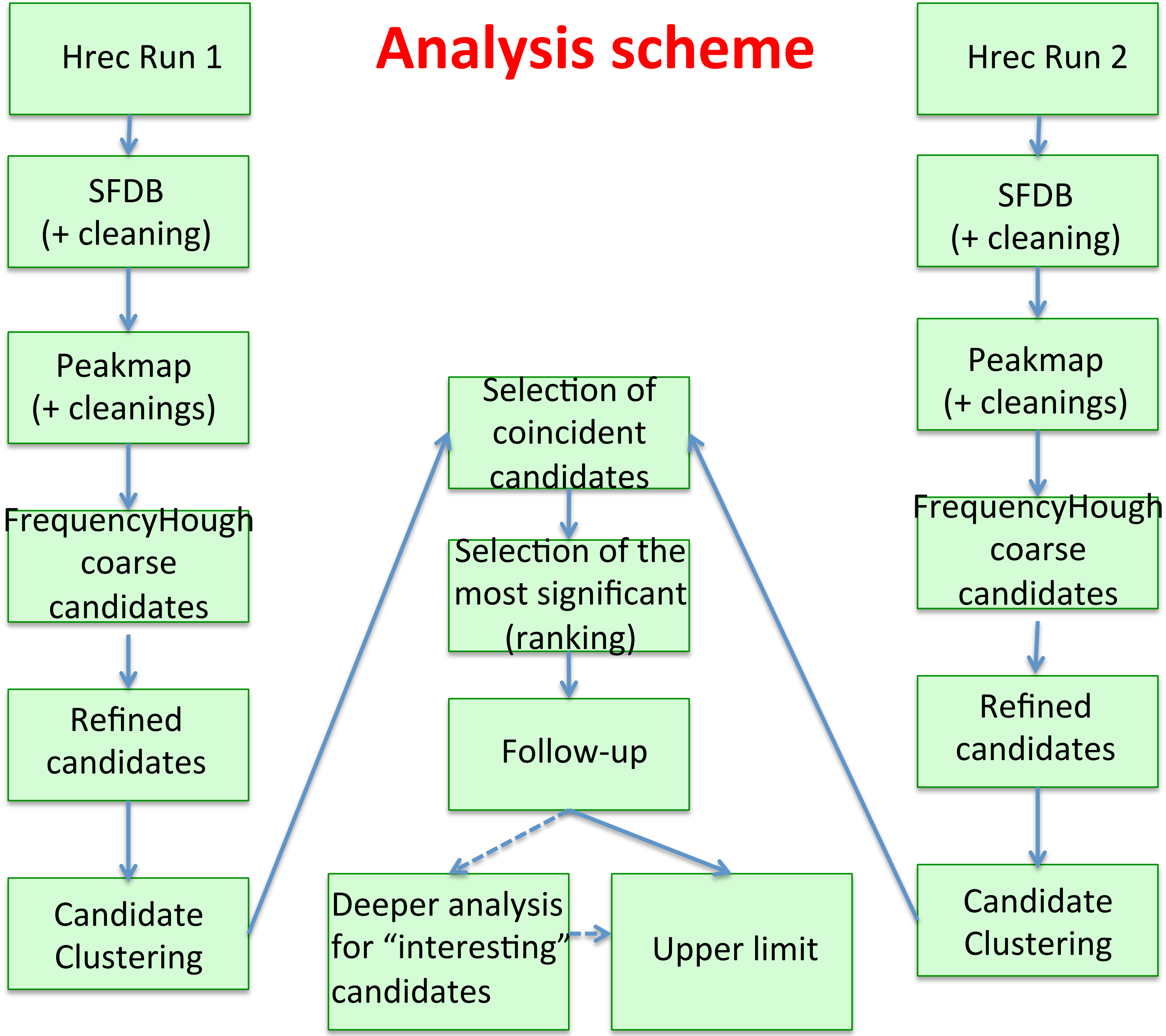
VSR2/4 FrequencyHough all-sky search: an update

Rome CW group

- VSR2 and VSR4 data
- Frequency band [20-128] Hz
- Initial FFT duration 8192 s

- Analysis description (focusing attention on parts never presented to the group)
- Preliminary upper limits (still to be reviewed)
- Paper draft at https://dcc.ligo.org/LIGO-P1500030

Analysis scheme



Hrec Run 1

SFDB
(+ cleaning)

Peakmap
(+ cleanings)

FrequencyHough
coarse
candidates

Refined
candidates

Candidate
Clustering

Hrec Run 2

SFDB
(+ cleaning)

Peakmap
(+ cleanings)

FrequencyHough
coarse
candidates

Refined
candidates

Candidate
Clustering

Selection of
coincident
candidates

Selection of the
most significant
(ranking)

Follow-up

Deeper analysis
for "interesting"
candidates

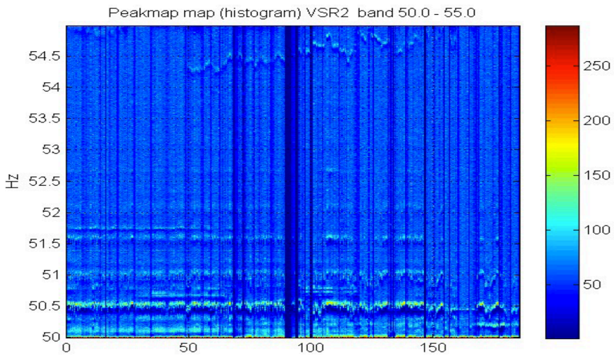
Upper limit

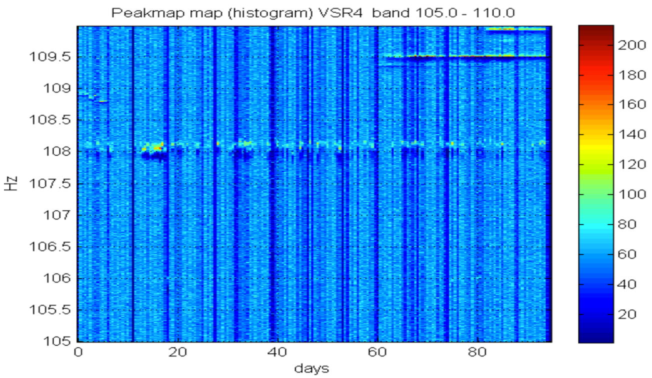
Peakmap

○ From the SFDB the peakmap is built with a threshold of 1.58 on the square root of the equalized spectra.

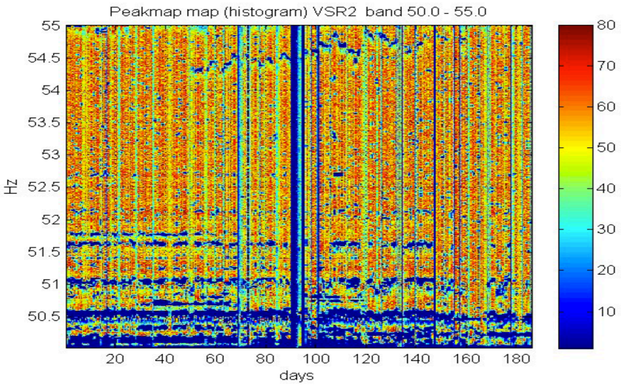
○ Two cleaning steps on the peakmap:
-removal of “wide/wondering lines”;
- removal of persistent lines with constant frequency

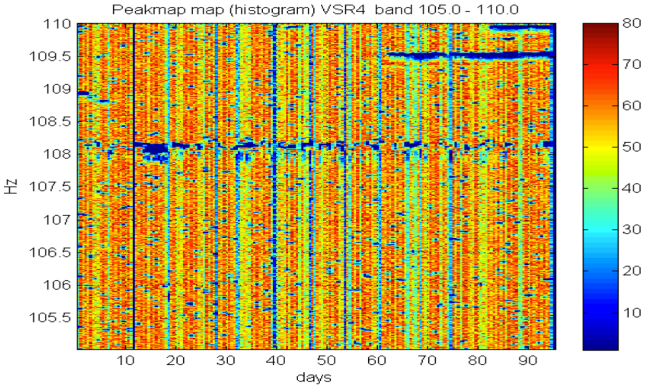
Peakmap “gross histogram” before veto





After veto





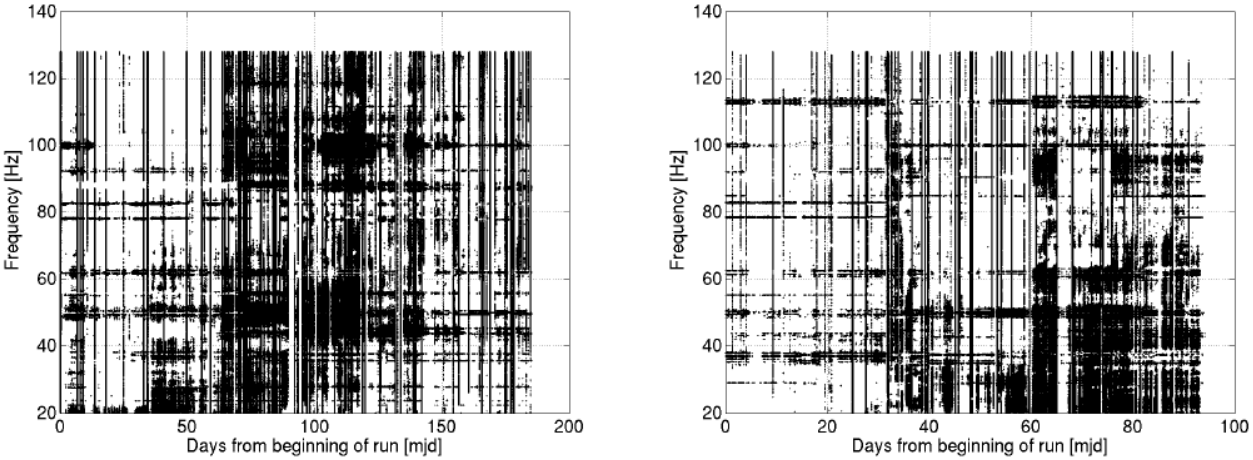


FIG. 4: Time-frequency plot of the peaks removed by the "gross histogram" cleaning procedure for VSR2 (left) and VSR4 (right).

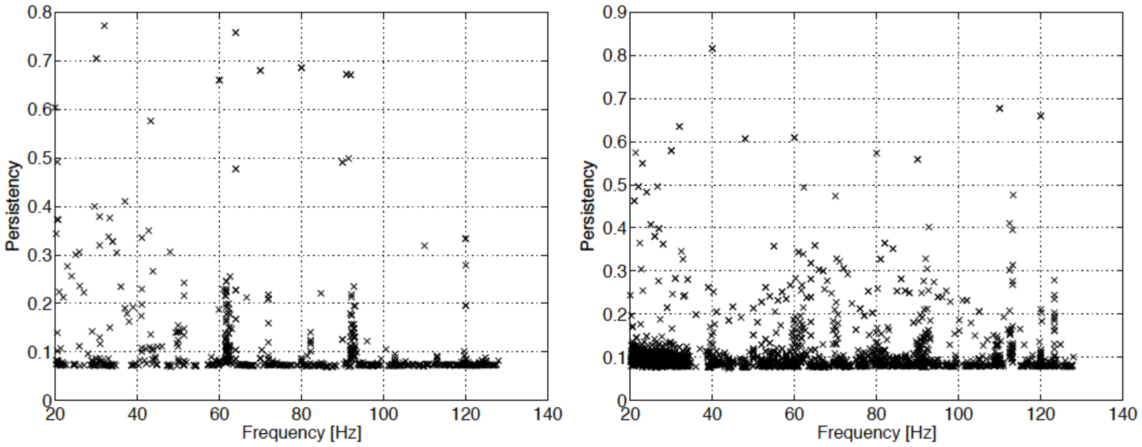
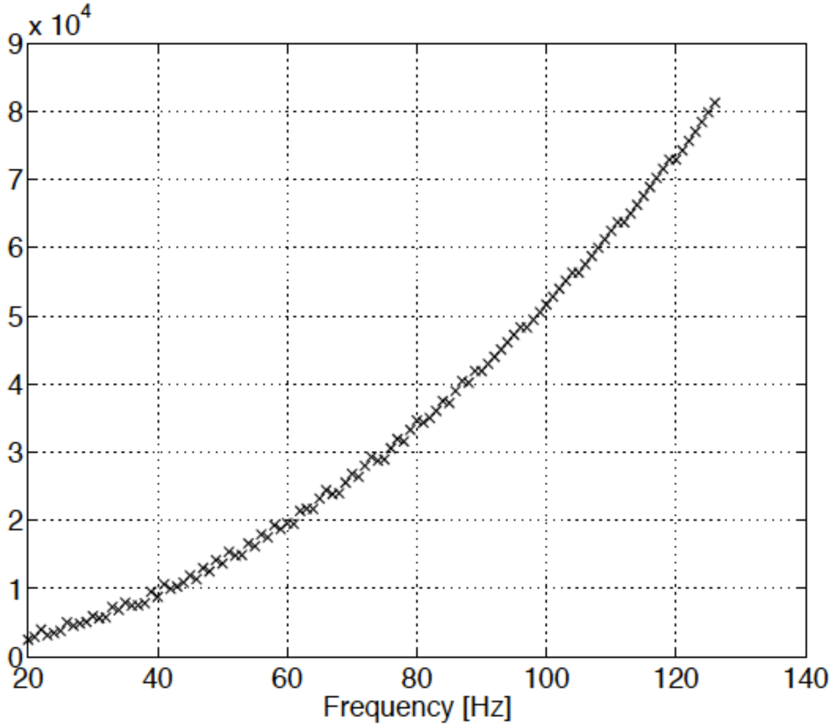


FIG. 6: Lines of constant frequency vetoed on the basis of the persistence, shown in the y-axis, for VSR2 (left) and VSR4 (right). We have removed 710 lines for VSR2 and 1947 lines for VSR4.)

