TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

pingar™

CSIRO

Fuelling a Radio Source or Just in the Way? Applying Machine Learning Techniques to Redshifted HI Absorption*

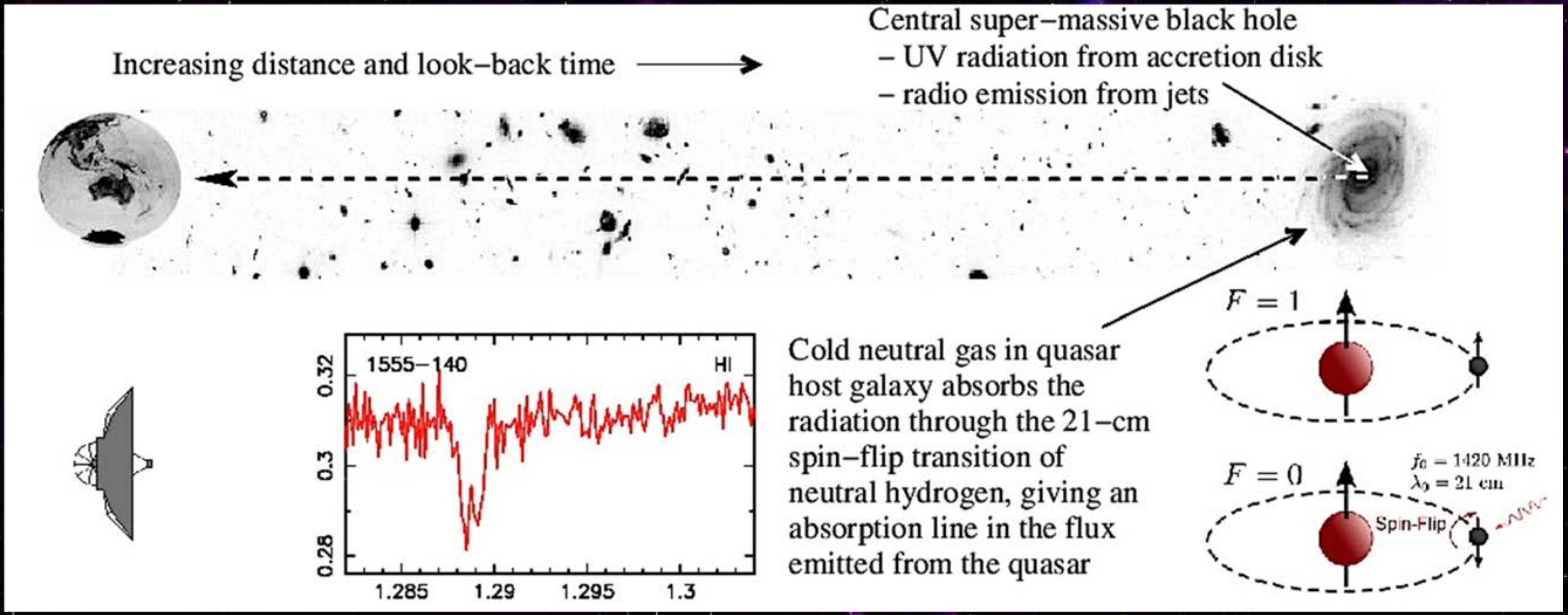
S. J. Curran, S. W. Duchesne, A. Divoli & J. R. Allison*

Absorption of the 21-cm transition of neutral hydrogen (HI) traces the cool component of the neutral gas, the reservoir star formation history. Also provides a useful probe of

- Baryonic mass density
- Evolution of large-scale structure
- Epoch of Reionisation
- Variations on the fundamental constants (α, μ, g_p)

Unlike the Lyman-α transition of HI (λ = 1216 Å), 21-cm can be observed at $z = 0$ by ground-based telescopes day or night (cf. $z > 1.7$).

