Works on OH data from Arecibo telescope

Aim:
- Find if OH exists in 79 MS lines-of-sight.
- Check if we can find OH in the clouds where there's more HI.

Data:
- Download the OH data from Arecibo Millennium Survey.
- Observations along 100 sightlines have been taken using On/Off-source method:

    '1325+32', '3C105', '3C109', '3C120', '3C123N', '3C123', '3C131', '3C132', '3C133', '3C138', '3C141.0', '3C142.1', '3C154E', '3C154', '3C167E', '3C167', '3C172.0', '3C18', '3C190.0', '3C192N', '3C192', '3C207', '3C208.0', '3C208.1', '3C223', '3C225aN', '3C225a', '3C225b', '3C234', '3C236', '3C237N', '3C237', '3C245', '3C264.0', '3C270', '3C272.1', '3C273N', '3C273', '3C274.1', '3C286', '3C287', '3C293', '3C298', '3C310', '3C315', '3C318', '3C33-1', '3C33-2', '3C333', '3C33', '3C348', '3C409', '3C410', '3C454.3', '3C64', '3C75-1', '3C75-2', '3C75N', '3C75', '3C78W', '3C78', '3C79', '3C93.1', '3C98-1', '3C98-2', '3C98', '4C07.32', '4C13.65', '4C13.67', '4C16.33', '4C19.44N', '4C19.44', '4C21.28', '4C22.12', '4C32.44', '4C35.23', 'CTA21', 'DW0742+10', 'G196.6+0.2', 'G197.0+1.1', 'HD149881', 'HD215733', 'HD93521', 'NRAO140', 'P0320+05', 'P0347+05', 'P0428+20', 'P0531+19', 'P0820+22', 'P1055+20N', 'P1055+20', 'P1117+14', 'Q1552+19', 'T0526+24', 'T0556+19', 'T0629+10', 'T0640+09', 'TauA', 'W49', 'W51c3'

- Data structure:

```
IDL> help, LA
** Structure <19883498>, 24 tags, length=122928, data length=122926, refs=1:
   SRCNAME         STRING    Array[100]
   RA1950          FLOAT     Array[100]
   DEC1950         FLOAT     Array[100]
   ELL             FLOAT     Array[100]
   BEE             FLOAT     Array[100]
   CFR_BD0         FLOAT     1420.41 MHz
   CFR_BD1         FLOAT     1665.40 MHz
   CFR_BD2         FLOAT     1667.36 MHz
   NRC             INT       100
   VLSR_BD0        FLOAT     Array[2048]
   VLSR_BD1        FLOAT     Array[2048]
   VLSR_BD2        FLOAT     Array[2048]
   I_EM_AVG_BD0    FLOAT     Array[2048]
   I_EM_MED_BD0    FLOAT     Array[2048]
   I_ABS_AVG_BD0   FLOAT     Array[2048]
   I_ABS_MED_BD0   FLOAT     Array[2048]
   I_EM_AVG_BD1    FLOAT     Array[2048]
   I_EM_MED_BD1    FLOAT     Array[2048]
   I_ABS_AVG_BD1   FLOAT     Array[2048]
   I_ABS_MED_BD1   FLOAT     Array[2048]
   I_EM_AVG_BD2    FLOAT     Array[2048]
   I_EM_MED_BD2    FLOAT     Array[2048]
   I_ABS_AVG_BD2   FLOAT     Array[2048]
   I_ABS_MED_BD2   FLOAT     Array[2048]
```

Note:

“”””

median versus average:

The average of a bunch of points is the least-squares value. That is, when you calculate the sum of the squares of the residuals,

sum_of_squares_of_residuals = \Sigma (\Delta y)_n^2

where (\Delta y)_n is the residual

(\Delta y)_n = y_n - y0

where y0 is the adopted result of combining all of the measurements y_n. If you adopt y0 as being the value that minimizes the sum of squares of residuals, then y0 is referred to as the 'least-squares value'. For a bunch of datapoints y_n, taking the average gives a value for y_0 that is the same as the least-squares value. So the average value is the same as the least-squares value.

When the probability density function (PDF) of the residuals follows a Gaussian (we call this 'Gaussian statistics'), then the least-squares value is the best estimate for y0. This is the 'AVG' spectrum.

In contrast, the median value is the one for which the number of ppositive residuals equals the number of negative ones. This is the same as the median. For Gaussian statistics, the median is NOT the best estimate. The median is, however, the best estimage when the PDF of the residuals follows a double-sided exponential PDF.

(see sections 0, 6, 13, especially 13.5, of my handout on least squares, available at

http://w.astro.berkeley.edu/~h

(click on 'lsfit2008.pdf').