	% after "gross histogram" veto	% after persistency veto	N. of peaks (after vetoes)
VSR2	89.7	89.5	191,771,835
VSR4	86.9	86.5	93,896,752

FrequencyHough Transform

- The cleaned peakmap is corrected for the Doppler effect for each sky direction, by shifting each frequency bins.
- A *coarse* sky grid is used at this stage.



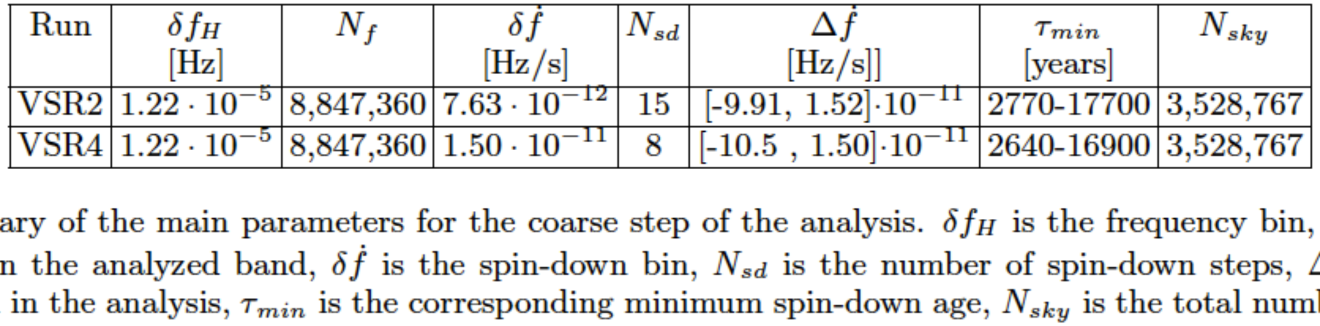
~3.5E6 points in total

FIG. 3: Number of sky patches in each 1 Hz band, from 20 Hz up to 128 Hz. The frequency on x-axis indicates the beginning frequency of each band. The number of patches increases with the square of the frequency, and is fixed by the highest frequency of each 1 Hz band.

- Each corrected peakmap is given in input to the FrequencyHough stage.

- Adaptive implementation is used

- We use an over-resolved grid in frequency (x10) and the "natural" spin-down step



Run	δf_H [Hz]	N_f	$\delta \dot{f}$ [Hz/s]	N_{sd}	$\Delta \dot{f}$ [Hz/s]]	τ_{min} [years]	N_{sky}
VSR2	$1.22 \cdot 10^{-5}$	8,847,360	$7.63 \cdot 10^{-12}$	15	$[-9.91,\ 1.52]\cdot 10^{-11}$	2770-17700	3,528,767
VSR4	$1.22 \cdot 10^{-5}$	8,847,360	$1.50 \cdot 10^{-11}$	8	$[-10.5\ ,\ 1.50]\cdot 10^{-11}$	2640-16900	3,528,767

TABLE II: Summary of the main parameters for the coarse step of the analysis. δf_H is the frequency bin, N_f is the number of frequency bin in the analyzed band, $\delta \dot{f}$ is the spin-down bin, N_{sd} is the number of spin-down steps, $\Delta \dot{f}$ is the range of spin-down covered in the analysis, τ_{min} is the corresponding minimum spin-down age, N_{sky} is the total number of sky patches.

Candidate selection

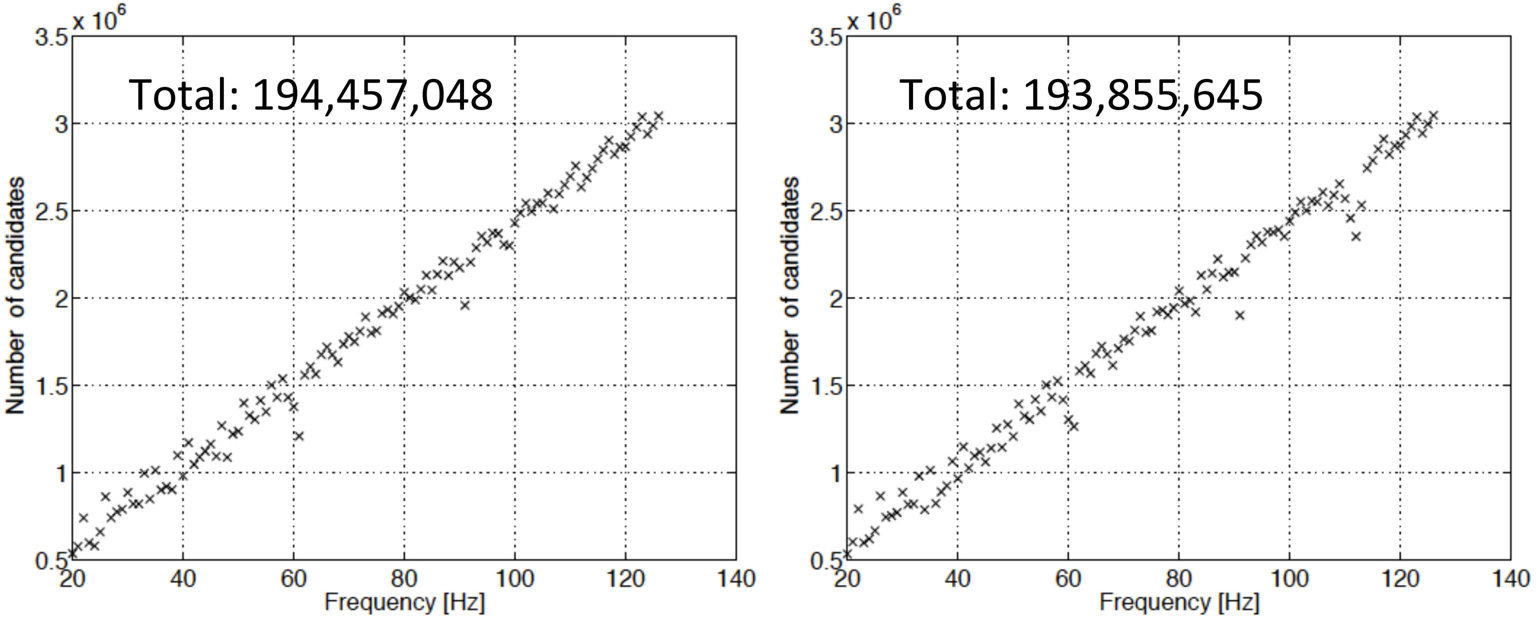
- We want to have a constant number of coincidences as a function of the frequency (i.e. the number of candidates in each run must increase linearly with the frequency);
- We fix the total number of candidates, 10^8 for each run (we would be able to handle up to $O(10^9)$ candidates);
- Within each 1Hz band we choose to have the same number of candidates for each sky patch, $N^{sky}_{cand;i}$ (i=20,… 127);

- Moreover, for each sky cell we divide the i-th band into $N^{sky}_{cand;i}$ sub-bands and select the strongest candidate in each of them (highest Hough amplitude);

- In this way the probability of being blinded by the presence of disturbances within a band is reduced

- The second strongest candidate is also selected, if well separated in frequency from the first;

- On each selected candidate a *refined* search is done using a refined grid around it (still using the FH): for each coarse candidate the strongest refined candidate is chosen.



Total: 194,457,048

Total: 193,855,645

FIG. 8: Number of candidates selected in each 1 Hz band for VSR2 (left) and VSR4 (right).

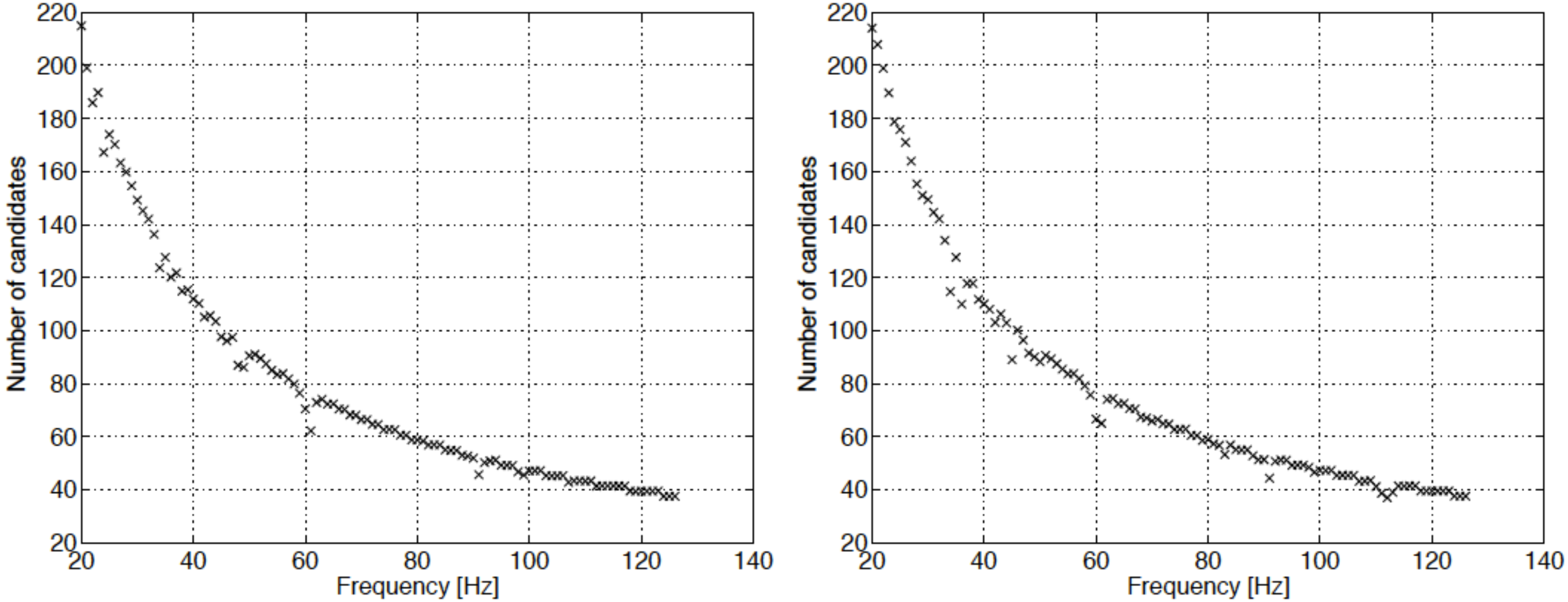


FIG. 9: Number of candidates selected in each 1 Hz band and in each sky patch for VSR2 (left) and VSR4 (right).

Candidate clustering and coincidences

- For computational efficiency reasons, candidates of the two runs are *clustered* before making coincidences.

- A cluster is a collection of candidates such that each of them has a distance d in the parameter space less than $d_{clust}=2$ from at least one other candidate of the same collection.

$$d = \|\vec{c}_1 - \vec{c}_2\| = \sqrt{k_\lambda^2 + k_\beta^2 + k_f^2 + k_{\dot{f}}^2}$$

$$k_\lambda = \frac{|\lambda_2 - \lambda_1|}{\delta\lambda}$$

And similarly for the other parameters

$$\delta\lambda = \frac{d\lambda_1 + d\lambda_2}{2}$$

- We find 94,153,784 clusters for VSR2 and 38,953,404 for VSR4.

- Coincident candidates are selected in the following way:

- determine coincident cluster pairs (i.e. those having at least a pair of candidates with distance less than $d_{coin}=2$).