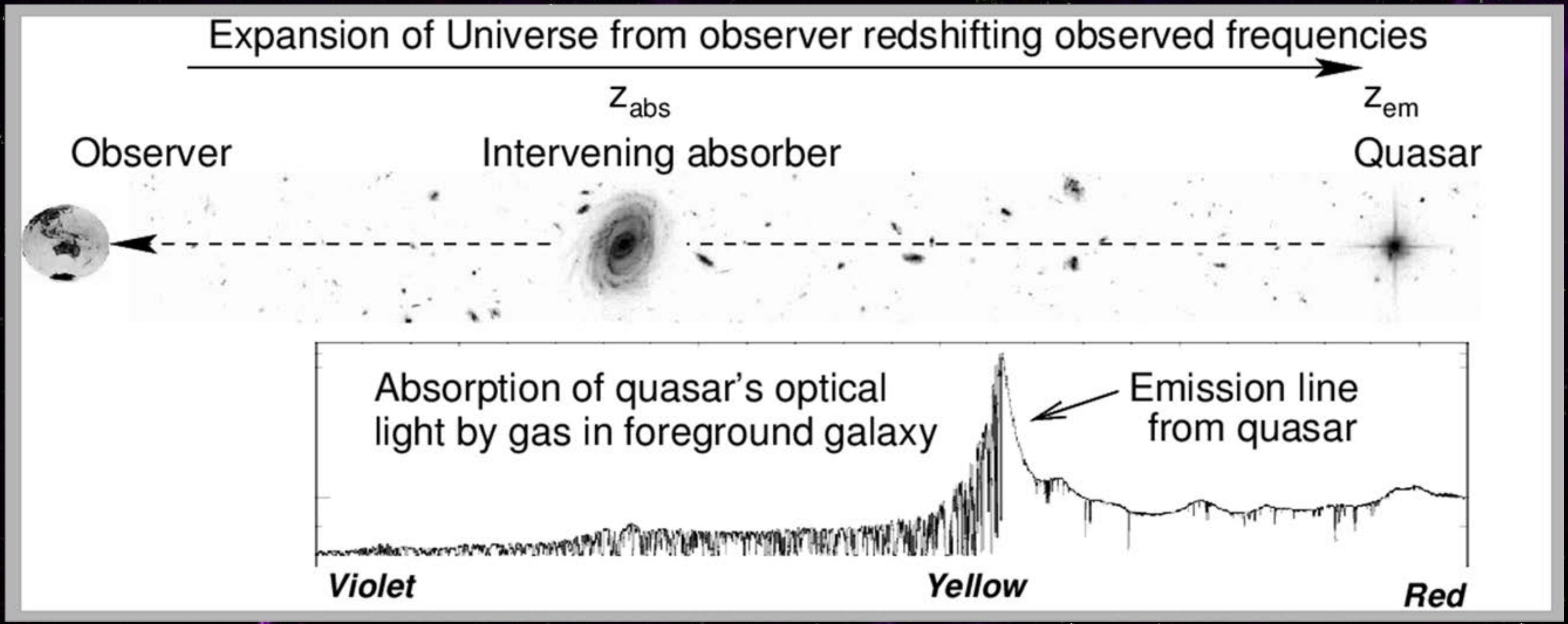
Unlike 21-cm emission, can be readily detected at $z \gtrsim 0.2$, since absorption strength only dependent upon column density and background flux.



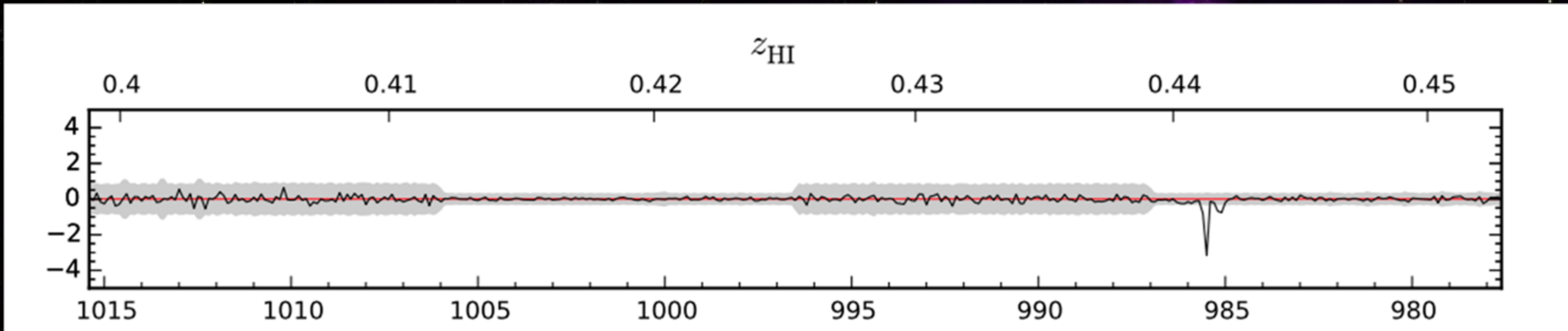
However, without optical spectrum (observationally expensive) it is not known whether the absorption occurs in a galaxy *intervening* the sight-line to a more distant radio source (quiescent) or is due to gas *associated* with the host of the radio source (active).