Now, in practice:

If the two spectra I_EM_AVG and I_EM_MED differ only because I_EM_MED is noisier, then you can be pretty sure that the PDF is Gaussian, i.e. that there are no interference spikes. So the proper thing to do is to take the average spectrum I_EM_AVG.

However, if the AVG spectrum has features that are not present in the MEDIAN spectrum, then the PDF is not gaussian, probably because of interference. Then one option is to simply use the median. Better is to look in detail at the data and edit out bad datapoints (using a program such as our 'flag_rfi.pro') to make the statistics Gaussian, or at least nearly so; then you can take the average of the edited data.

So the idea here is to use the difference between AVG and MED an indicator for whether or not you need to edit the data.

“”””

- BD0 is HI at 1420.4 MHz. BD1 is OH at 1665 MHz and BD2 is OH at 1667 MHz.
- Each spectrum has 2048 channels with the bandwidth of ~0.77643MHz and the step of ~0.37912 kHz (0.06829 km/s).

- I_EM_AVG_BD* is the Off-source spectra, $T_{off\text{-}source}$.
- I_ABS_AVG_BD* is the on-source antenna temperature ***minus*** the off-source antenna temperature: $(T_{on\text{-}source} - T_{off\text{-}source}) = T_{continuum} * e^{-\tau}$.

On the emission lines, the continuum baseline is the off-source system temperature (actually, twice that; what's plotted is Stokes I, which is the sum of the two orthogonal polarizations). This is normally about 50 K. It can be higher if the background continuum is high, which might happen near the Galactic plane.

On the absorption lines: the continuum baseline is the on-source antenna temp minus the off source antenna temp (again, times two because it's Stokes I). Thus, in the case of 3C131, the source deflection (on - off) is about 47.7 K. This is not related to the off-source system temp, except insofar as if the source is extended it might be producing some antenna temp in the 'off-source' position. You can convert the vertical scale of the 3C131 absorption plot to exp(-tau) simply by dividing all of the spectral points by 47.7 K.

$T_{off\text{-}source} = I\_EM\_AVG\_BD^*$
$T_{abs} = T_{on\text{-}source} - T_{off\text{-}source} = I\_ABS\_AVG\_BD^*$
For the on/off-source measurements, I use the following equations to calculate the excitation temperature T_{ex} and Opacity $e^{-\tau}$:
$T_{off\text{-}source} = T_{bg} * e^{-\tau} + T_{ex}(1-e^{-\tau})$
$T_{on\text{-}source} = (T_{bg} + T_c) * e^{-\tau} + T_{ex}(1-e^{-\tau})$

So that:

$$e^{-\tau} = (T_{on\text{-}source} - T_{off\text{-}source})/T_{continuum}$$
$$T_{ex} = (T_{off\text{-}source} - T_{bg} * e^{-\tau}) / (1-e^{-\tau})$$

- In practice, $T_{off\text{-}source}$ and $T_{on\text{-}source}$ include the receiver temperature, T_{rx}, which is about 50K (in Stokes I). Thus, two equations above need to incorporate T_{rx}:

$$T_{off\text{-}source} = T_{bg} * e^{-\tau} + T_{ex}(1-e^{-\tau}) + T_{rx}$$
$$T_{on\text{-}source} = (T_{bg} + T_c) * e^{-\tau} + T_{ex}(1-e^{-\tau}) + T_{rx}$$

- Now, what is T_{rx}? Off the line, the equation for $T_{off\text{-}source}$ is

$$T_{off\text{-}source} = T_{bg} + T_{rx}$$

- T_{bg} contains the cosmic background, 2.8 K, plus the Galactic contribution, which I estimate from the Haslam et al. 408 MHz survey and apply a spectral index of 2.8 for brightness temperature. With this, the Galactic component typically comes out to be about 0.6 K or so, so this means that

$$T_{off\text{-}source} \sim 3.4 + T_{rx}$$
$T_{off\text{-}source}$ is 51.1K, so T_{rx} is 47.7K.

- Among 100 sources observed, I find there are 23 sources where OH is present, they are:

Source	l	b
3C105	187.63	-33.61
3C109	181.83	-27.78
3C123	170.58	-11.66
3C131	171.44	-7.80
3C132	178.86	-12.52
3C133	177.73	-9.91
3C154	185.59	4.00
3C167	207.31	1.15
3C18	118.62	-52.73
3C207	212.97	30.14
3C409	63.40	-6.12
3C410	69.21	-3.77
3C454.3	86.11	-38.18
3C75	170.26	-44.91
4C13.67	43.50	9.15
4C22.12	188.05	0.05
G196.6+0.2	196.64	0.17
G197.0+1.1	196.98	1.10
P0428+20	176.81	-18.56
P0531+19	186.76	-7.11
T0526+24	181.36	-5.19
T0556+19	190.09	-2.17
T0629+10	*201.53*	*0.51*

- As mentioned above, in order to estimate the continuum background brightness temperature for OH main lines, I use the all-sky continuum map from Haslam et al. 408 MHz survey and then apply a spectral index of 2.8 for brightness temperature.