- On the basis of global cluster parameters, pairs of clusters which cannot be in coincidence are not taken into account (thus reducing the computational effort).

- For each pair of coincident clusters take as coincident candidates those with the **smallest distance** (i.e. we select one pair of coincident candidates for each pair of coincident clusters).

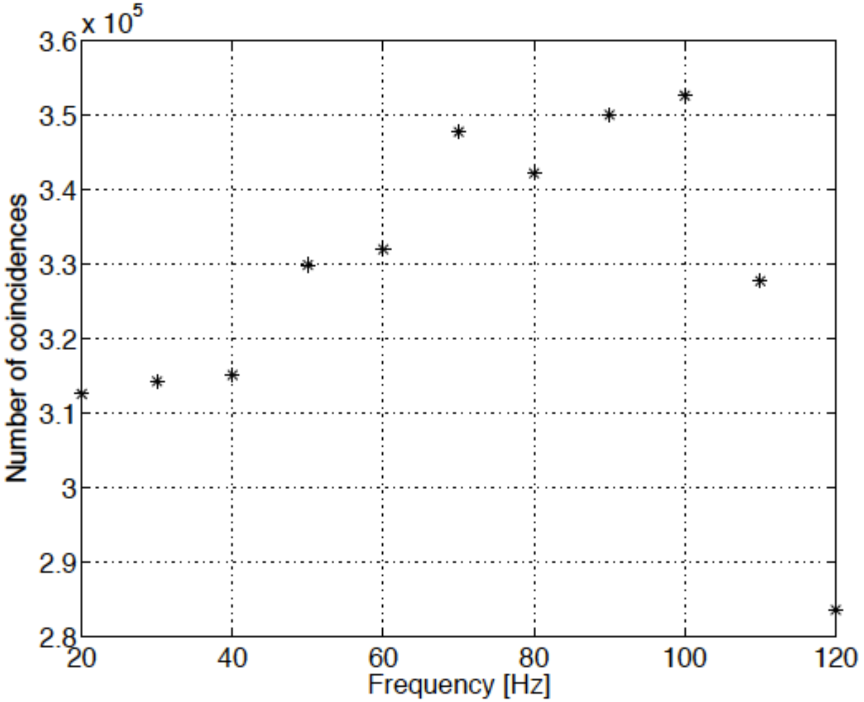


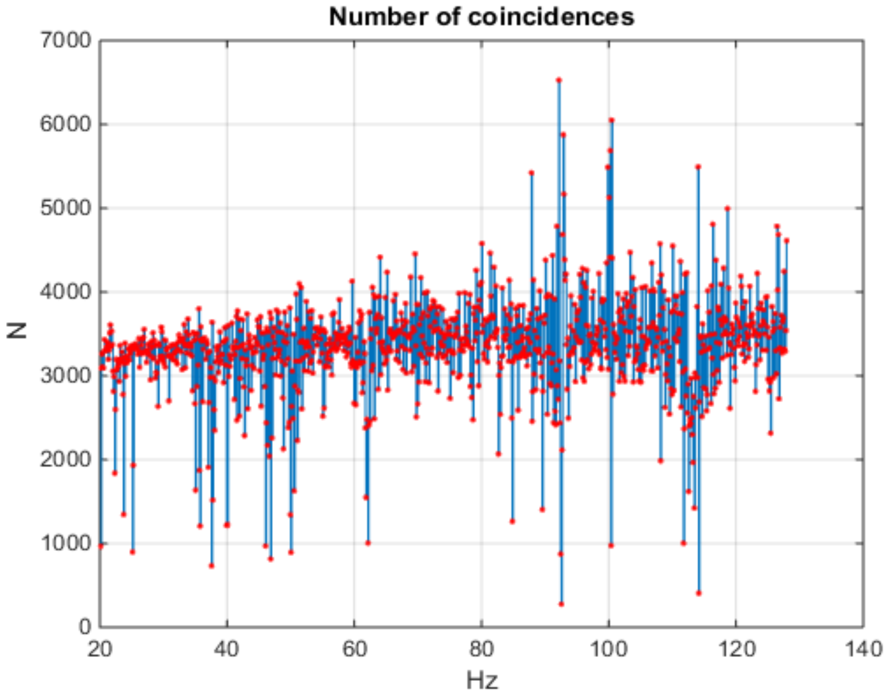
FIG. 11: Number of coincidences, divided into 10 Hz wide bands, as a function of the beginning frequency of each band

Total number of coincidences: 3,608,192

Coincident candidates are subject to a *ranking* procedure in order to select the most significant.

Ranking

- Coincident candidates are divided in bands of 0.1Hz



- In each band (and for each detector): order candidates in descending order of the Hough amplitude

- Assign a rank to each of them: from 1/N to the highest to 1 to the smallest (N: number of candidates in the band)

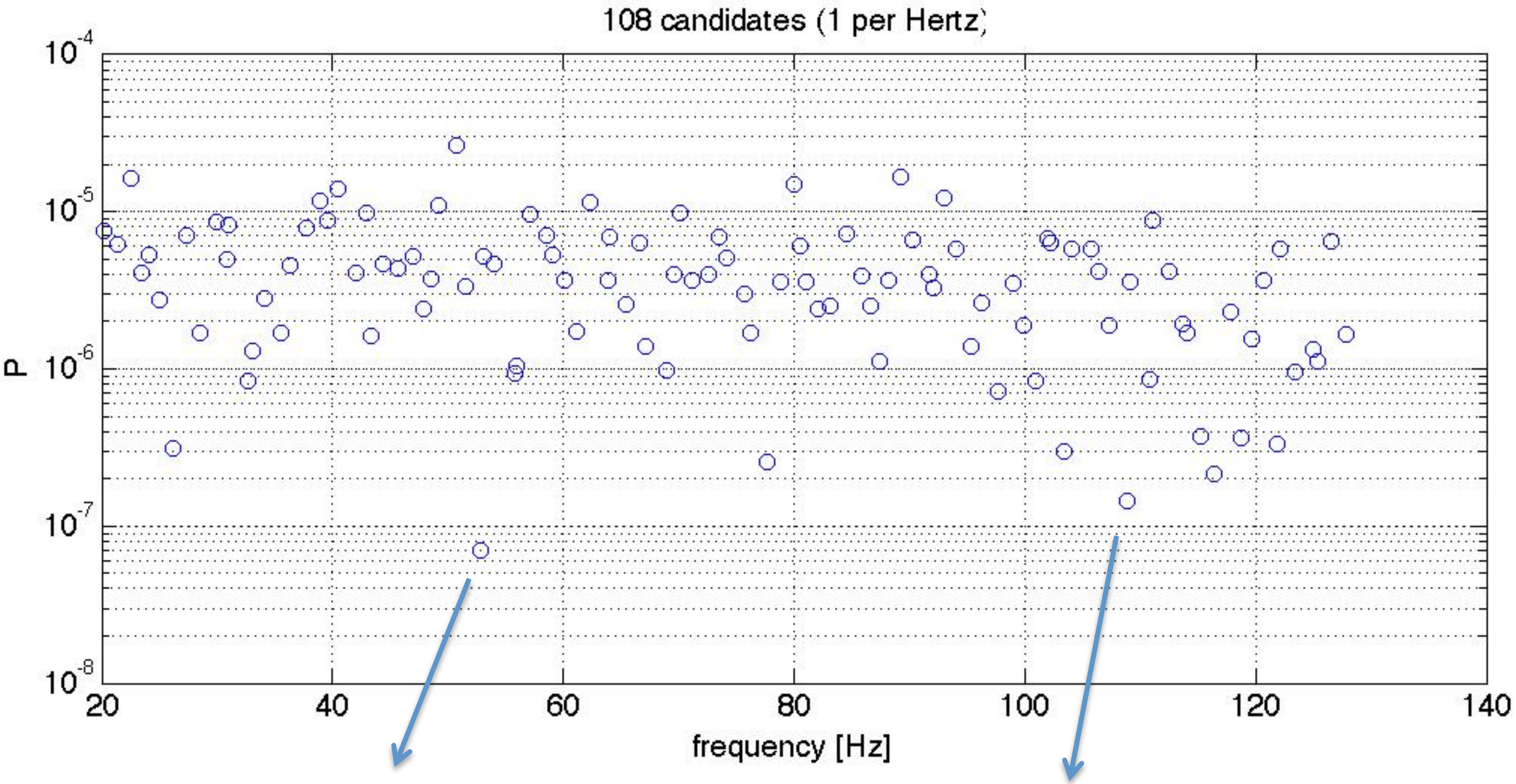
- Make the product of the ranks of coincident candidates.

  This can be seen as the probability of having a value of the product of the ranks equal or larger than the value we are considering (possible range: 1/N^2 → 1)

- Select the most significant candidate (i.e. that with the smallest rank product) for each 0.1Hz band: 1080 candidates.

- Among these select the most significant in each 1Hz band: 108 candidates. These are subject to the follow-up.

- Again, we do not want to be overwhelmed by some large disturbance.

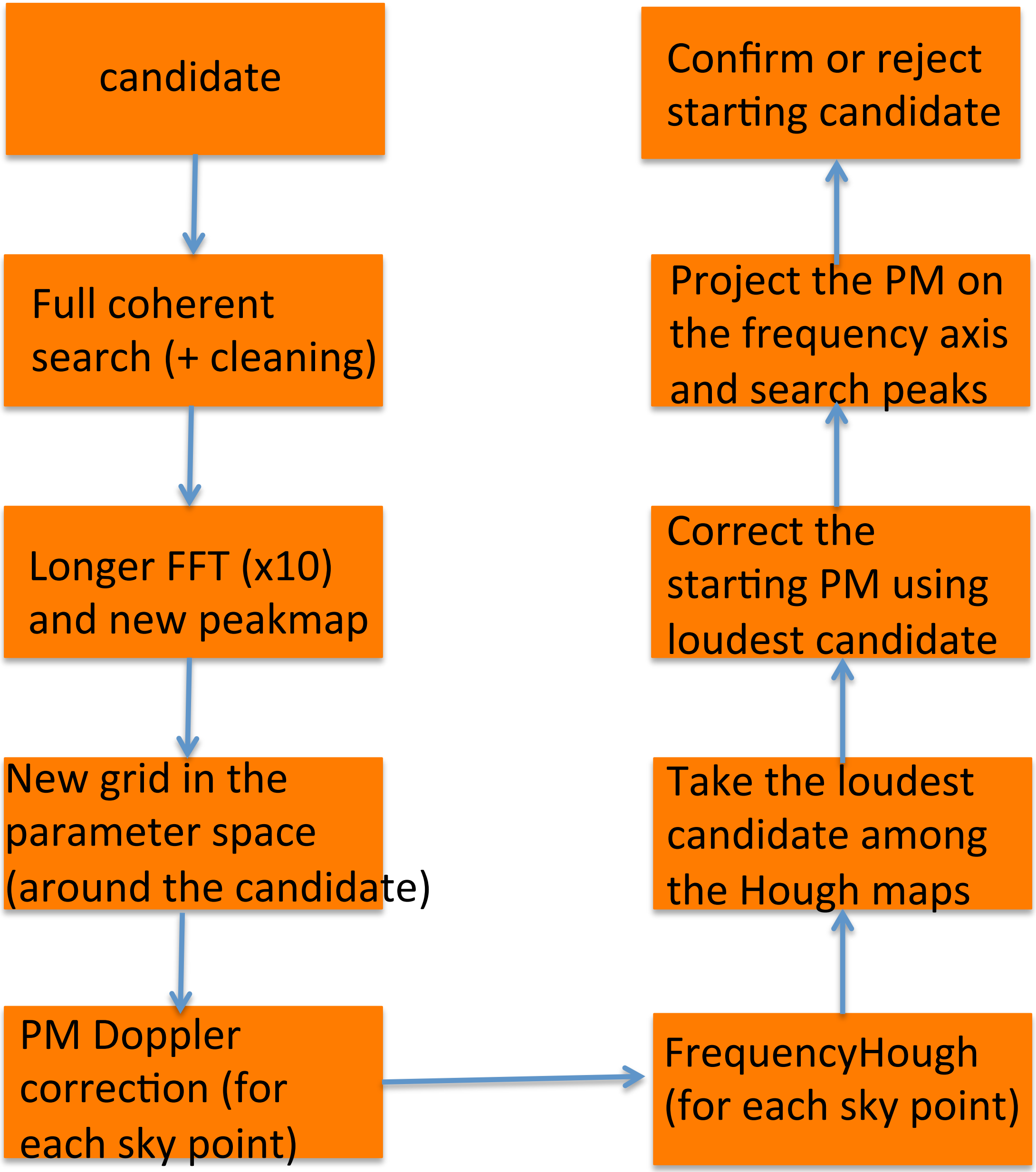


pulsar5

pulsar3

- The follow-up consists of various steps.

