All-sky search for isolated pulsars using the Hough transform

Badri Krishnan
Alicia Sintes & Maria Alessandra Papa
For the LIGO scientific collaboration
badri.krishnan@aei.mpg.de

Max Planck Institut für Gravitationsphysik
GWDGW 2004, Annecy, France
LIGO DCC No. G040538-00-Z
Organization of the talk

- The Hough transform
- Expected sensitivity for S2
- Search parameters and the search pipeline
- Statistical properties
- Setting the thresholds
- Hardware injections
- The search results
- The Monte-Carlo strategy
- Future work
The Hough transform is an incoherent method based on looking for patterns in the time-frequency plane. Less sensitive but computationally inexpensive compared to full coherent search.

- Start with coherent stretches of data and combine them incoherently – present search combines 30 Min long SFTs.

- Similar to power summing (Stack-slide) algorithm but instead of adding power, after sliding we add zeros and ones depending on whether power in SFT bin meets a certain criteria – input to Hough algorithm is set of zeros and ones.

- Advantage of Hough is computational speed – can consider large region in parameter space at once without stepping through templates one-by-one.

- Final result of Hough search is a histogram in parameter space.
The Hough transform

- Time-frequency pattern follows Doppler shift equation

\[ f(t) - f_0(t) = f_0(t) \frac{v(t) \cdot n}{c} \]

- \( f(t) \rightarrow \) observed frequency at time \( t \)
- \( f_0(t) \rightarrow \) intrinsic signal frequency at time \( t \)
- \( v(t) \rightarrow \) detector velocity at time \( t \)
- \( n \rightarrow \) sky-position
The Hough transform

- Smallest signal that can be detected with 1% false alarm and 10% false dismissal in ideal stationary Gaussian noise:

\[ \langle h_0 \rangle = \frac{8.54}{N^{1/4}} \sqrt{\frac{S_n}{T_{coh}}} \]

- \( T_{coh} = 30\text{Min} \) and \( N \) is number of SFTs
- Will be eventually combined with \( F \)-statistic in a hierarchical scheme
- We analyze data from S2 run (14 Feb – 14 Apr 2003) of the LIGO detectors using this method
- S2 data has 1761 SFTs from H1, 687 from L1 and 1384 from H2
- Less sensitive than full directed 2-month coherent search by factor of \( \sim 5 \)
Expected sensitivity for S2

Sensitivity of Hough search for S2 (10% Fd, 1%FA)

- H1
- H2
- L1

Frequency

Sensitivity ($h_0$)
Expected sensitivity for S2

- This implies following astrophysical reach of the search

\[ d = \frac{16\pi^2GN^{1/4}I_{zz}\epsilon f^2}{8.54c^4} \sqrt{\frac{T_{coh}}{S_n(f)}} \]

\[ d^{L1} = 30.4\text{pc} \left( \frac{I_{zz}}{10^{38}\text{kg-m}^2} \right) \left( \frac{f}{300\text{Hz}} \right)^2 \left( \frac{\epsilon}{10^{-6}} \right) \sqrt{\frac{10^{-43}\text{Hz}^{-1}}{S_n}} \]

\[ d^{H2} = 18.3\text{pc} \left( \frac{I_{zz}}{10^{38}\text{kg-m}^2} \right) \left( \frac{f}{300\text{Hz}} \right)^2 \left( \frac{\epsilon}{10^{-6}} \right) \sqrt{\frac{4 \times 10^{-43}\text{Hz}^{-1}}{S_n}} \]

\[ d^{H1} = 22.2\text{pc} \left( \frac{I_{zz}}{10^{38}\text{kg-m}^2} \right) \left( \frac{f}{300\text{Hz}} \right)^2 \left( \frac{\epsilon}{10^{-6}} \right) \sqrt{\frac{3 \times 10^{-43}\text{Hz}^{-1}}{S_n}} \]
Search parameters

- All sky search with \( \sim 3 \times 10^5 \) sky locations (frequency dependent)

\[
\delta \theta = \frac{c}{2vfT_{coh}}
\]

In practice, cannot analyze whole sky all at once – break up sky into 23 patches of roughly equal area (required because we set grid in stereographic plane)

- Frequency band: 200-400 Hz
- One spindown parameter

\[
\delta f_{(1)} = \frac{1}{\frac{T_{obs}}{T_{coh}}} \approx -1.1 \times 10^{-10} \text{Hz/s}
\]

Largest spindown parameter: \( -1.1 \times 10^{-9} \text{Hz/s} \implies 11 \) spindown values considered
- Search takes \( \sim 0.3 \) days for each detector

GWDAW-2004 – p.8/22
The search pipeline

1. Calibrated 1800s input SFTs
2. Calculate power spectrum for each SFT using running median
3. Select frequency bins by simple thresholding
4. Using set of zeros and ones perform the Hough transform
5. Select parameter space points using threshold on number counts
6. Obtain histogram in parameter space
Statistical properties

