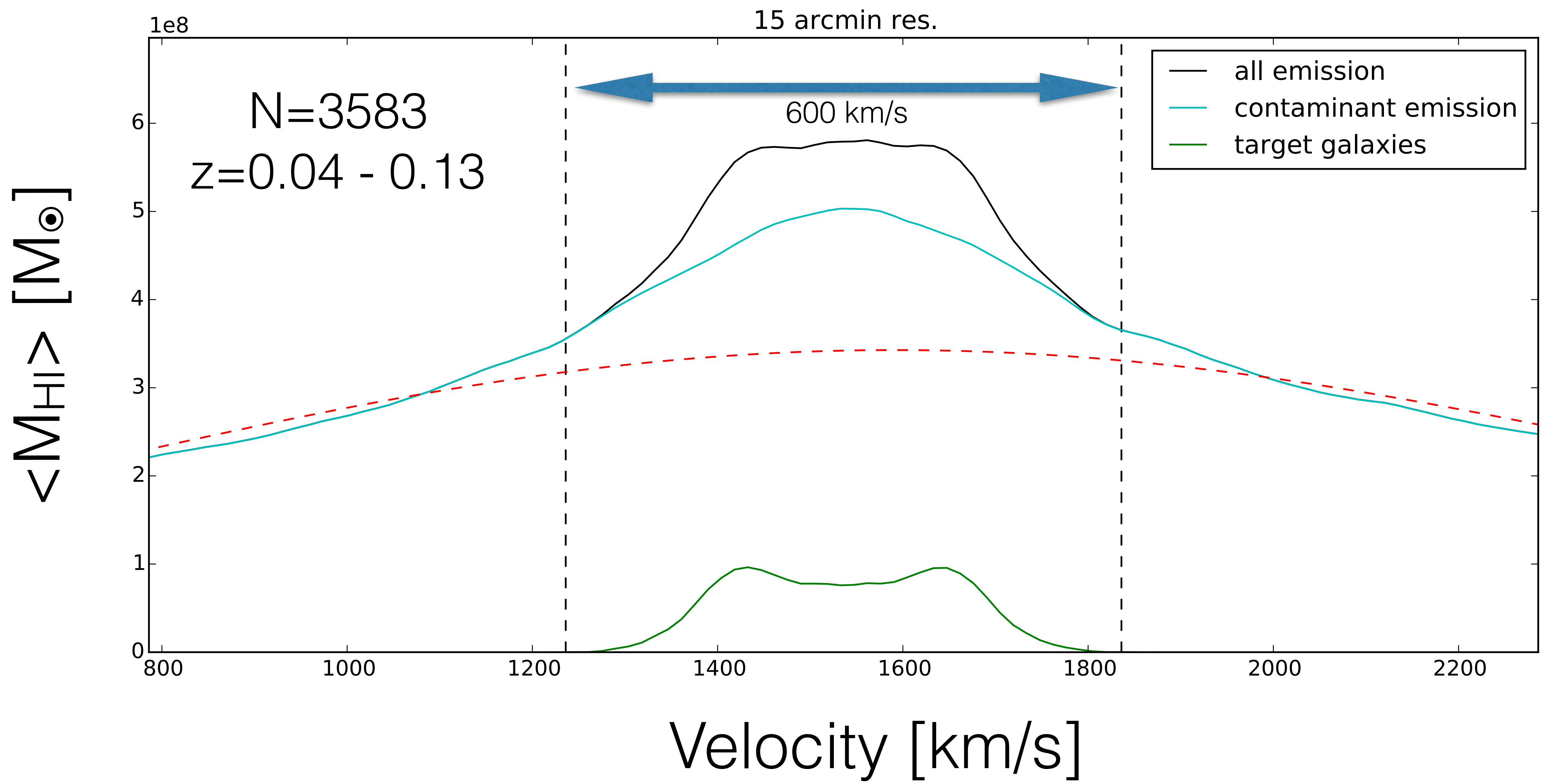
Mock Parkes co-add:



$<M_{HI}> = 5.84 \times 10^{10}\ M_{\odot}$ ✗

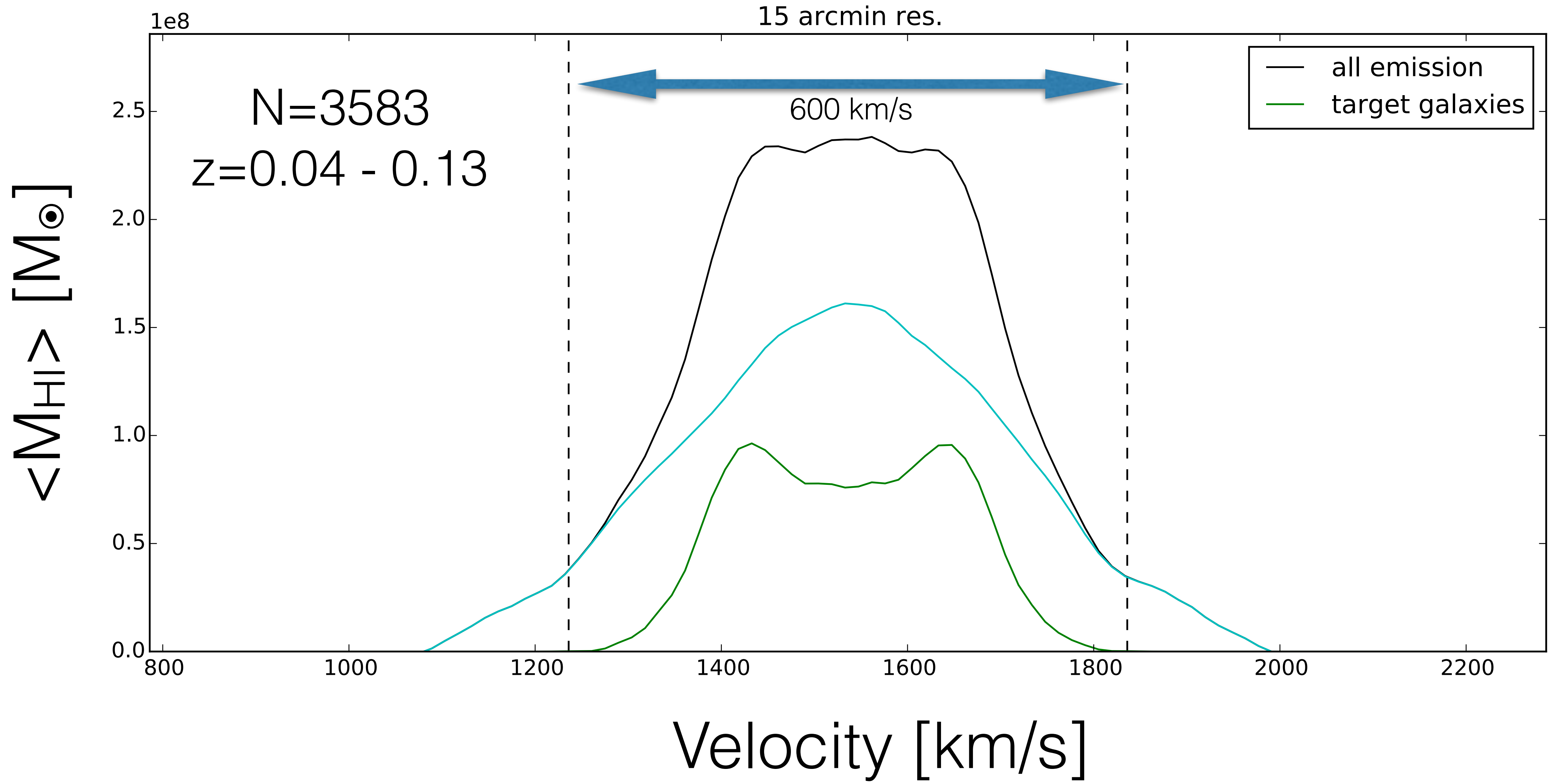
$<M_{HI}> = 2.08 \times 10^{9}\ M_{\odot}$ ✓

Typical Parkes co-add:



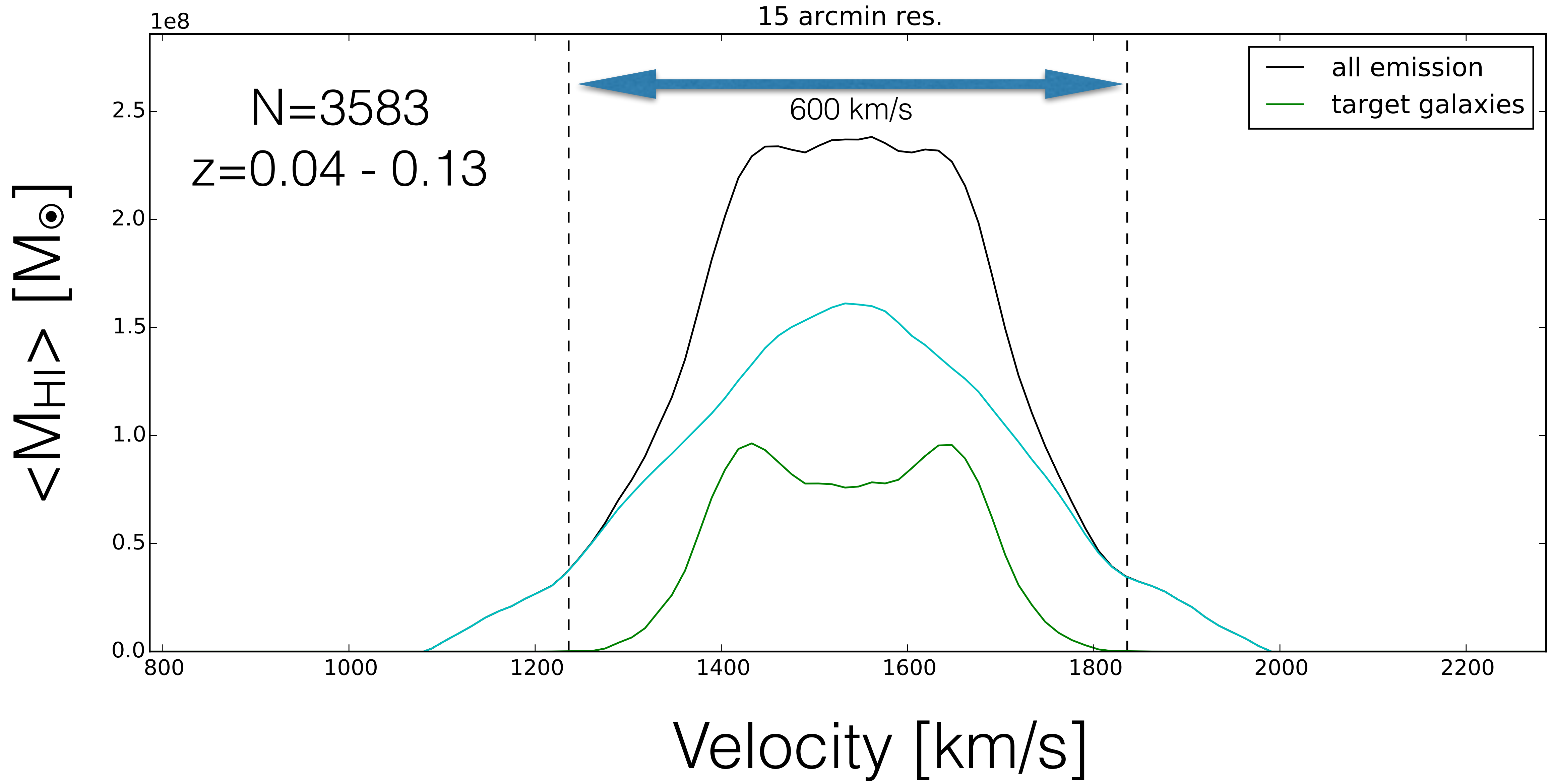
~ 5.63 x 10^{10} $M_{\odot}$ contaminant mass per galaxy

Typical Parkes co-add:



~ 4.51 x 10^9 $M_\odot$ contaminant mass per galaxy

Typical Parkes co-add:



$\langle M_{HI} \rangle = 6.62 \times 10^9\ M_\odot$

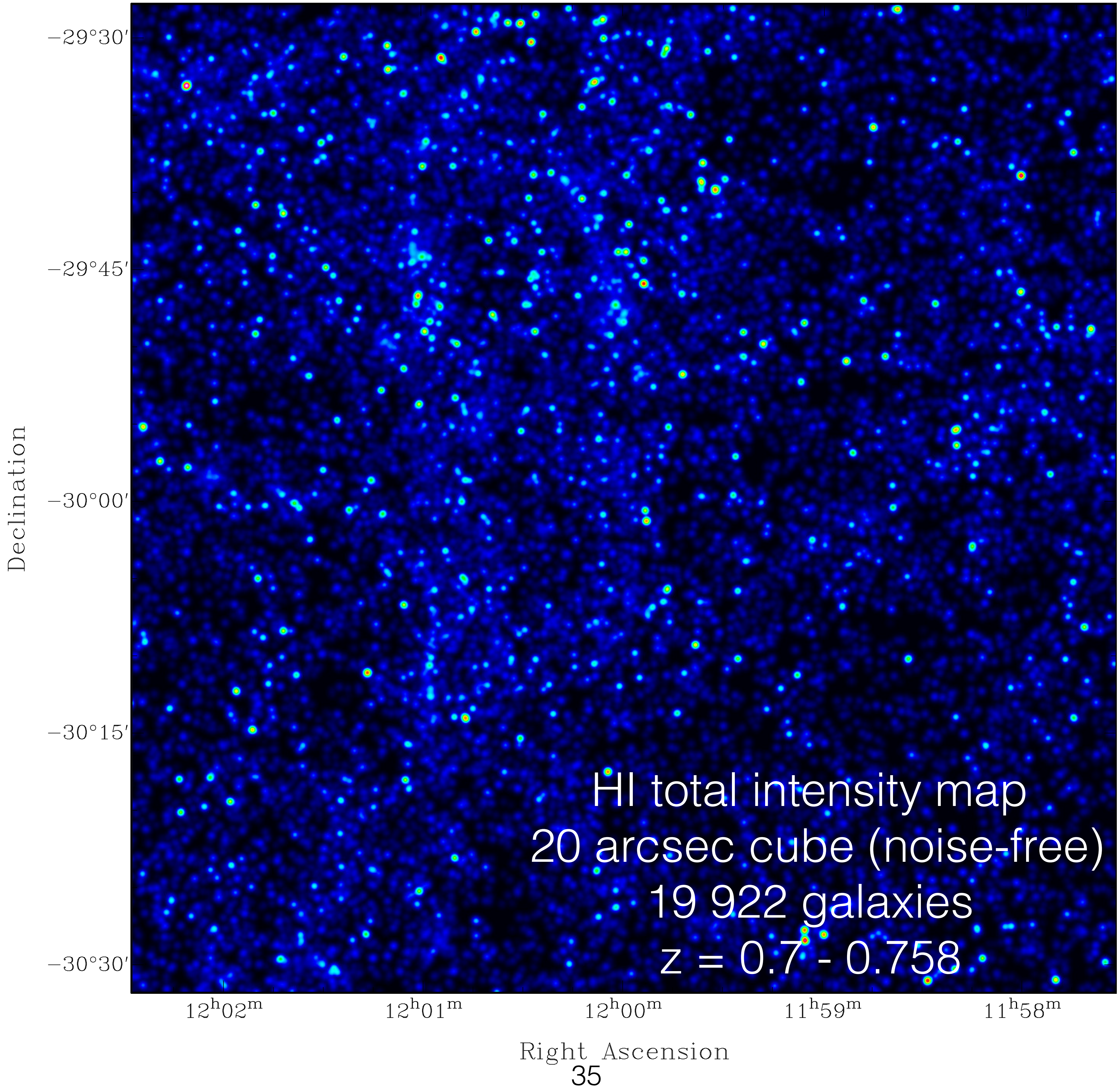
mock

$\langle M_{HI} \rangle = (6.93 \pm 0.17) \times 10^9\ M_\odot$

Delhaize + (2013), 3277 2dFGRS galaxies

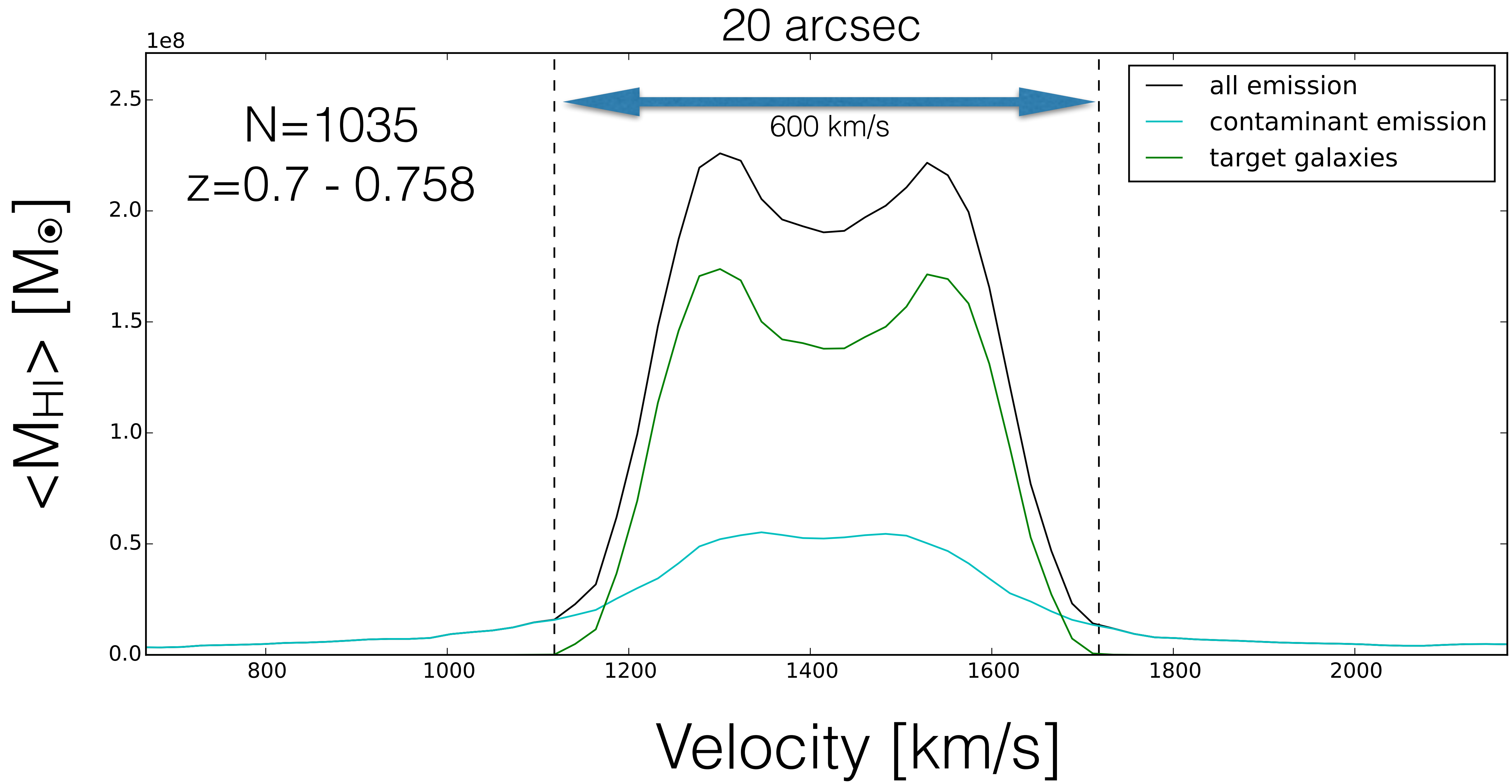
High-z mock stacking experiment

- Generate high-z synthetic cube:
  - 0.7 sq deg
  - $z = 0.7 - 0.758$
  - 19 922 galaxies ($1.03 \times 10^{13}\ M_\odot$)
  - Spatial resolution $\theta = 20$ arcsec.
  - Extract spectra at (1035) galaxies with $M_* > 10^{10}\ M_\odot$



−29°30′
−29°45′
−30°00′
−30°15′
−30°30′
Declination
12h02m
12h01m
12h00m
11h59m
11h58m
Right Ascension
HI total intensity map
20 arcsec cube (noise-free)
19 922 galaxies
z = 0.7 - 0.758

Noise-free LADUMA co-add



Fractional mass contribution from target galaxies: 0.73

Average confused mass: $1.03 \times 10^9\ M_\odot$

Summary

- We have the methods & machinery to produce realistic synthetic HI data cubes.

Summary

- We have the methods & machinery to produce realistic synthetic HI data cubes.
- Synthetic products will be useful to MeerKAT survey teams.

Summary

- We have the methods & machinery to produce realistic synthetic HI data cubes.
- Synthetic products will be useful to MeerKAT survey teams.
- Mock HI stacking experiments show high levels of source confusion in low-z stacking experiments.

Summary

- We have the methods & machinery to produce realistic synthetic HI data cubes.
- Synthetic products will be useful to MeerKAT survey teams.
- Mock HI stacking experiments show high levels of source confusion in low-z stacking experiments.
- Are current Ω_{HI} measures too high?
- Source confusion will not dominate LADUMA data, especially if we know how to correct for it.