$T_{bg} = 2.8 + T_{bg408} * (408./v)^{2.8}$

where:

v: frequencies of two OH main lines (1665.402 MHz or 1667.359MHz)

T_{bg408}*: Continuum brightness Temperature at 408 MHz, which is the average of the values within the area of Arecibo beam width (3.5 arcminutes) around 23 lines-of-sight, as illustrated in Figure ??.*

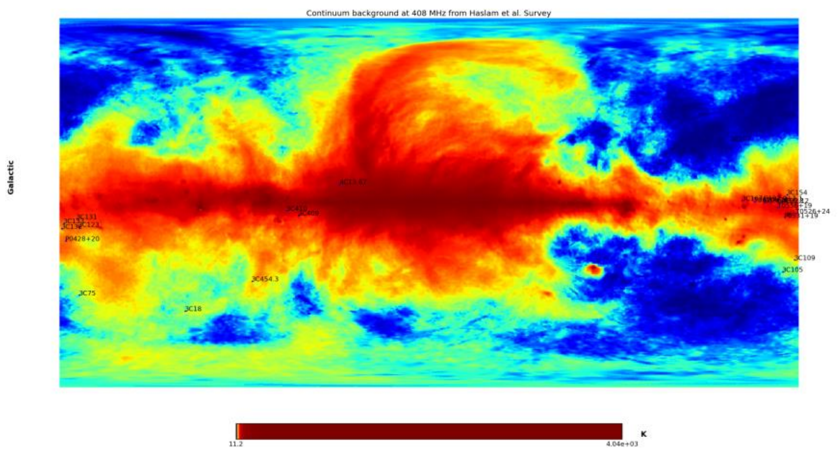

Figure ??: All-sky map of Continuum brightness Temperature at 408 MHz from Haslam et al.