Follow-up



candidate

Full coherent search (+ cleaning)

Longer FFT (x10) and new peakmap

New grid in the parameter space (around the candidate)

PM Doppler correction (for each sky point)

FrequencyHough (for each sky point)

Take the loudest candidate among the Hough maps

Correct the starting PM using loudest candidate

Project the PM on the frequency axis and search peaks

Confirm or reject starting candidate

1. **Full coherent search** using candidate parameters (using the same pipeline used for targeted searches).

We start from the SFDB (FFT duration: 8192 seconds).

The output is a down-sampled (1Hz) time-series corrected for Doppler effect, spin-down, Einstein delay.

A further step of outliers removal is done.

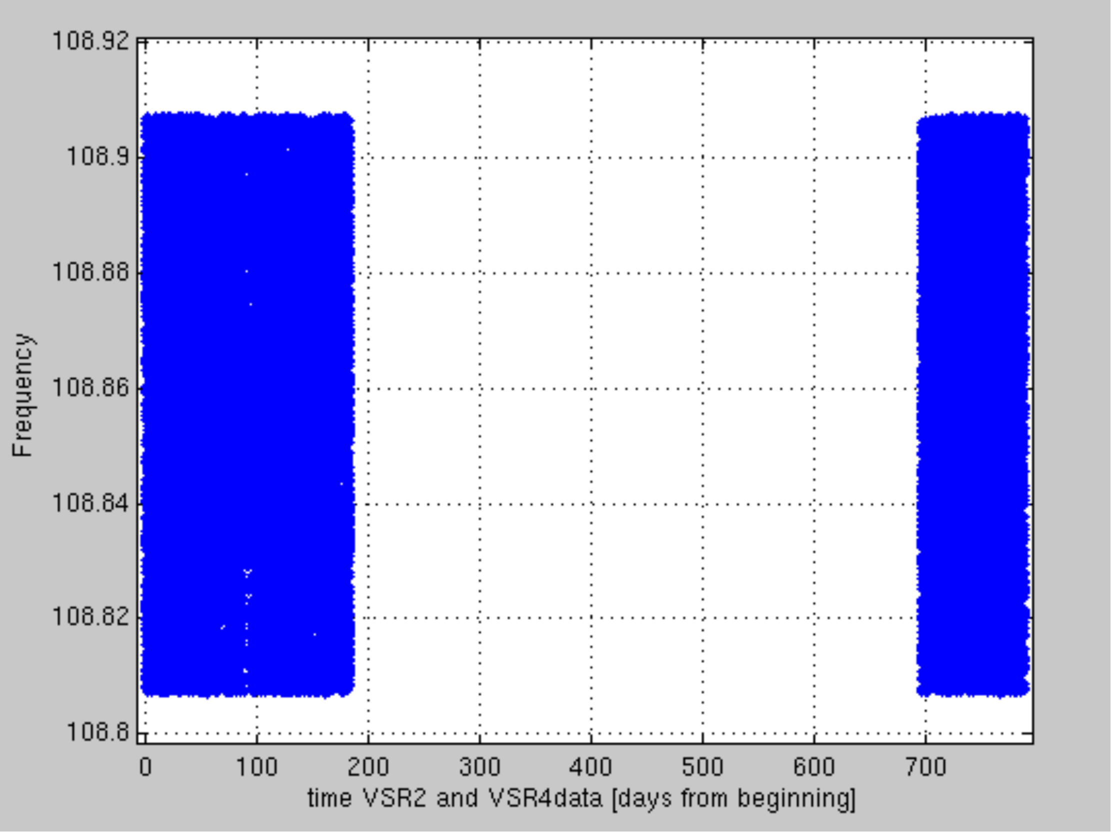
2. **From the corrected data compute new, longer, FFTs and a new peakmap**. Enhancement factor: 10 → T_{FFT}=81920 seconds

Done separately for VSR2 and VSR4.
We consider ±0.05 Hz around candidate frequency.

We select peaks with a threshold of 2.34 (while we used 1.58 in the first step).

The resulting probability of selecting a noise peak is less than 4E-3, i.e. ~20 times smaller than with the initial threshold.

3**. A single peakmap is built** from the two separate peakmaps (one from VSR2 and one from VSR4).



This peakmap is the input of a new FrequencyHough step.

4. **A new grid is built** in the parameter space (around the candidate)

Frequency resolution: Df = 1/(10*Tfft) = 1.22* 10^{-6} Hz
The covered frequency band is 0.1Hz around candidate frequency.

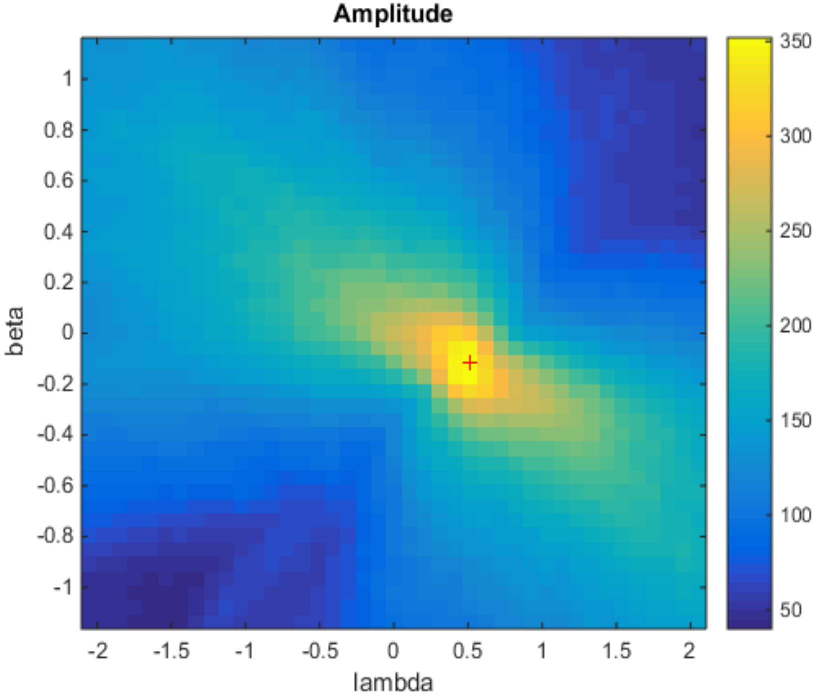
Spin-down resolution: Dfdot=1.5 10^{-12} Hz/s.
It covers ±1 coarse bin around candidate spin-down

Sky: 41*41=1681 patches around candidate position.
They cover ±0.75 times the coarse patch.

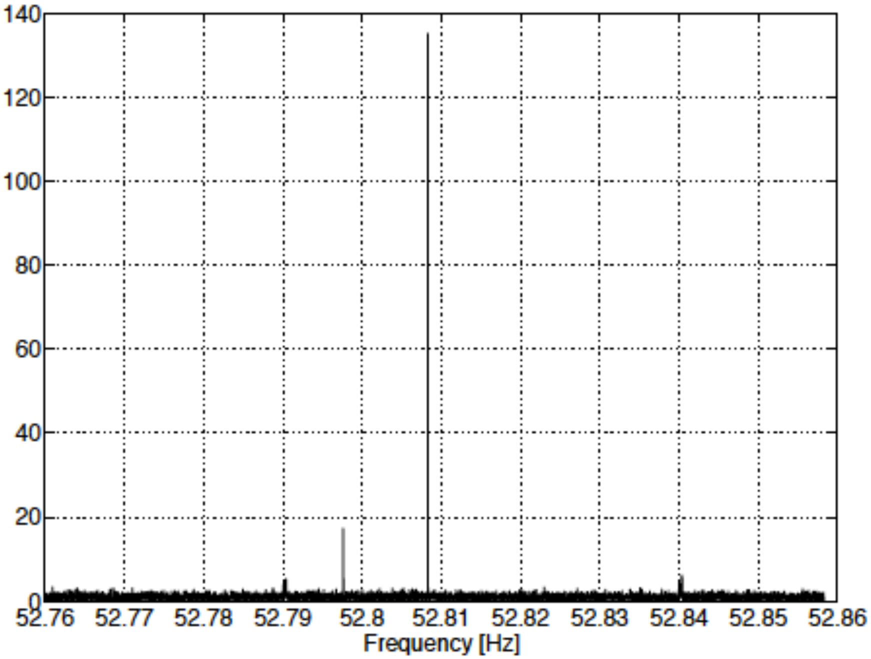
5. For each sky position in the grid: **the peakmap is Doppler corrected** and **a frequencyHough transform is computed**: 1681 transforms.

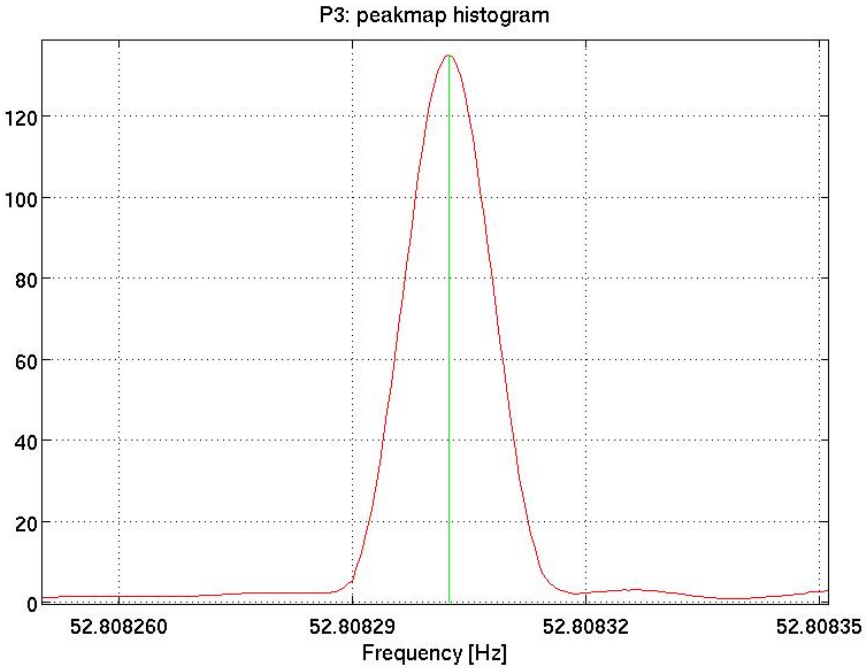
Then, the loudest peak among all the 1681 Hough maps is selected.

Plot of the maximum of the Hough maps for each point in the sky (around P5)



6. **Correct the initial peakmap** using the signal parameters estimated in the last step and **project it** on the frequency axis.



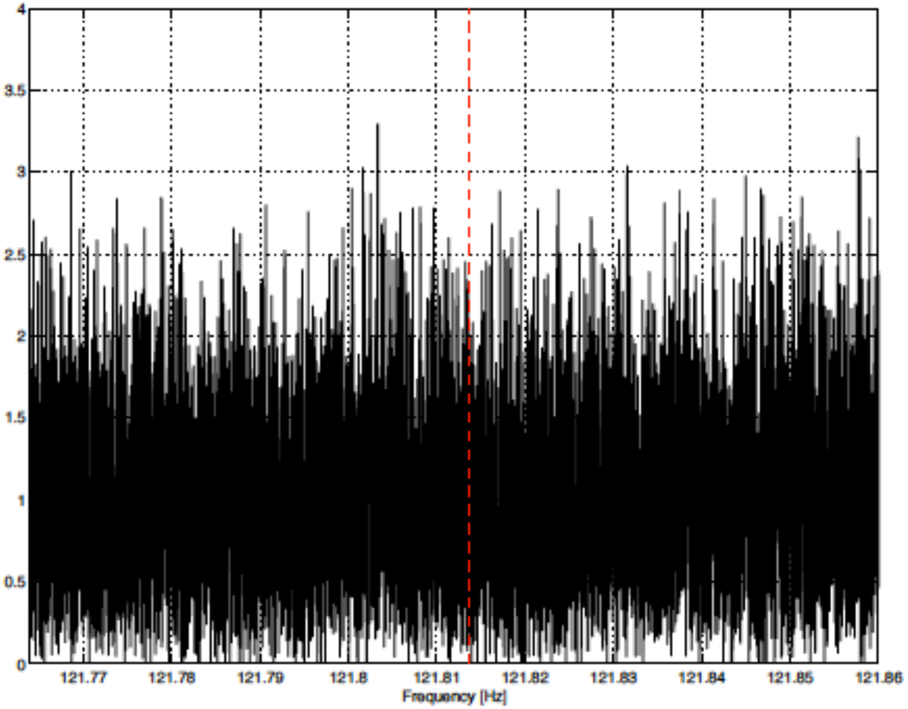


Finally, we **search for significant peaks** in a band of 2 coarse bins around the initial reference frequency.