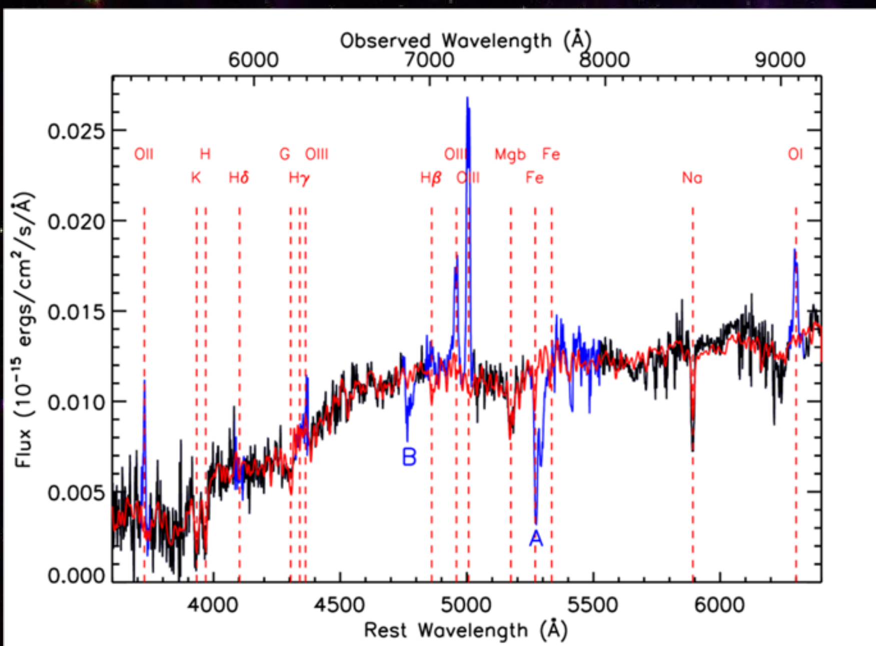
Furthermore, optical pre-selection of targets introduces biases to both types:

- *Associated* - at high redshift these are very UV luminous in source rest-frame ionising the gas, beyond detection (*Curran et al*. 2008) and possibly ionising the whole galaxy (*Curran & Whiting*, 2012)
- *Intervening* - biases against
  - optically obscured sight-lines (*Ellison et al.*, 2005)
  - absorbers rich in molecular gas (*Curran et al*. 2006,2011), which provide probes of physical and chemical conditions and anchor lines for changes in μ



For example, for a 1015 – 710 MHz (z = 0.4 – 1.0) scan with the *Boolardy Engineering Test Array* of the *Australian Square Kilometre Array Pathfinder* (ASKAP, *Allison et al.*, 2015)



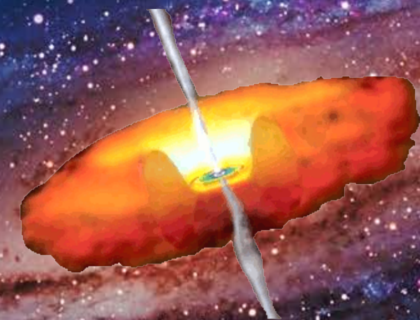


No previous spectroscopy of PKS B1740−517 and photometric redshifts range z = 0.347 – 0.63, which corresponds to a difference of 2 Gyr in look-back time!