- *Patches around lines-of-sight*

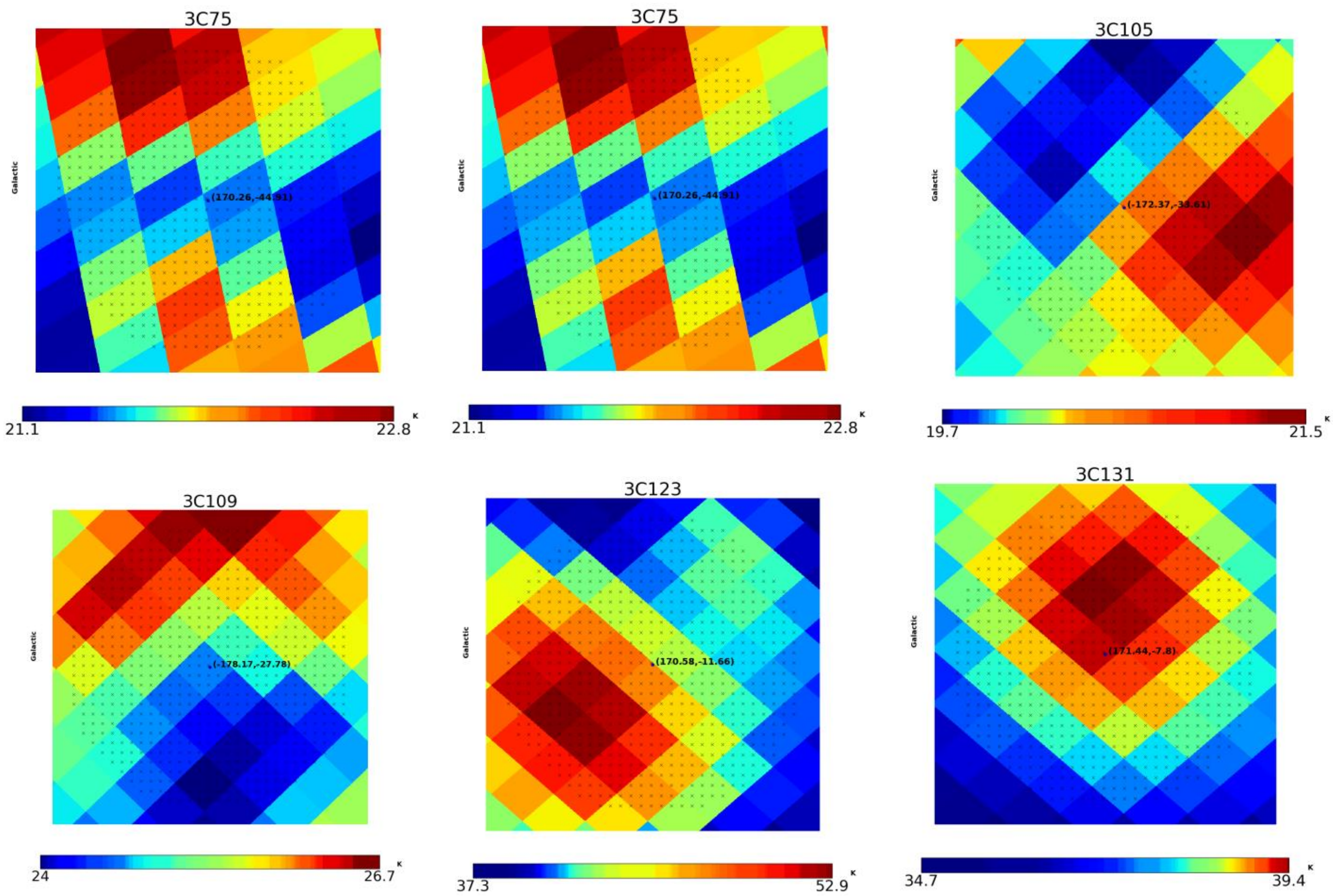

Figure ??: Patches of *continuum background temperature at 408MHz* T_{bg408} map around point-sources showing the pixels within the round area of beam width 51 arcminutes.

- Ningyu and Dili also used the MS data and we found the same result. Namely, the sources with OH but without CO are:

3C18, 3C109, 3C409, 3C132, P0531+19

(not really clear on P0531+19 because of its weak absorption lines).

The sources 3C109, 3C409, 3C132, P0531+19 are found in the regions with high N_{HI}. Please see the figure *??*.