- The rest of the band is used to estimate the "background" (408 intervals).

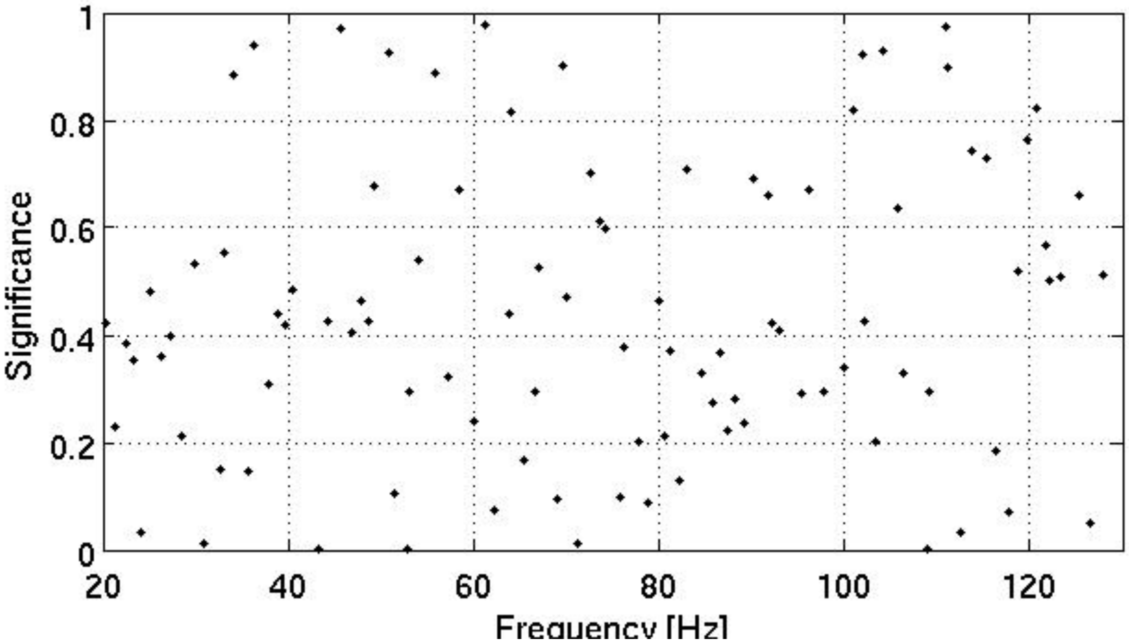
- Given the loudest peak in the central band, we compute the fraction of background sub-bands in which the loudest peak is larger than that in the search band .

- If this number is smaller than 1%, the candidate is considered "interesting" and will be further analyzed.



Peakmap histogram for candidate at 121.81Hz

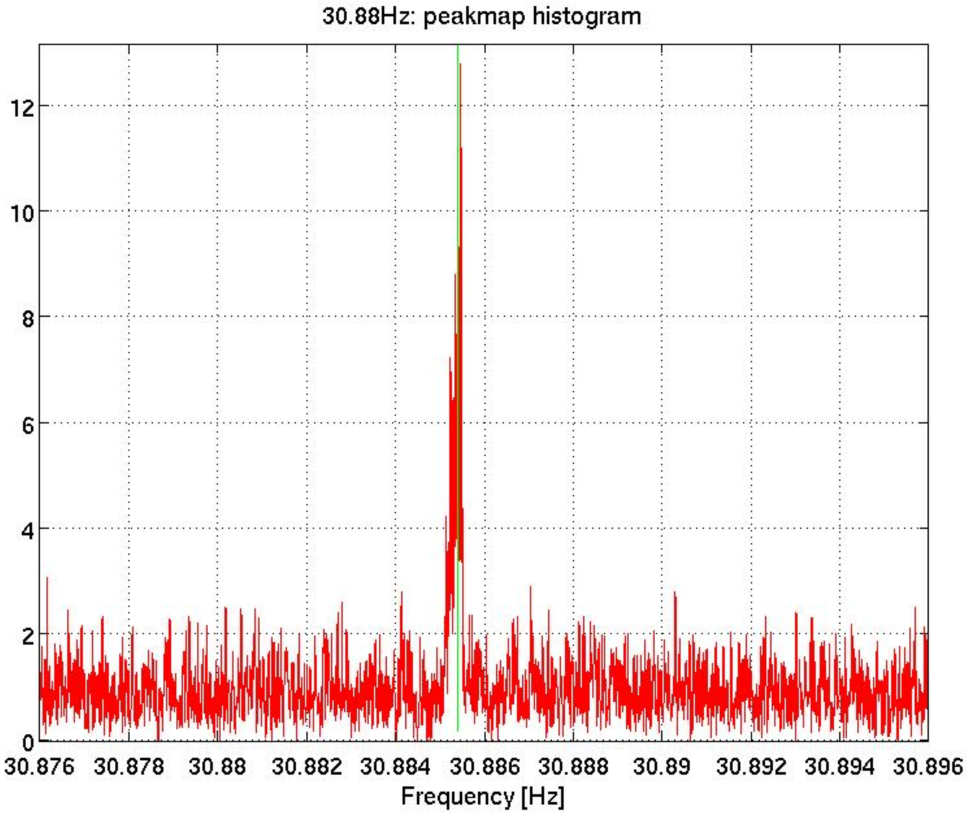
- Follow-up still to be done for 10 candidates

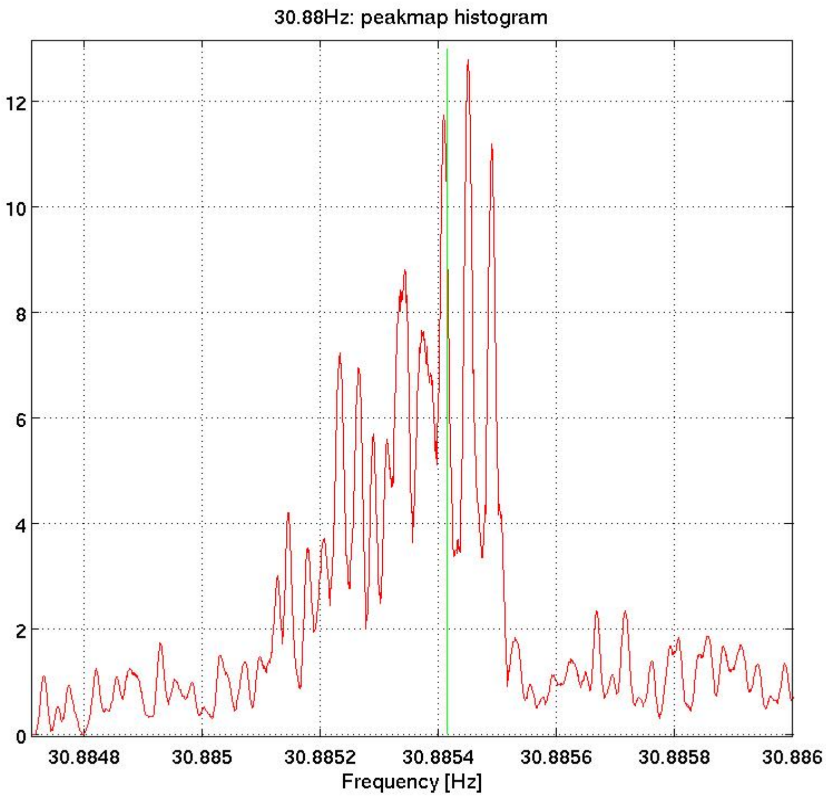


- Apart from the two HI there are 3 “interesting” outliers at 30.88Hz, 43.30Hz and 71.20Hz.

- We have analyzed these in some more detail, and they appear to be due to noise.

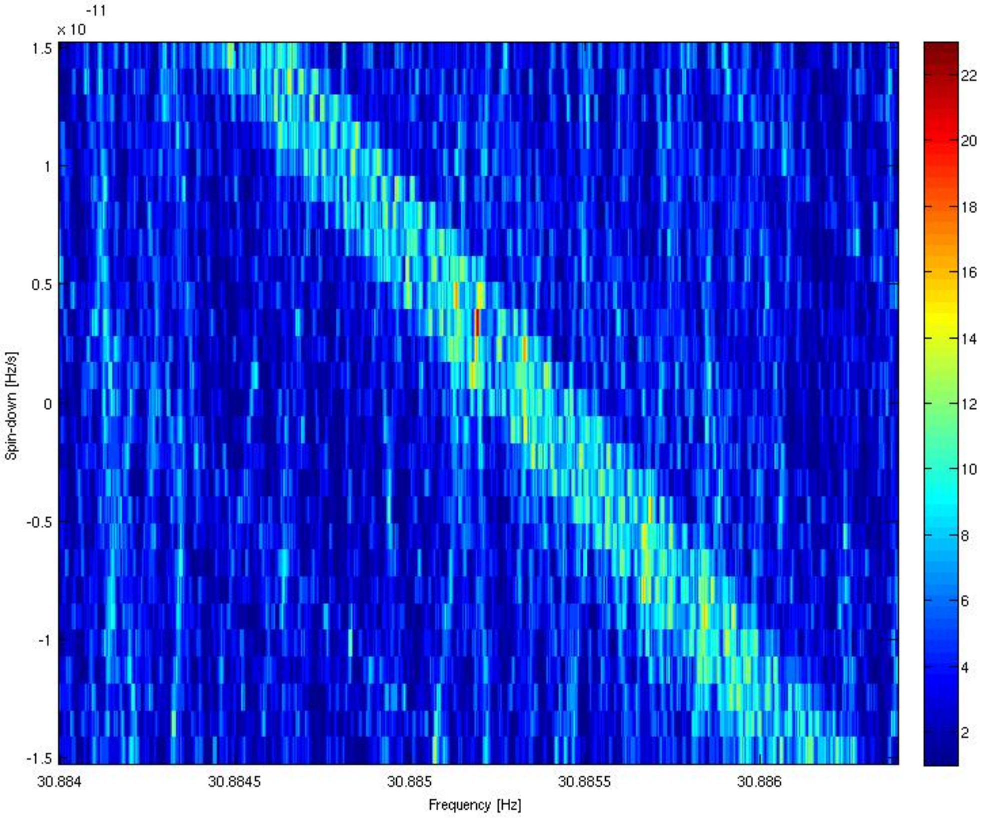
Candidate 30.88Hz





p	f	fdot	λ	β
0.011	30.88541	-9.19E-12	235.40	50.99

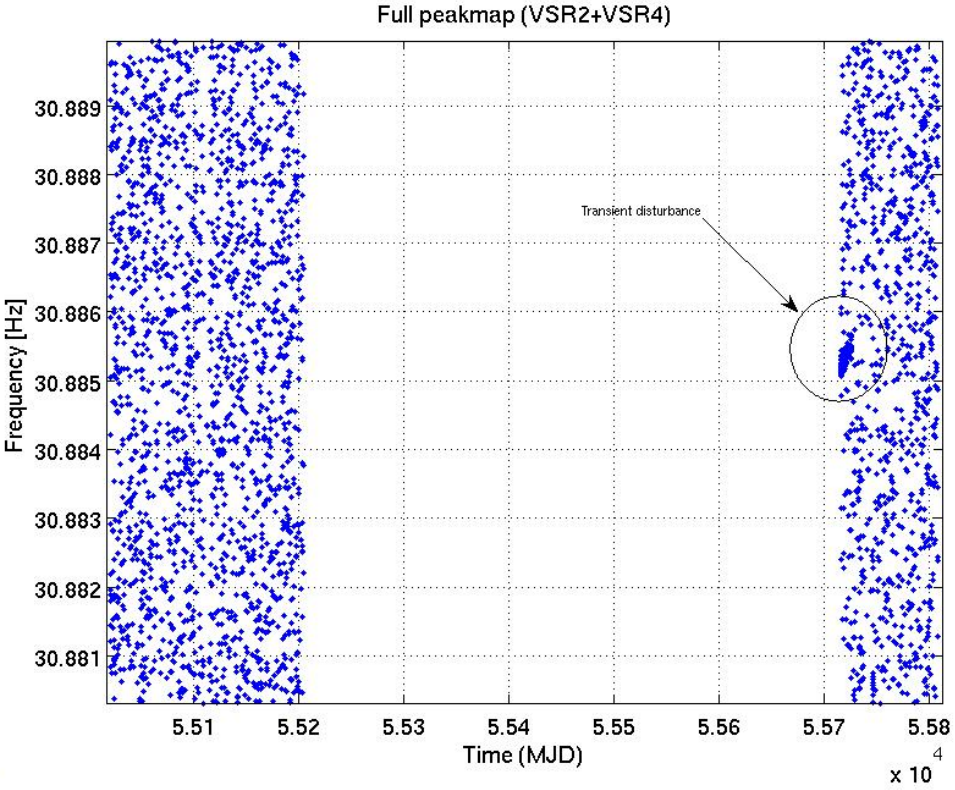
Looking at the Hough map:

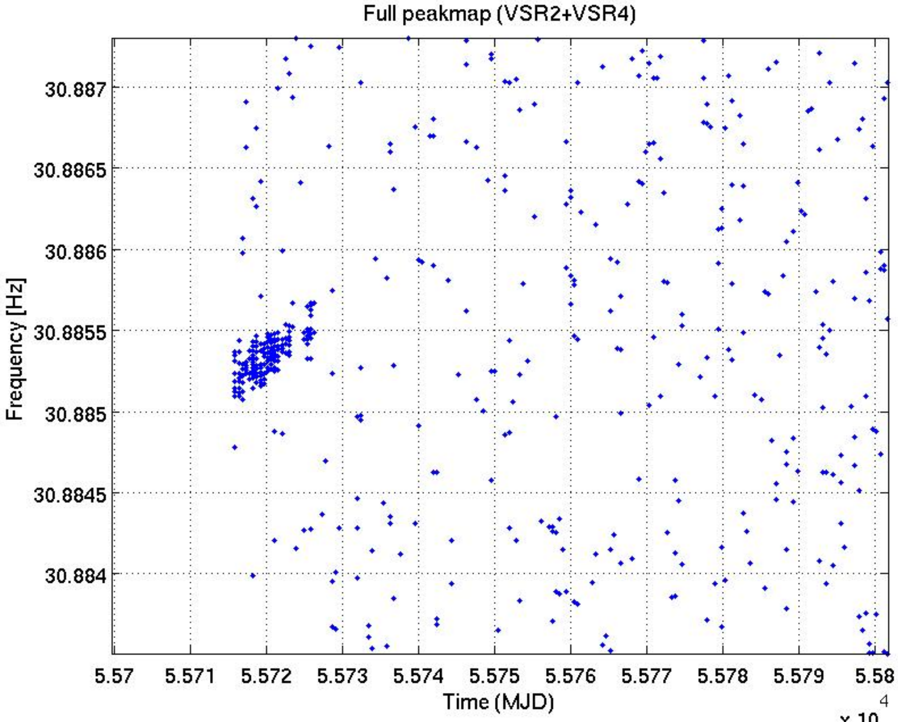


It is a single ‘large’ strip: we expect this is produced by some kind of ‘temporary’ disturbance in only one of the two runs.

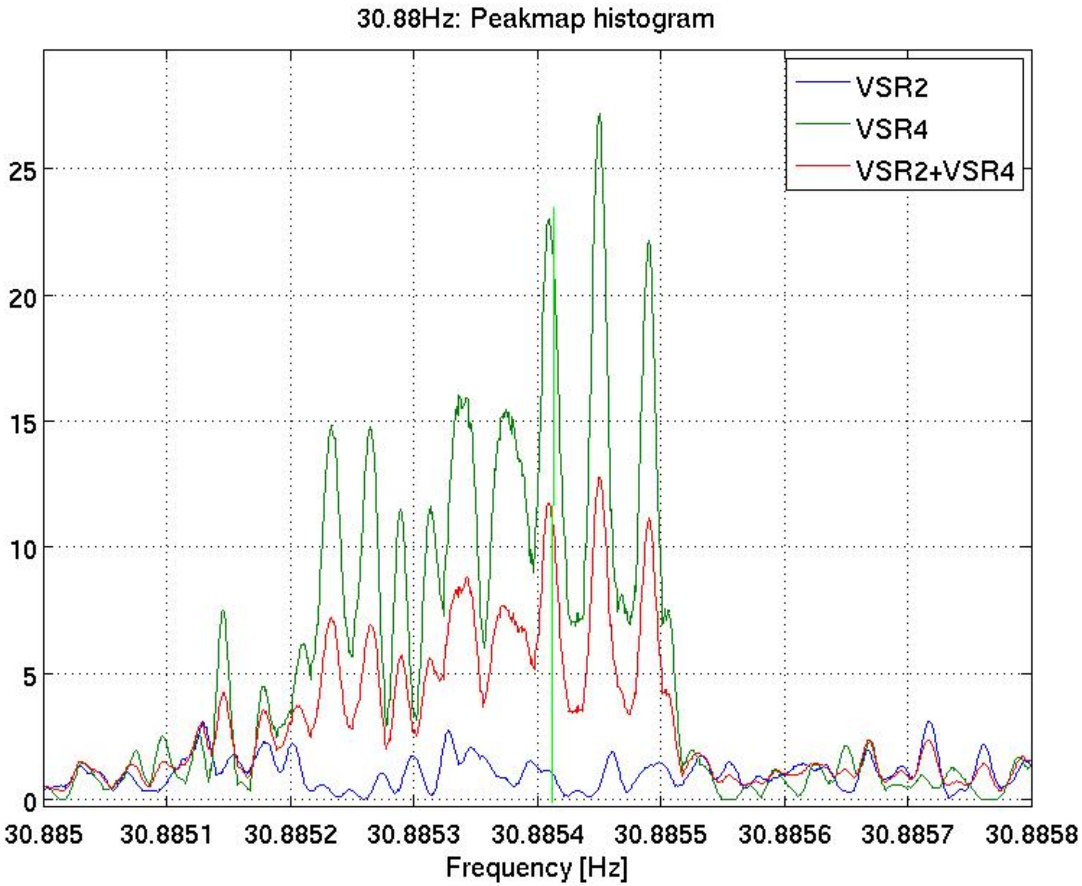
Looking at the peakmap:

There is an excess of points at the right frequencies over about 10 days at the beginning of VSR4.



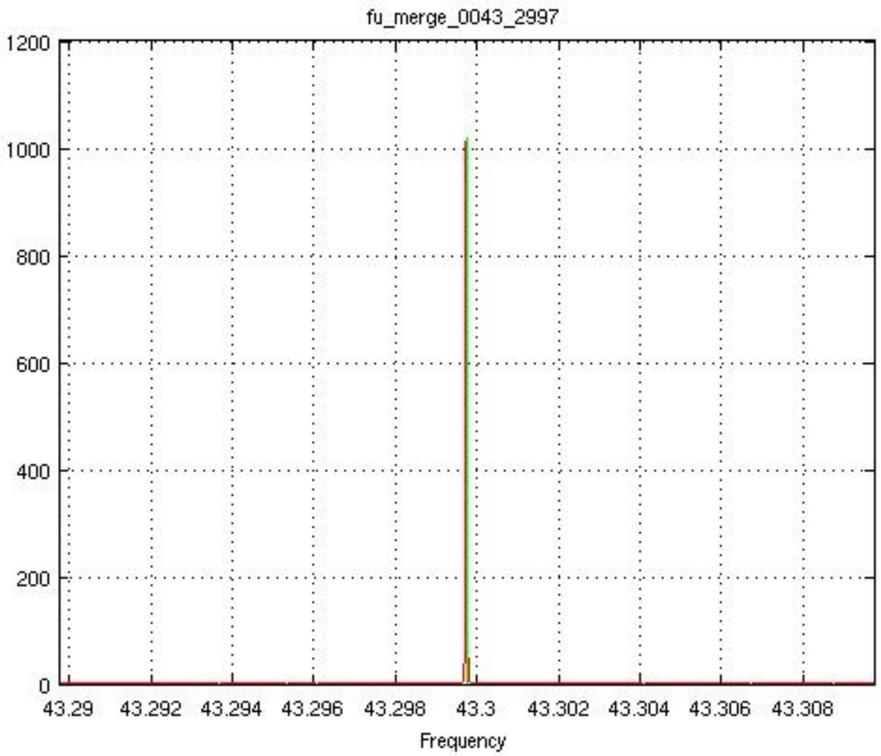


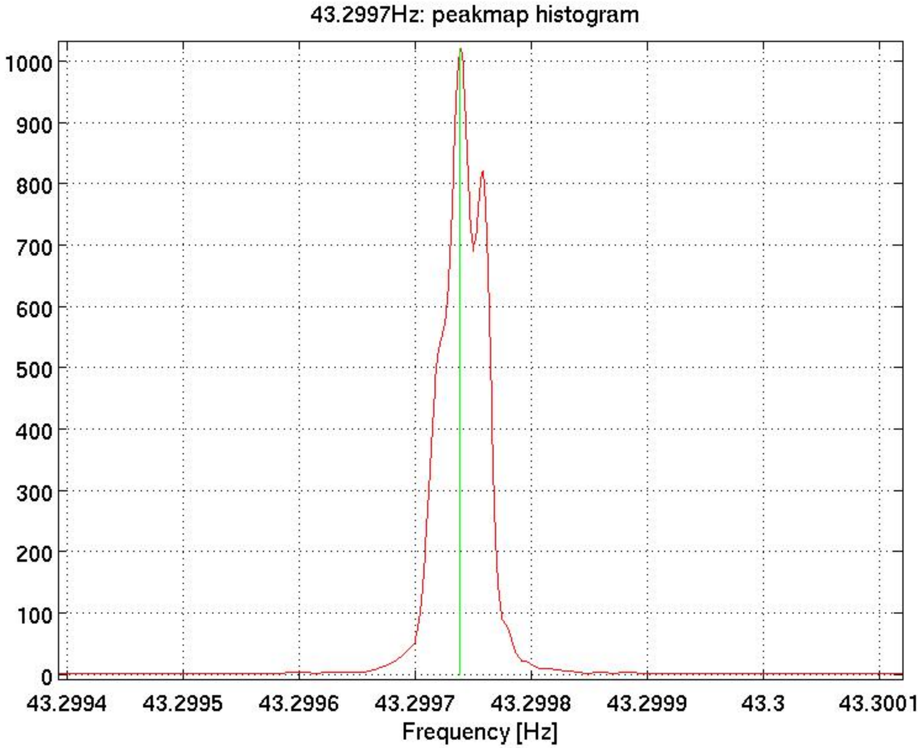
Indeed, by looking at the two runs separately we have a significant candidate only in VSR4:



We conclude this candidate is due to noise.