1.5 hours of Director's Discretionary Time *on Gemini-South* confirms that absorption is associated with $z_{\rm opt}$ = 0.44230 ± 0.00022 cf. $z_{\rm HI}$ = 0.44129230 ± 0.00000040

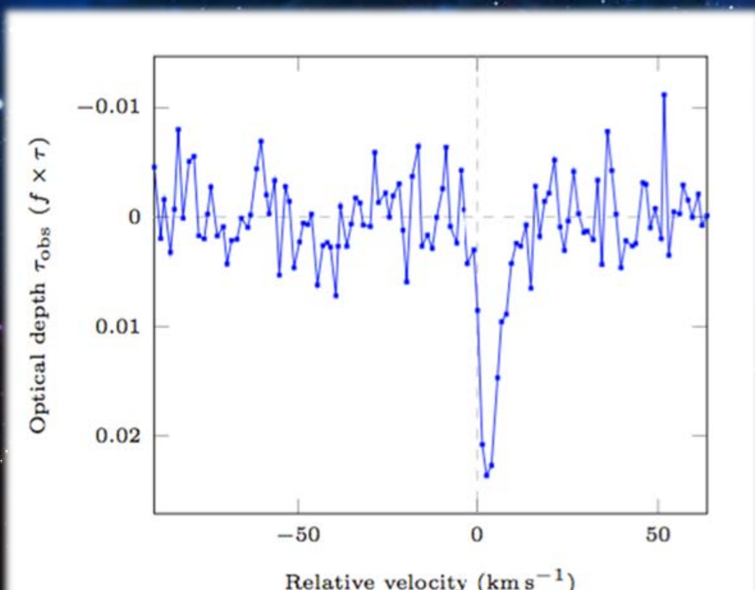
This is one line-of-sight with a six-antenna prototype. The full *First Large Absorption Survey in HI* (FLASH) on ASKAP will search 150,000 sight-lines with 36 antennas!

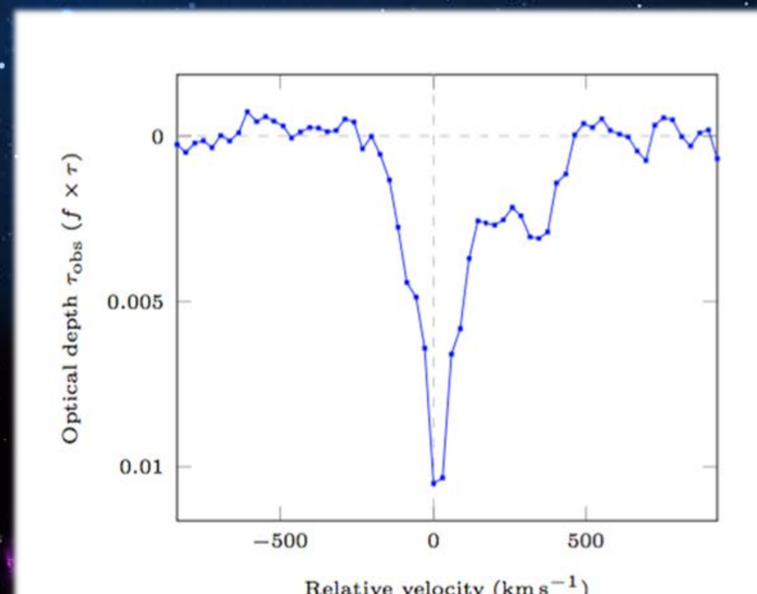
Want to determine whether *associated* or i*ntervening* without need for optical redshift
- populations of active and quiescent galaxies which may be too faint/obscured for optical identification



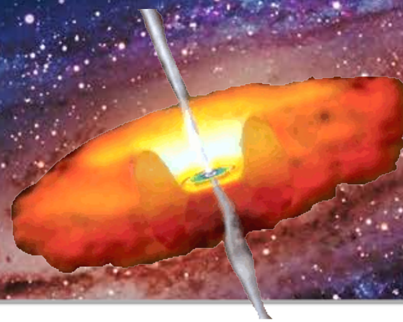
Intervening - absorption of background radio source by gas in bulk of quiescent galaxy