Peak selection statistics

- We set threshold $\rho_{th}$ on normalized SFT power $\rho$ to select frequency bins
- Noise floor estimated by running median for each sft
- In absence of signal $\rho$ follows exponential distribution
- False alarm rate for peak selection is $\alpha = e^{-\rho_{th}}$
- In presence of signal, distribution of $2\rho$ is the non-central $\chi^2$ distribution

$$p(\rho|\lambda) = e^{-\rho-\lambda/2} I_0(\sqrt{2\rho\lambda})$$

where $\lambda = 4|\tilde{h}|^2 / (T_{coh} S_n)$ is the coherent SNR and non-centrality parameter

- Presence of signal increases peak selection rate

$$\eta \approx \alpha + \frac{1}{2}\alpha\rho_{th}\lambda$$
Statistical properties

Hough map statistics

- In absence of signal distribution of number counts is binomial
  
  \[ p(n|N) = \binom{N}{n} \alpha^n (1 - \alpha)^{N-n} \]

- Select candidates using threshold \( n_{th} \) on number counts

- False alarm rate is
  
  \[ \alpha_H(n_{th}, \rho_{th}, N) = \sum_{n=n_{th}}^{N} p(n|\rho_{th}) \]

- Mean and standard deviation are
  
  \[ \bar{n} = N\alpha \quad \sigma^2 = N\alpha(1 - \alpha) \]
Hough map statistics

- In presence of signal distribution of number counts is binomial only if amplitude modulation and non-stationarity is neglected

\[ p(n|\lambda) = \binom{N}{n} \eta^n (1 - \eta)^{N-n} \]

- False dismissal rate is

\[ \beta_H(n_{th}, \rho_{th}, \lambda, N) = \sum_{n=0}^{n_{th}-1} p(n|\rho_{th}, \lambda) \]

- One possible way to choose thresholds \((\rho_{th}, n_{th})\) is to minimize \(\beta_H\) for a given choice of \(\alpha_H\). For weak signals this leads to \(\rho_{th} = 1.6\) (\(\alpha = 0.20\)) and

\[ n_{th} = N\alpha + \sqrt{2N\alpha(1 - \alpha)} \text{ erfc}^{-1}(2\alpha_H^*) \]
Hardware Injections

Two artificial pulsar signals were injected into all LIGO IFOs during S2 for a 12h period with $h_0 = 2 \times 10^{-21}$, $\psi = 0$, $\cos i = 0$, $\phi_0 = 0$

Pulsar P1 with constant intrinsic frequency:
- $f = 1279.123\text{Hz}$
- $\alpha = 5.147162\text{rad}$, $\delta = 0.376696\text{rad}$

Pulsar P2 with spindown:
- $f = 1288.901\text{Hz}$
- $\alpha = 2.34567\text{rad}$, $\delta = 1.23456\text{rad}$
- $\dot{f} = -10^{-8}\text{rad/s}$

Both signals are clearly detected but observation time of 12h too small to pin-down pulsar parameters exactly
Hardware Injections

Pulsar P1: L1 data

- **Map 2442**
  - Right ascension [radians]: 4.9, 5, 5.1, 5.2, 5.3
  - Declination [radians]: 0, 0.2, 0.4, 0.6, 0.8

- **Map 2222**
  - Right ascension [radians]: 5.135, 5.14, 5.145, 5.15, 5.155
  - Declination [radians]: 0.34, 0.35, 0.36, 0.37, 0.38, 0.39

- **Map 2662**
  - Right ascension [radians]: 4.9, 5, 5.1, 5.2, 5.3
  - Declination [radians]: 0.6, 0.8, 1, 1.2, 1.4

GWDIAW-2004 – p.14/22
Hardware Injections

Pulsar P2: H1 data

map 2013

map 2014
Search results

Hough number count maximized over whole sky and spindowns – H1 data
Search Results

- Expected threshold on number counts for $\alpha_H = 10^{-10}$ should ideally be 459 – should lead to ~ a few candidates every 1Hz band
- Agrees with results in “good” frequency bands
- There are many sharp lines which have been smeared out due to Doppler effect
- Apart from violin modes, 60 Hz lines, 16Hz data acquisition lines, there are many 0.25 Hz harmonics
Hough number count maximized over whole sky and spindowns – L1 data

Search results for L1 data

Maximum number count

Frequency
Search results

Hough number count maximized over whole sky and spindowns – H2 data

![Search results for H2](chart.png)
Distributions in clean bands agree very well with expected binomial distribution but not in bands with strong disturbances.
Setting Upper limits

- We set all-sky upper limits over frequency bands of ~ 1Hz – in progress
- To avoid setting unnecessarily bad limits, we should exclude a priori the known broad features – violin modes, 60Hz lines etc.
- Upper limits will be set by Monte-Carlo injections
- Aim is to set 95% upper limits based on the loudest event in each band, i.e. find the smallest $h_{0}^{95}$ such that
  \[
  \sum_{n=n_{\text{max}}}^{N} p(n|h_{0}^{95}) = 0.95
  \]
  where $n_{\text{max}}$ is largest number count in that band
- We will inject signals with random parameters and random mismatch from search templates and find $h_{0}^{95}$ satisfying above equation.
- Monte-Carlo results being validated within LSC
To-do list

- Validating the S2 upper limits
- Analyze S3 data
- Compare results with other incoherent methods
- Medium term plan – Combine $\mathcal{F}$-stat with Hough algorithm to get longer coherent integration times and thus better sensitivity
- Longer term plan – Use Hough as part of a multi-stage hierarchical search