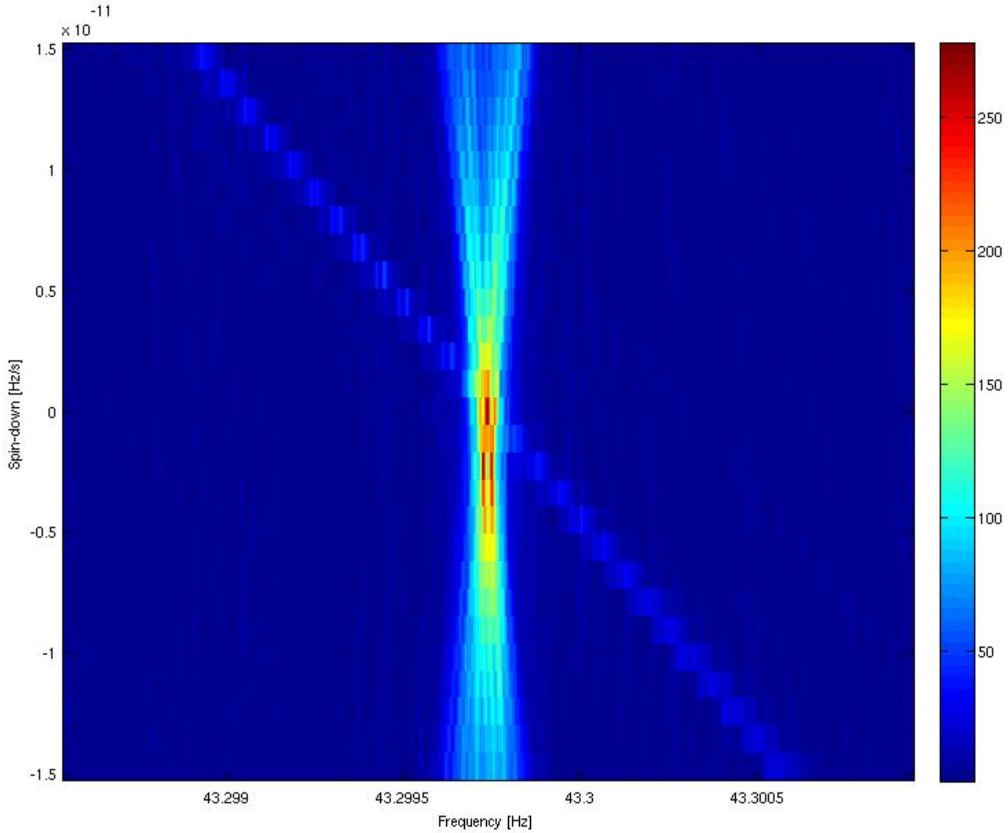
Candidate 43.30Hz





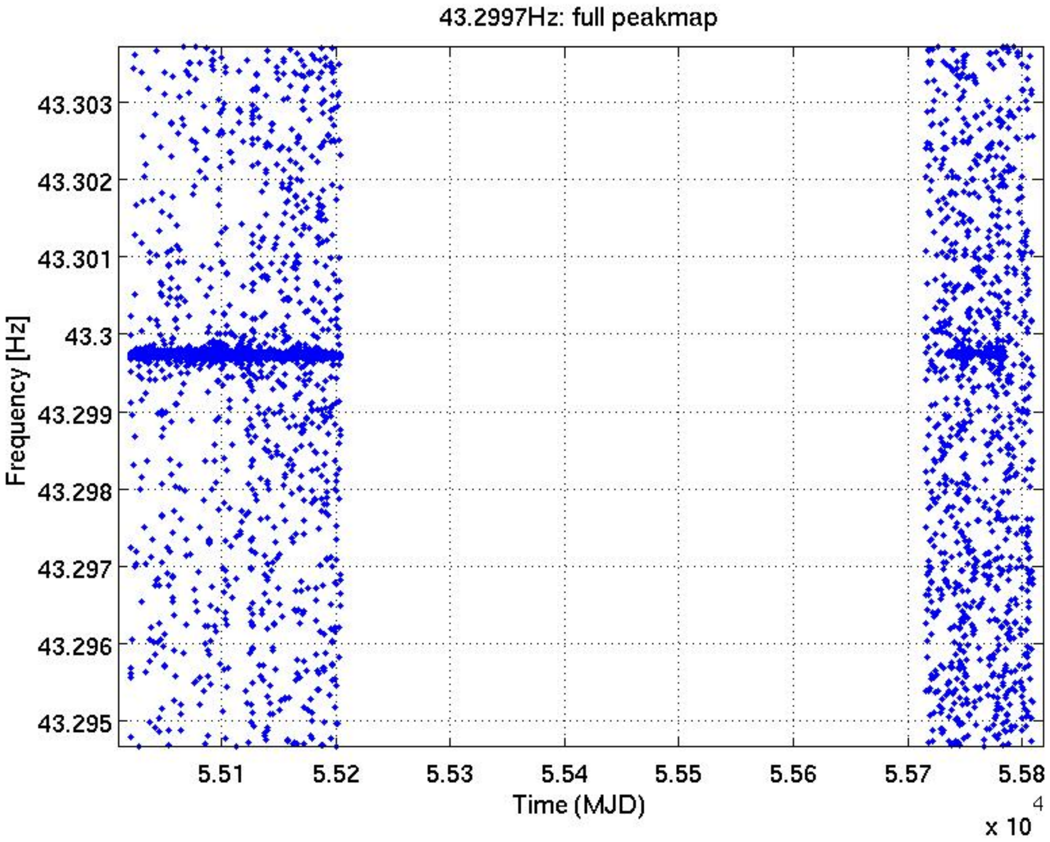
p	f	fdot	λ	β
<5.5E-3	43.29974	-5.04E-12	198.30	89.42

The candidate is at the ecliptical pole, where constant frequency “signals” tend to collect.



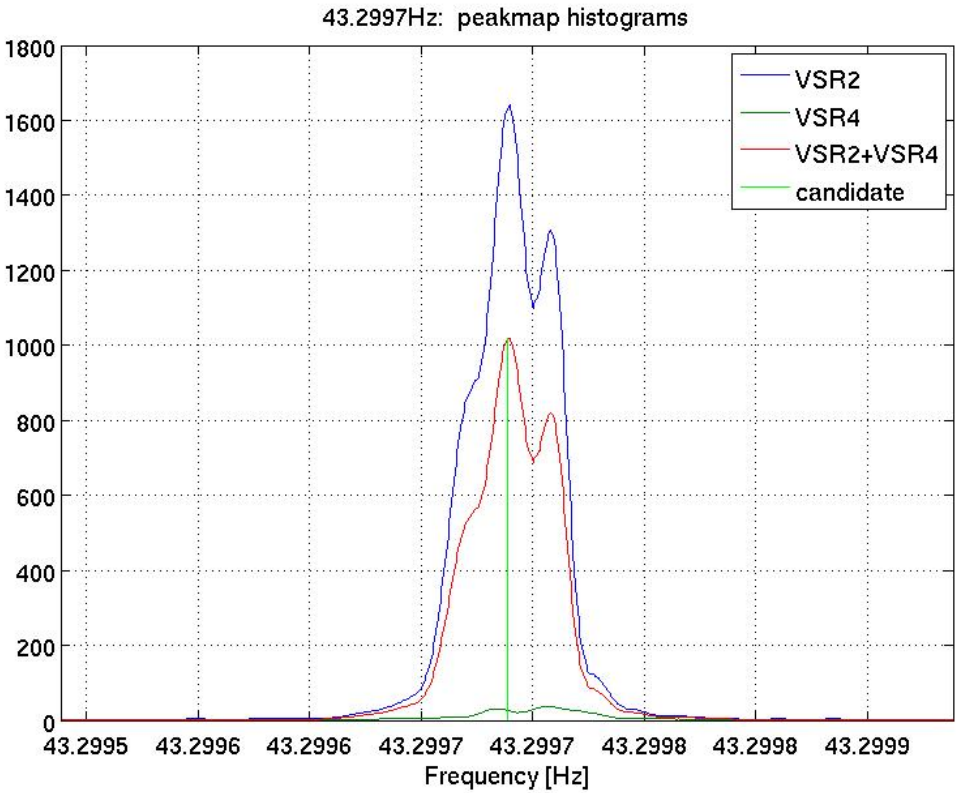
The signal appears to be very strong in VSR2 and very weak in VSR4.

This is also clear from the peakmaps, see next slide.



While in VSR2 the signal appears to be always present, in VSR4 it seems to be present only in the central part.

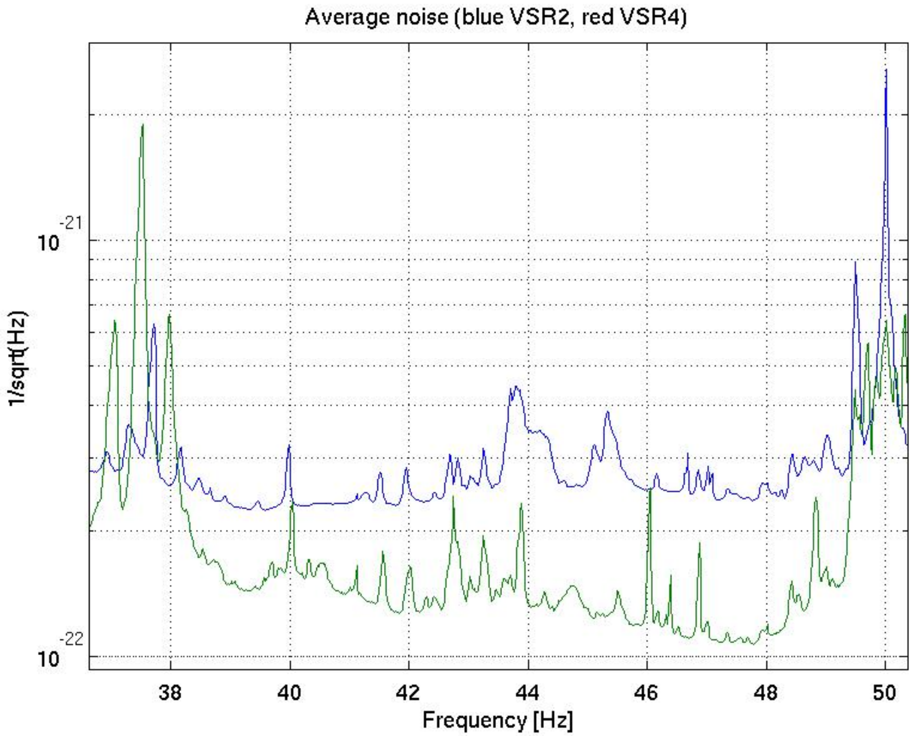
This is confirmed looking at the peakmap histograms of the two runs separately, see next slide.



Candidate in VSR2 much stronger than in VSR4

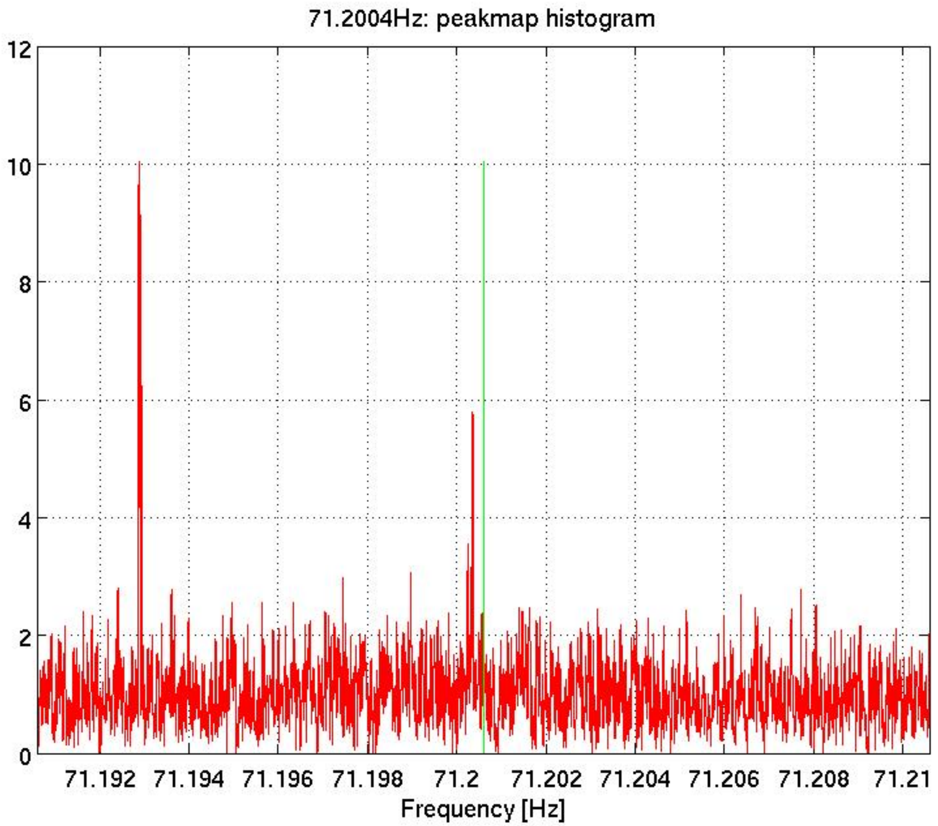
We conclude this candidate is due to noise.

What could produce such a large peak in the histogram?

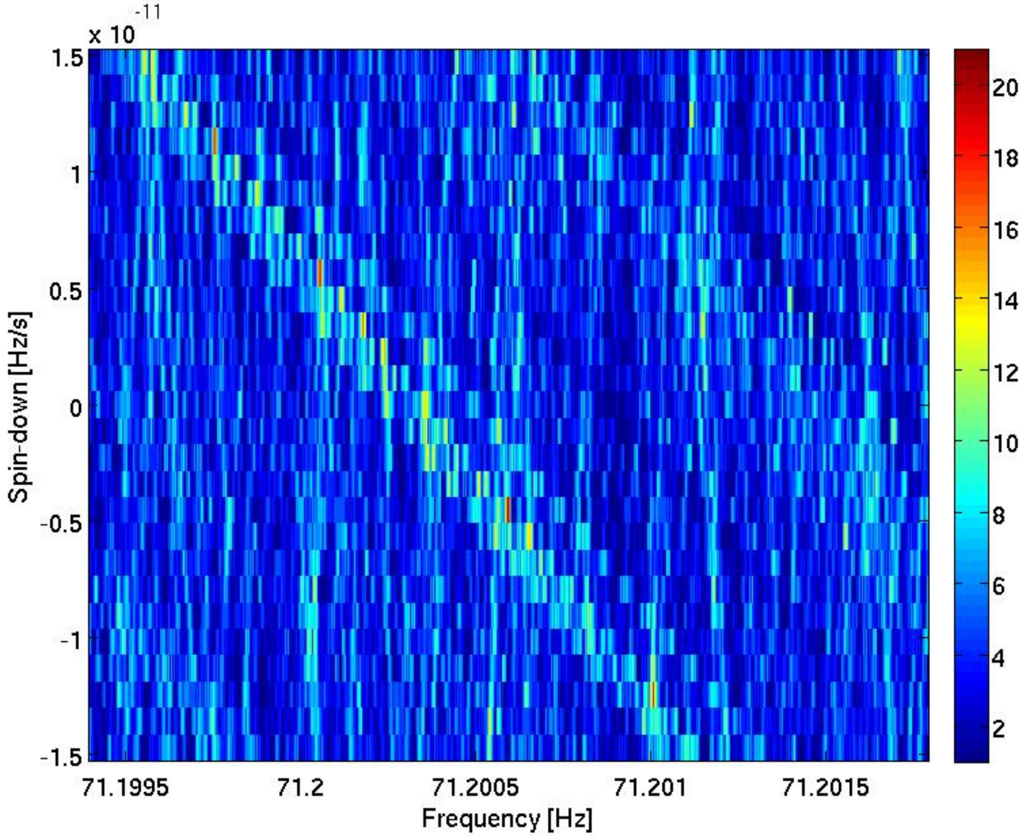


This is a region heavily polluted by lines due to cooling fans (in racks etc.).