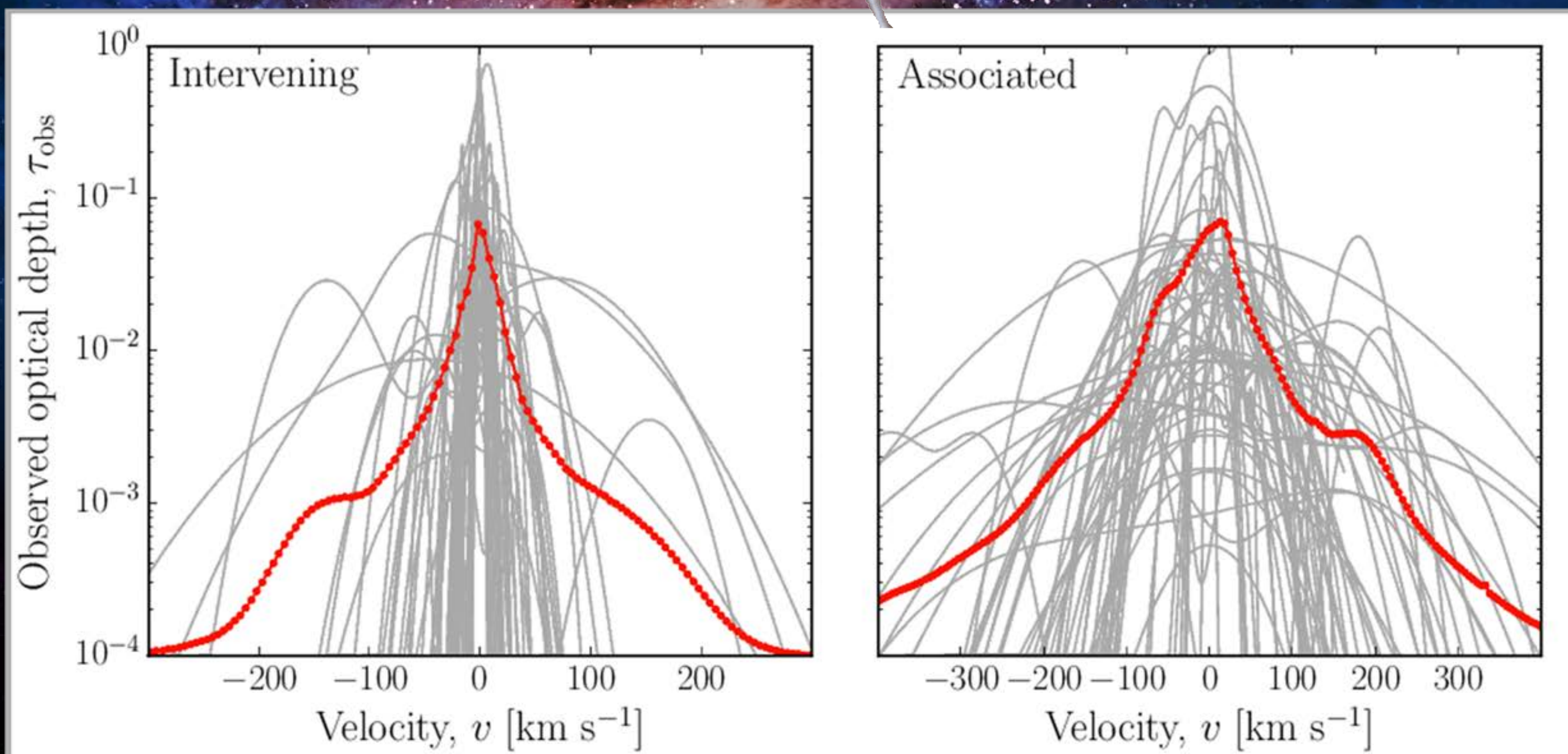
Associated - additional wide component from high velocity gas close to the AGN