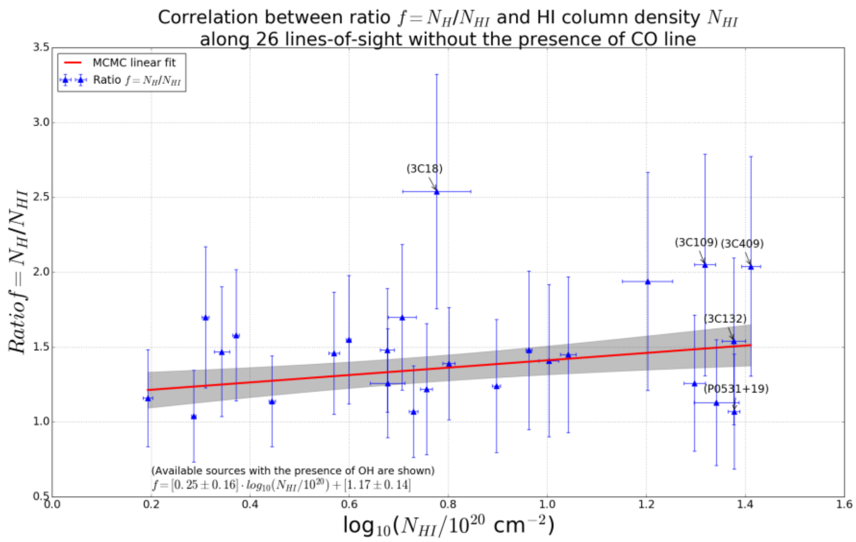

Figure ??: <title>

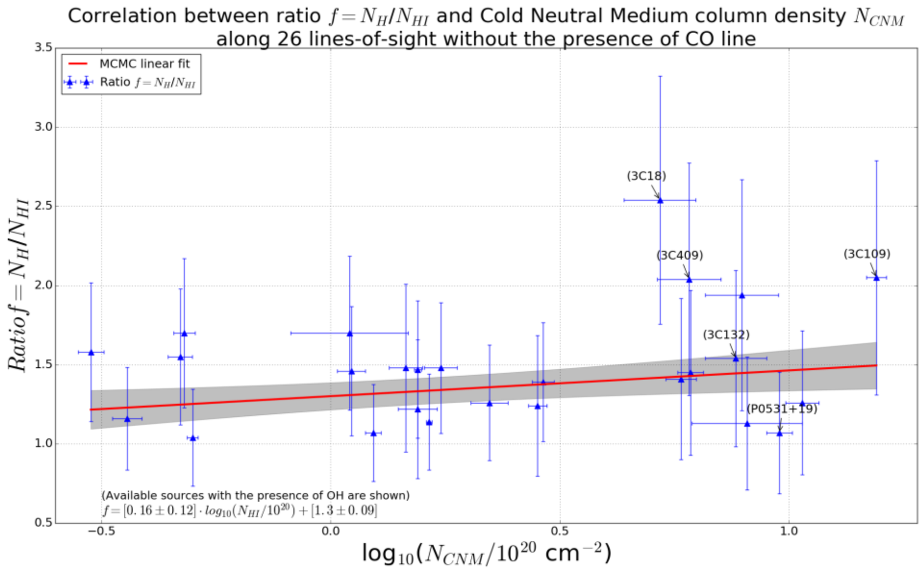

Figure ??: <title>

- I use the OH data from MS to calculate the Column Densities of OH1665 and OH1667 for each line-of-sight:

Step 1: I compute the opacity $e^{-\tau}$ from On/off-source measurements. And I plot the Absorption spectra.

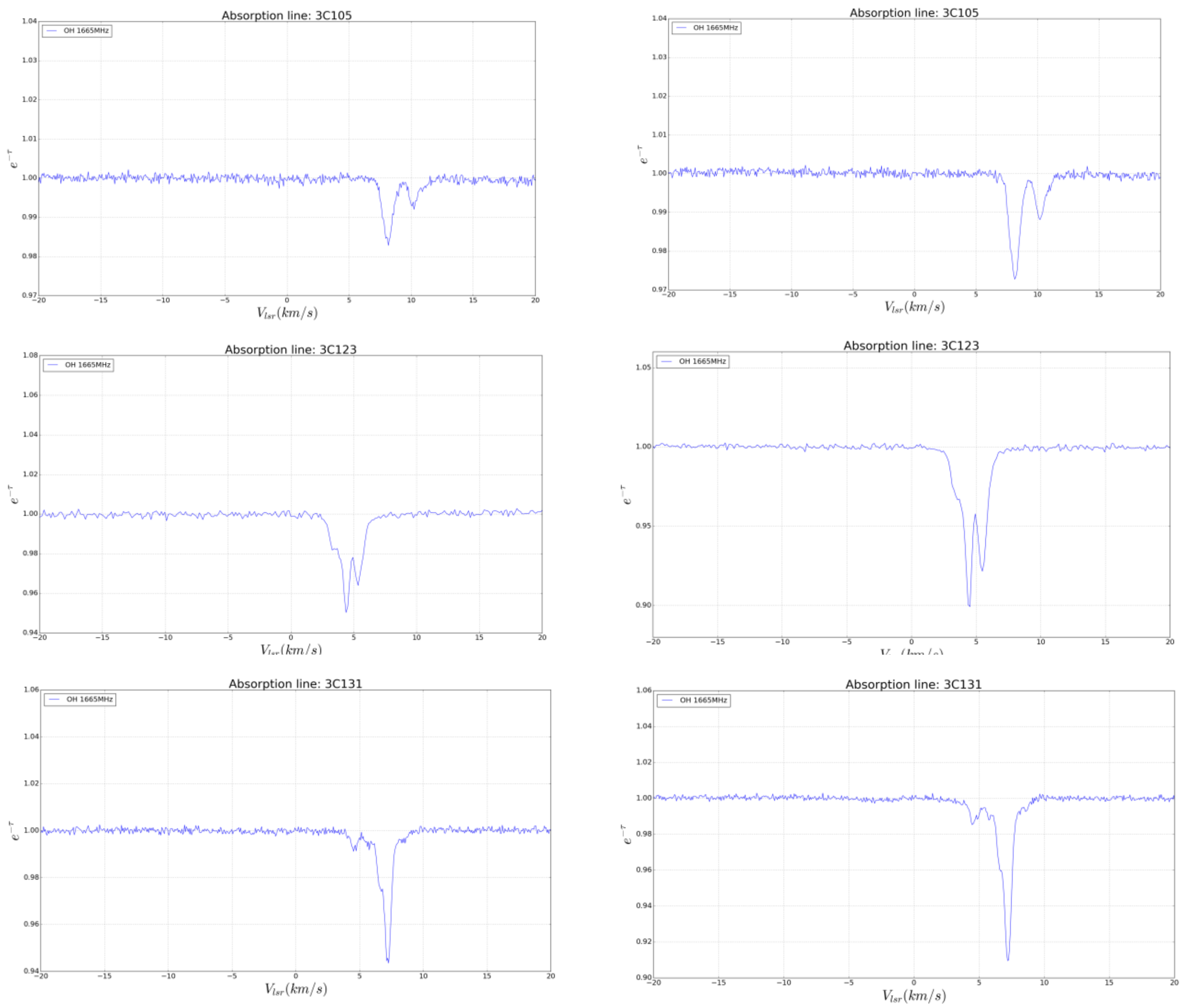

Figure ??: Absoption spectra of OH 1665MHz (left column) and OH 1667MHz (right column)

Step 2: I apply Gaussian fit for opacity $e^{-\tau}$ spectra just to obtain the velocity ranges (FWHM) of the peaks.