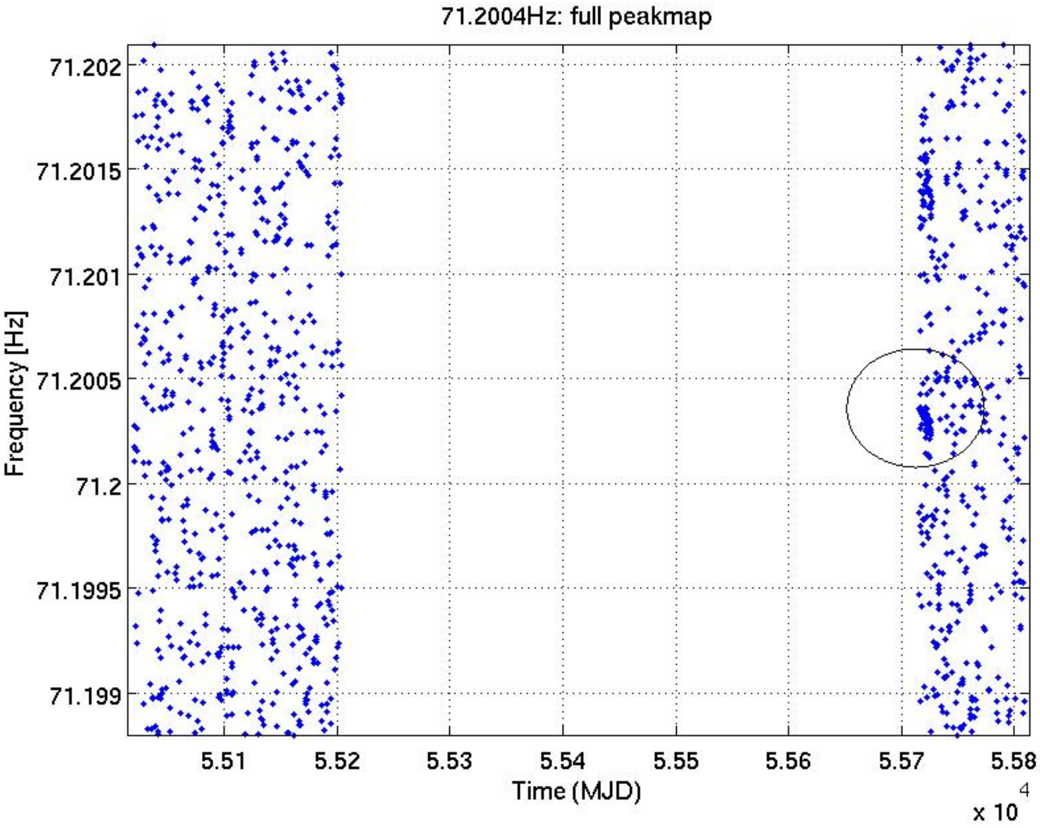
Candidate 71.20Hz



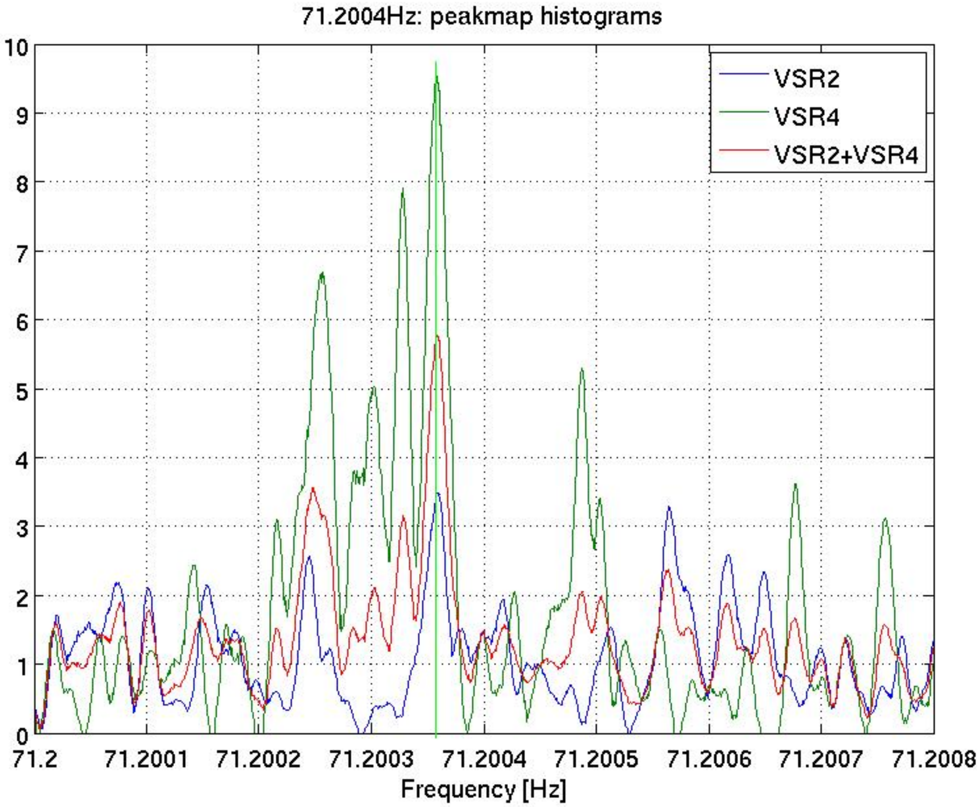
p	f	fdot	λ	β	CR_{raw}	CR_{corr}
1.12E-3	71.20035	1.13E-11	162.83	-55.27	10.64	10.07



Also in this case the strip is rather large, and it seems that mainly one run contributes. Again this suggests it is due to some temporary peak excess in one of the two runs.



Again a peak excess at the beginning of VSR4.



The maximum for VSR2 is significantly smaller than for VSR4.

The full CR is smaller than that of VSR4 alone.

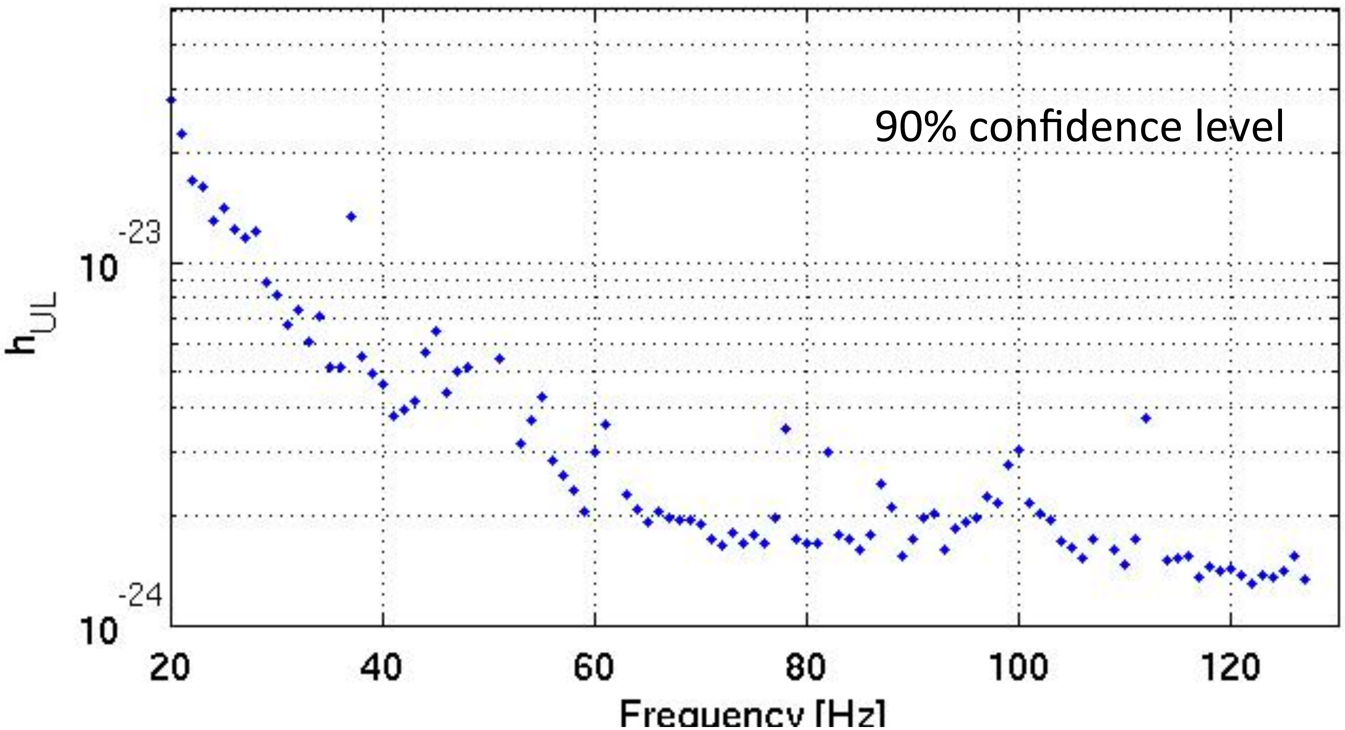
Also in this case we think we can safely reject the candidate.

Upper limit computation (TO BE REVIEWED)

- We compute the upper limit for each 1Hz band

1. inject set of signals at various fixed amplitudes and with random parameters in both VSR2 and VSR4 data;
2. run the analysis (in a small volume around each injection) and select candidates (no clustering)
3. find the pair of nearest coincident candidates
4. check if they are coincident with the injected signal
5. if yes check if their Hough amplitudes (both in VSR2 and VSR4) are larger than those of the actual candidate
6. compute the fraction of injections for which 4. AND 5. are verified
7. find the amplitude for which such fraction is 90%

PRELIMINARY!



Backup slides

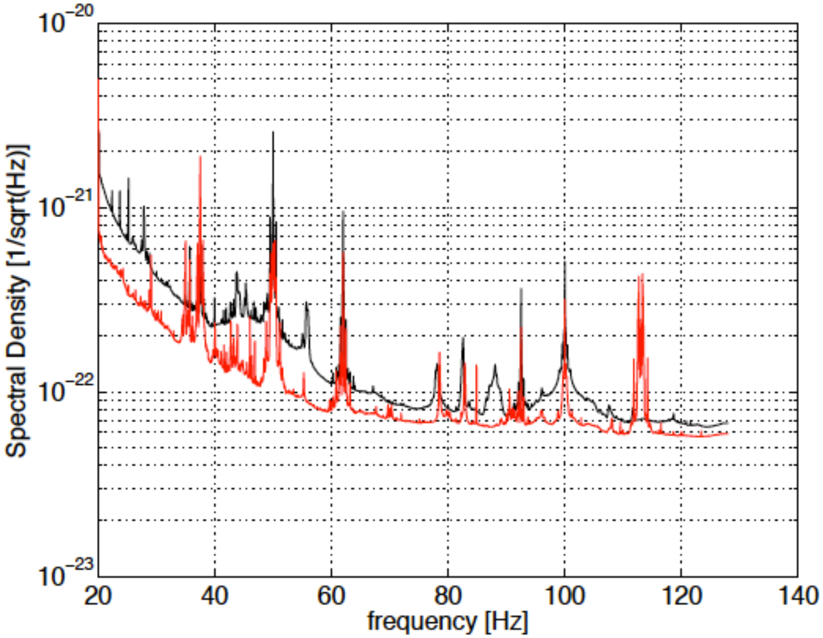


FIG. 1: VSR2 (darker, black in the color version) and VSR4 (lighter, red in the color version) average spectral density, in units of $\frac{1}{\sqrt{Hz}}$, in the frequency range from 20 Hz up to 128 Hz.

Examples of grids