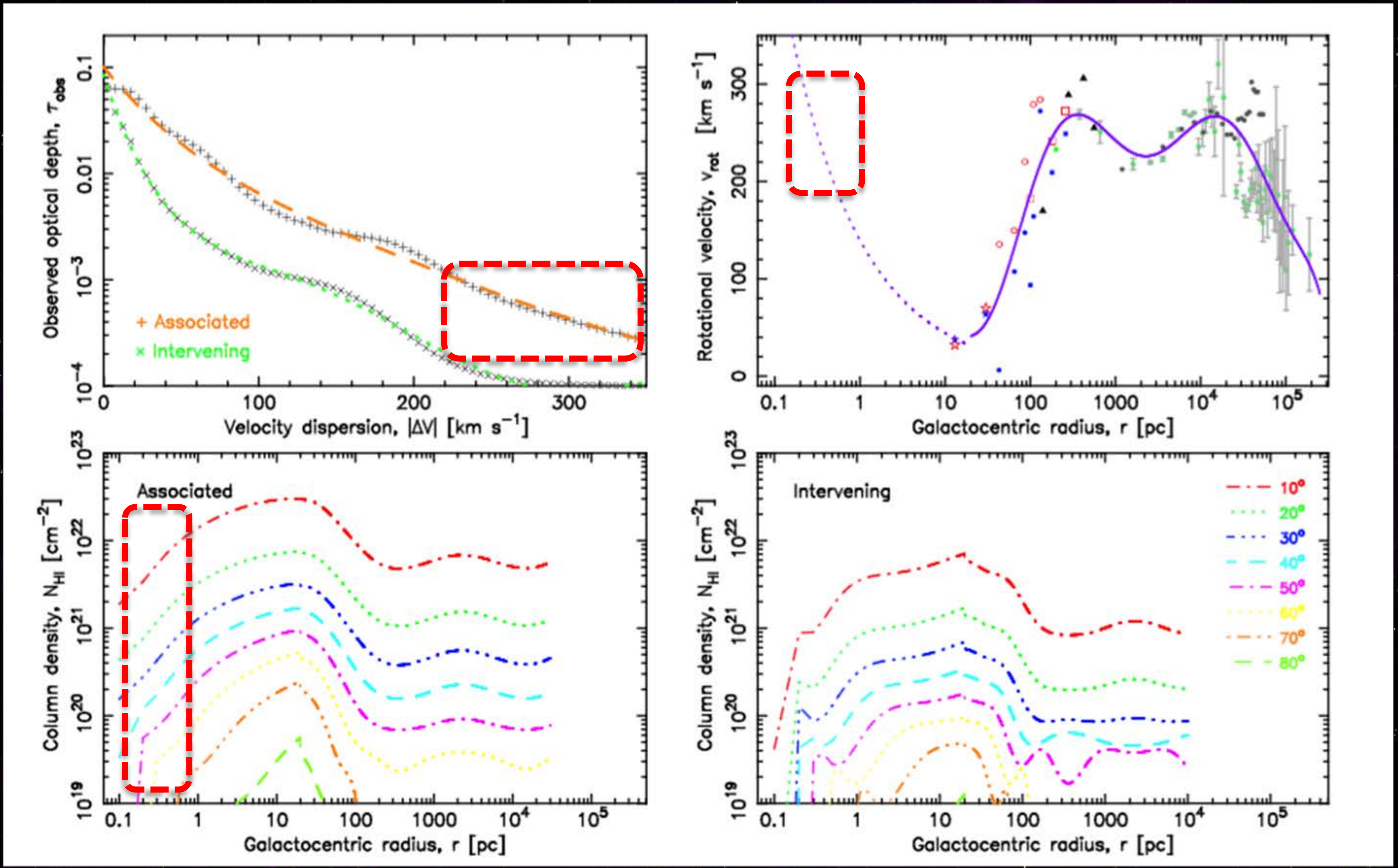


Although the mean associated displays wider low optical depth wings (at $\tau \lesssim 10^{-3}$), there are also some narrow associated and some wide intervening profiles, making it difficult to predict the absorber type based on profile width alone





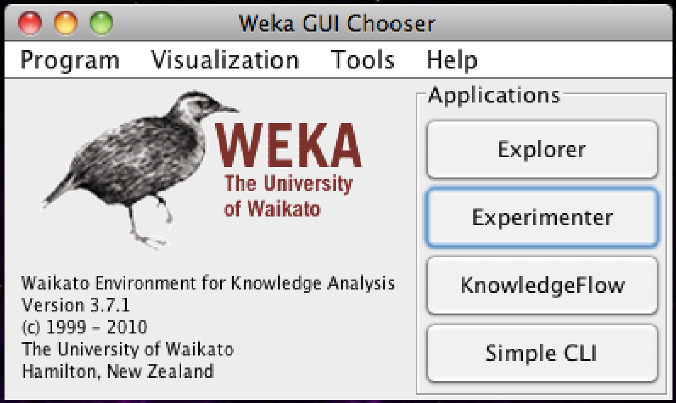
Combining mean optical depth-dispersions with rotation curve (Circinus & Galactic)