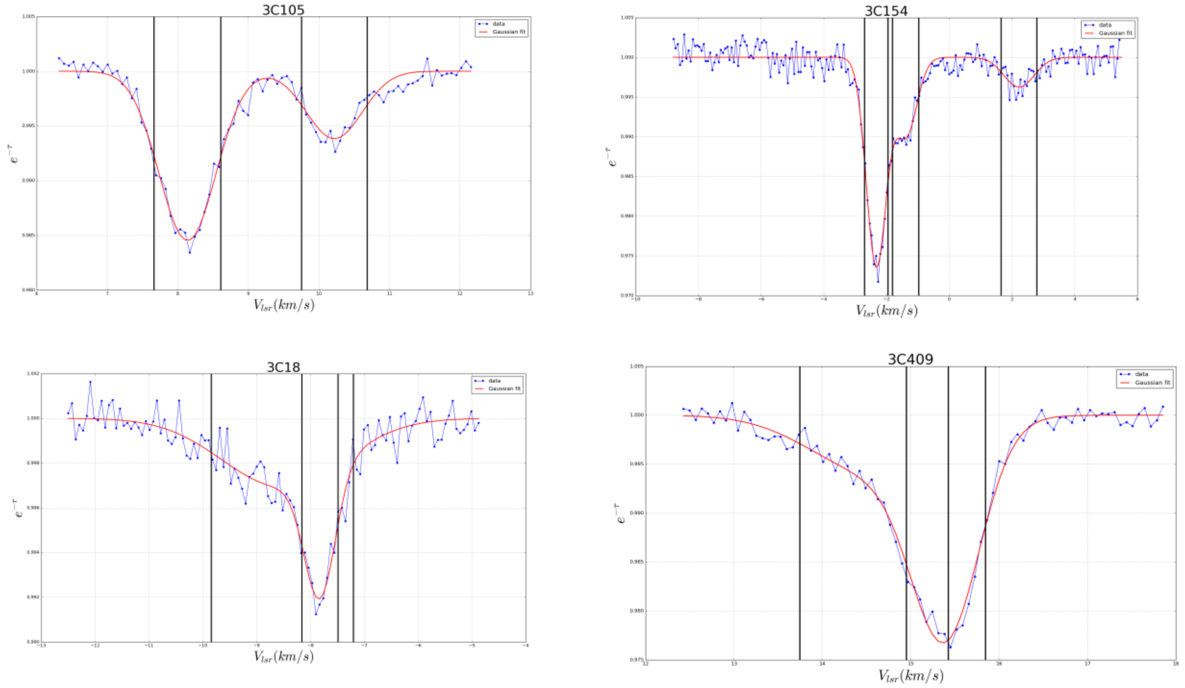

Figure ??: Opacity spectra (blue) and Gaussian fit (red) with the velocity ranges (FWHM) of the peaks.

Step 3: I use the $e^{-\tau}$ (in fact τ values) to calculate the Excitation Temperature T_{ex}. From τ spectra, I derive the 1-σ standard deviation of τ.

I see that the T_{ex} spectra are quite flat within the velocity ranges of the peaks. Thus, I choose ONLY the ranges within the FWHM [v_0-$FWHM/2$, v_0+FWHM/2] to calculate the Mean values of T_{ex} with the criterion that τ within the velocity range must be large than $2\sigma_\tau$.