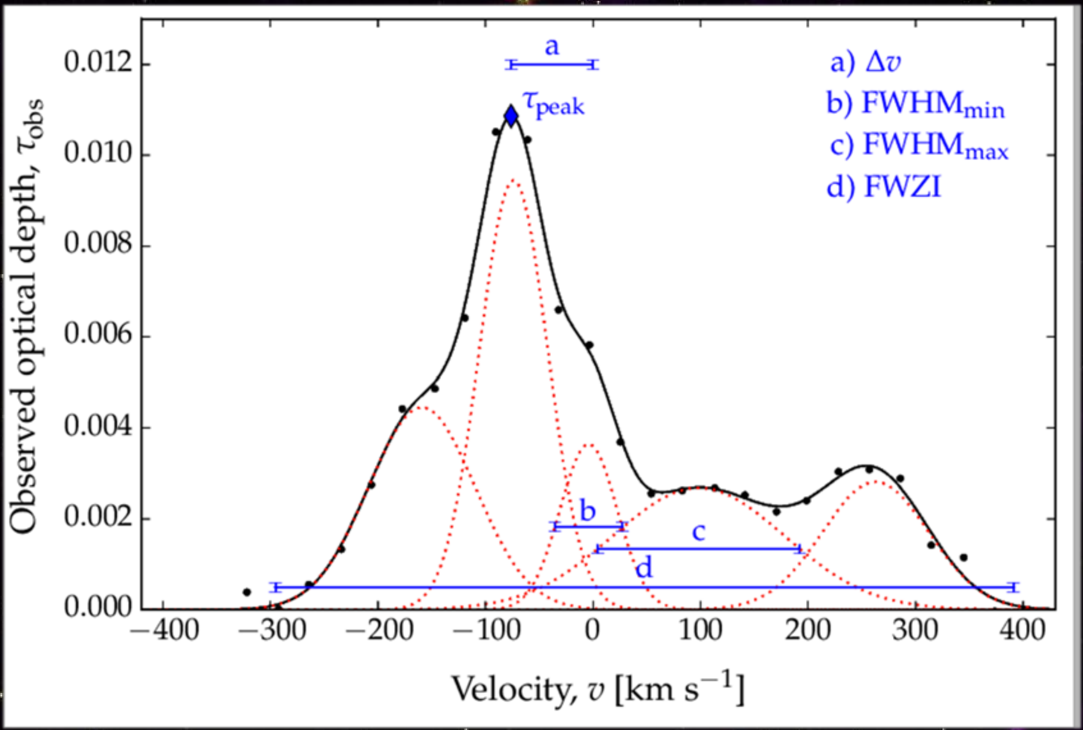
MACHINE LEARNING

We classify each spectrum by the following features

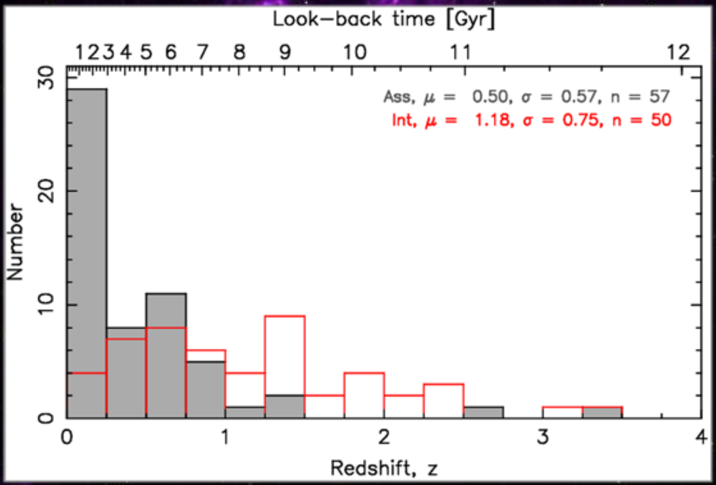
- The number of Gaussian fits (components)
- Full-Width Zero Intensity
- Full-Width Half Maximum *
- Peak optical depth *
- The average velocity offset of the components from the mean weighted redshift



However, we do not use the redshift itself, as associated absorption occurs predominately at $z \lesssim 1$, due to ionisation of gas by UV from the AGN (Curran & Whiting 2012)



Furthermore, the aim is to predict absorber type with no prior knowledge of the redshift of the background continuum source



MNRAS 462, 4917

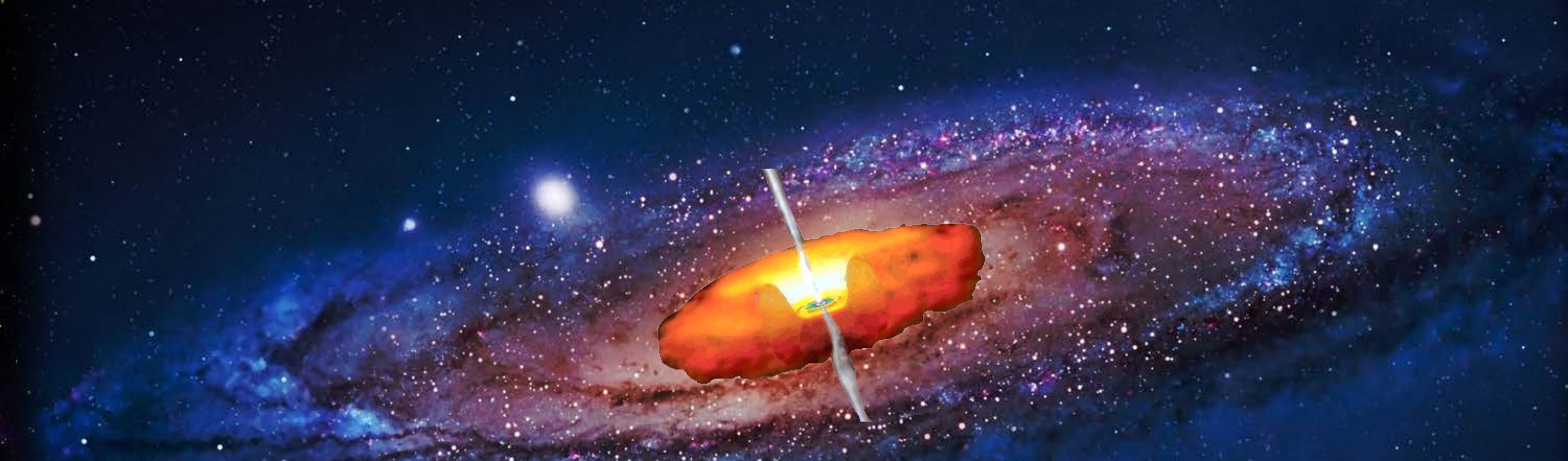
Testing several different algorithms, we obtain a ≈ 80% accuracy

Algorithm	P [%]	R [%]	F [%]	A [%]	$\|\overline{\sigma}\|$	κ
	Whole (55 associated/43 intervening)					
Bayesian Network	81.2	80.6	80.7	80.6	0.194	0.611
Sequential Minimal Optimisation	80.6	78.6	78.6	78.6	0.214	0.577
Classification Via Regression	80.0	79.6	79.7	79.6	0.302	0.590
Logistic Model Tree	80.9	80.6	80.7	80.6	0.346	0.610
Random Forest	81.6	81.6	81.6	81.6	0.305	0.422

- As expected, this is mostly driven by the line-width information.
- Optical depth is not an important feature at this stage.
- Surprisingly, the number of components is not an important feature (at least at this stage) – we expect the associated profiles to be more complex. However, as higher spectral resolution becomes available with the SKA this may change.

Still a sample sample (≈ 50 of each type), however, as more are added expect predictive power of classifier to improve in preparation for the large number of SKA (and precursor) sight-lines.

Associated absorption is generally wider than intervening



$\mathrm{FWHM_{eff}} \equiv \int \tau \, dv / \tau_{peak}$