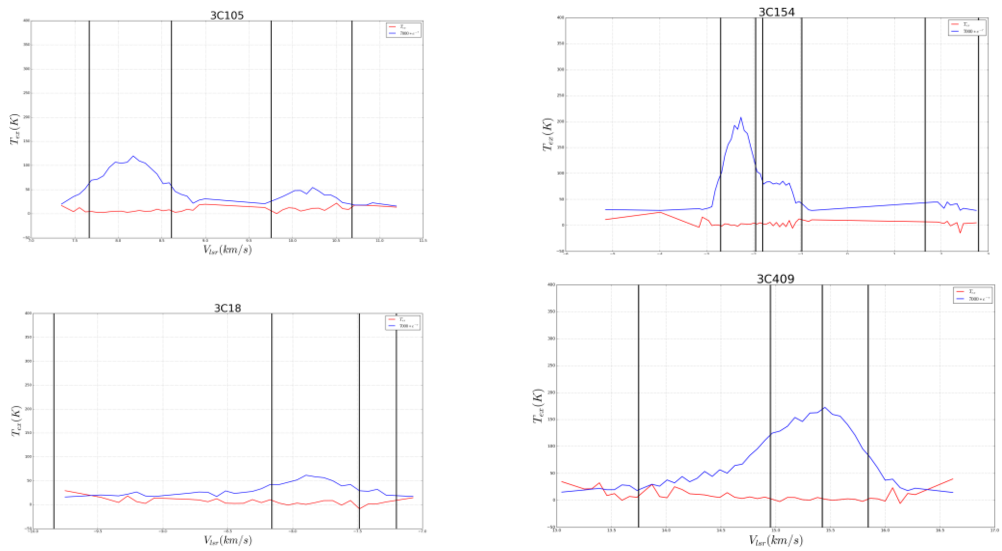

Figure ??: Excitation spectra T_{ex} within the FWHM velocity ranges.

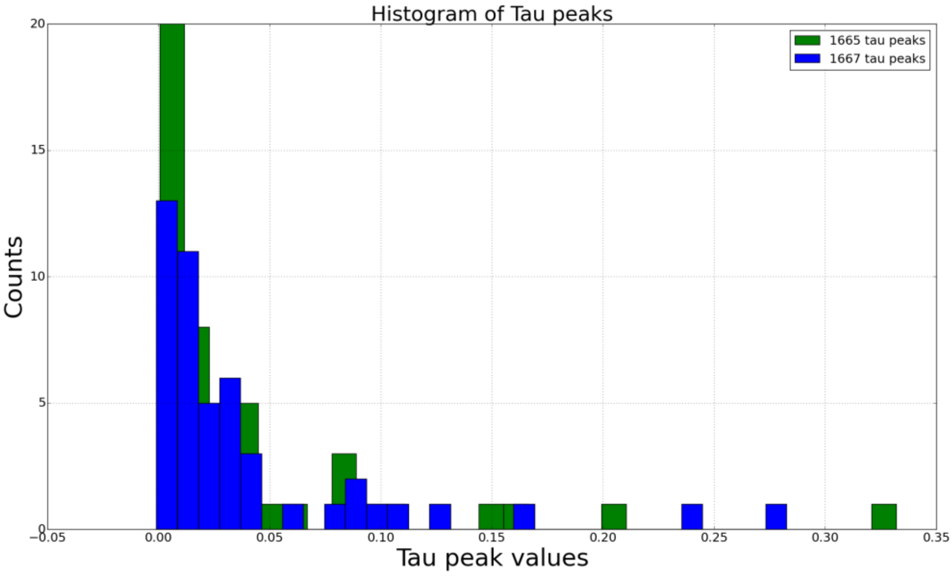

Figure ??: Histogram of Tau peak values

Step 4: Compute the Column Densities of $OH_{1665MHz}$ and $OH_{1667MHz}$ using the following equations:

$N_{OH}[cm^{-2}] = [8\pi\nu k\Sigma g_i/(c^2 g_l A_{ul} h)] * T_{ex}\int\tau_\nu dv$

where:

$c = 3x10^8$ (m/s)

$\Sigma g_i = 5+3+5+3 = 16$

g_l = 3 for 1665MHz line, const = 3.99757843817e14

g_l = 5 for 1667MHz line, const = 2.21841824609e14

And Einstein coefficients are given by Destombes, 1977:

- Line 1612MHz, $A_{ul} = 1.302x10^{-11}$
- Line 1665MHz, $A_{ul} = 7.177x10^{-11}$
- Line 1667MHz, $A_{ul} = 7.778x10^{-11}$
- Line 1720MHz, $A_{ul} = 9.496x10^{-12}$

Results:

1. There exists the main-line anomalies of OH, this means these main lines are not, in general, in Local Thermal Equilibrium.

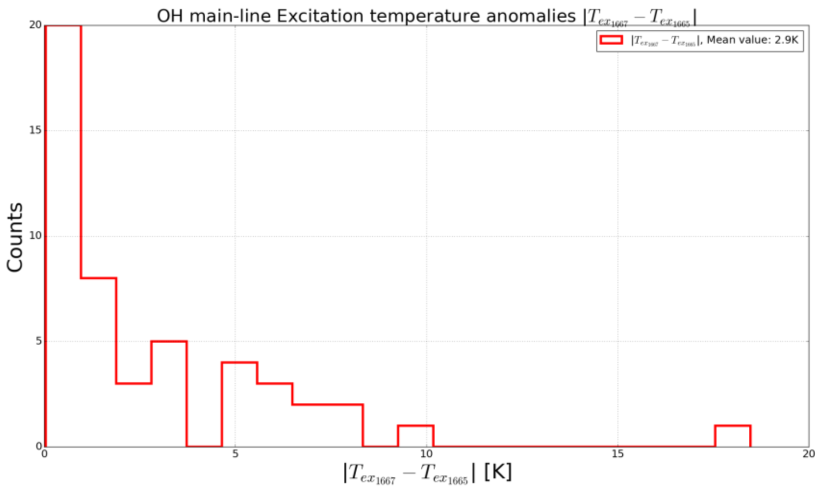

Figure ??: Excitation temperature anomalies of 2 main lines of Hydroxyl. Mean value is 2.9K.

2. The OH/H_2 ratio abundance is $N_{OH} \sim 10^{-7} \cdot N_H$ (cm^{-2}), this result is consistent with other work [1] .

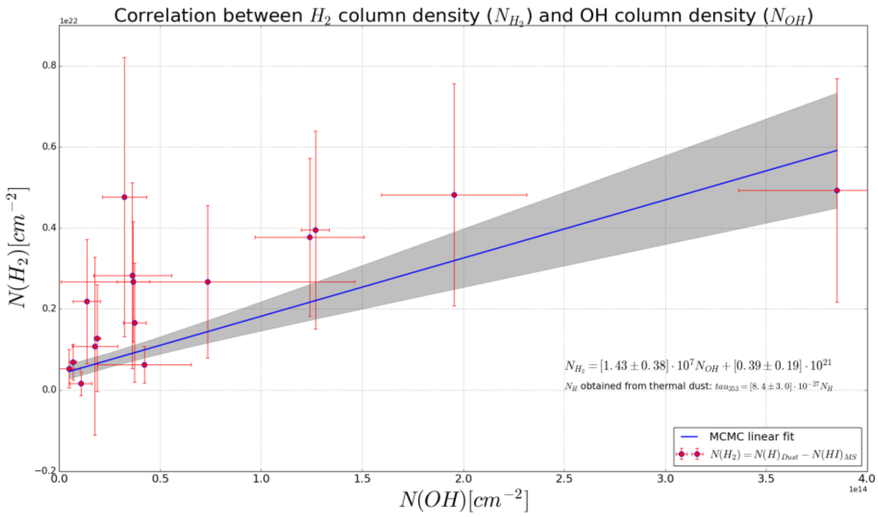

Figure ??: OH/H 2 ratio abundance: $N_{OH}/N_{H2} \sim 10^{-7}$.

3. We find OH in the regions where the values of N_{HI} are high.

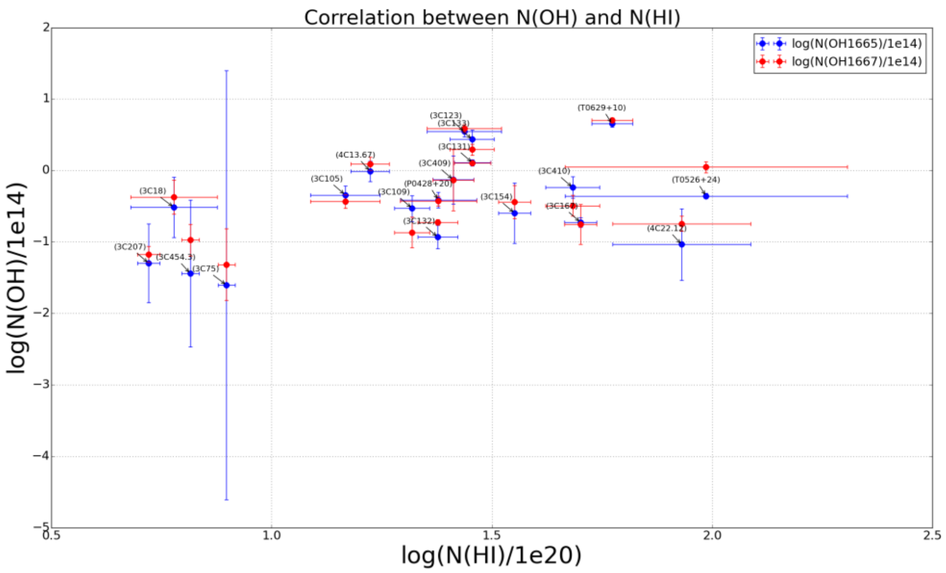

Figure ??: Correlation between N_{OH} and N_{HI}.

References:

[1] Liszt, H., & Lucas, R. 1999, in Highly Redshifted Radio Lines, ed. C. L. Carilli, S. J. E.