LASER INTERFEROMETER GRAVITATIONAL WAVE OBSERVATORY
- LIGO -
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Document Type **LIGO-T070124-00-Z** 2007/06/05

Hough search with improved sensitivity

Badri Krishnan and Alicia M. Sintes

Distribution of this draft:
Continuous Waves Upper Limit Group

California Institute of Technology
LIGO Project - MS 51-33
Pasadena CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project - MS 20B-145
Cambridge, MA 01239
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

WWW: http://www.ligo.caltech.edu/

Processed with LaTeX on 2007/06/05

Contents

Abstract

Description of improvements made to Hough S2 driver to improve its sensitivity.

1 Introduction

To improve the sensitivity of the non-demodulated Hough search we must take into account the amplitude modulation of the signal, and also the changing noise floor of the different SFTs. One way to do this is is to change the peak selection threshold based on the sky location and the SFT. This should work, at least in principle, but it has proven somewhat hard to understand. The alternative is to consider a statistic which is not an integer number count but rather

$$n = \sum_{i=1}^{N} w_i n_i \tag{1}$$

where N is the number of SFTs, n_i is as before either 0 or 1 depending on whether the power crosses the threshold. The new feature is the weighting w_i for each SFT. This is what is considered by Palomba et. al. [1] and is similar to what is considered in the power-flux search. When $w_i = 1$ then we recover the usual Hough number count, but we have the freedom to choose it more optimally. We can consider weights to be positive and normalized such that

$$\sum_{i=1}^{N} w_i = N\,. \tag{2}$$

With this normalization we still get a number lying between 0 and N, however as we shall see later, the sensitivity is independent of the overall normalization of the weights. The relative ratio between the weights is what matters. The quantity which determines the sensitivity is $(a^2+b^2)/S_n$ where a and b are the usual amplitude modulation functions which depend on the sky position and time, and the noise floor estimate S_n. The correct way to choose the weights [1] is

$$w_i \propto \frac{a_i^2+b_i^2}{S_n^{(i)}} \tag{3}$$

where the proportionality factor can be removed by normalizing according to (2). Here a_i and b_i are the values of a and b at the mid time of the SFTs and $S_n^{(i)}$ is the noise floor estimate for the i^{th} SFT. With this choice of weights, we get greater contribution at the more sensitive sky locations and from SFTs which have low noise.

2 Statistical Properties and Sensitivity

This section discusses the statistics of the weighted Hough transform and derives the formula for the expected sensitivity. Since the details are lengthy, let us first summarize the results. The notation follows what is used in the Hough methods paper [2].

First, the optimal choice of the peak selection threshold remains the same, i.e. $\rho_{\text{th}} = 1.6$ which leads to a peak selection probability of $\alpha = e^{-\rho_{\text{th}}} = 0.2$. Given this threshold, the number count threshold for a fixed overall false alarm rate α_H is

$$n_{\text{th}} = \alpha A + \sqrt{2||\mathbf{w}||^2\alpha(1-\alpha)}\,\text{erfc}^{-1}(2\alpha_H) \tag{4}$$

where

$$A = \sum_{i=1}^{N} w_i \quad \text{and} \quad ||\mathbf{w}||^2 = \sum_{i=1}^{N} w_i^2 \,. \tag{5}$$

These equations do not assume any particular normalization of the weights. Equation (4) is to be compared with the corresponding equation derived in [2]

$$n_{\text{th}} = N\alpha + \sqrt{2N\alpha(1-\alpha)}\,\text{erfc}^{-1}(2\alpha_H)\,. \tag{6}$$

Equation (4) reduces to this when all the weights are unity in which case $A = ||\mathbf{w}||^2 = N$.

The sensitivity is defined to be the smallest amplitude that would cross the threshold (4) for a given false alarm rate α_H and false dismissal rate β_H. The result is:

$$h_0 = 3.38\mathcal{S}^{1/2}\left(\frac{||\mathbf{w}||}{\mathbf{w}\cdot\mathbf{X}}\right)^{1/2}\sqrt{\frac{\langle S_n^{(i)}\rangle}{T_{\text{coh}}}} \tag{7}$$

where $\langle S_n^{(i)}\rangle$ is the average of all the $S_n^{(i)}$ (it could be other kinds of averages, such as the harmonic mean, median etc.), and

$$\mathcal{S} = \text{erfc}^{-1}(2\alpha_H) + \text{erfc}^{-1}(2\beta_H)\,. \tag{8}$$

and we have defined the N-dimensional vectors

$$\mathbf{w} = (w_1, w_2, \ldots, w_N) \tag{9}$$

$$\mathbf{X} = \langle S_n^{(i)}\rangle\left(\frac{a_1^2+b_1^2}{S_n^{(1)}}, \frac{a_2^2+b_2^2}{S_n^{(2)}}, \ldots, \frac{a_N^2+b_N^2}{S_n^{(N)}}\right) \tag{10}$$

Equation (7) is to be compared with the earlier formula derived in [2]:

$$h_0 = 5.34\frac{\mathcal{S}^{1/2}}{N^{1/4}}\sqrt{\frac{S_n}{T_{\text{coh}}}}\,. \tag{11}$$

This equation was derived by averaging the beam pattern functions over the sky and taking unit weights.

From eq. (7), it is clear that the scaling of the weights does not matter because $w_i \to kw_i$ for any constant k leaves h_0 unchanged. More importantly, it is also clear that the sensitivity is best, i.e. h_0 is minimum, when the inner product $\mathbf{w}\cdot\mathbf{X}$ is maximum. This happens when the two vectors are proportional to each other:

$$w_i \propto X_i \propto \frac{a_i^2+b_i^2}{S_n^{(i)}}\,. \tag{12}$$

This shows that this choice of weights maximises the sensitivity (and minimizes the false dismissal). This also shows that the gain in sensitivity as compared to the case with $w_i = 1$ is when the variance of the X_i is large.

The following subsection gives the proofs of the above statements.

2.1 Details

To simplify life, let us first consider each of the SFTs to have the same Gaussian noise with PSD S_n. Thus, let us start by considering just the amplitude modulation effects. The extension to SFTs with different PSDs will be straightforward. The notation is the same as used in [2].

The peak selection probability in the absence of a signal is $\alpha = e^{-\rho_{\rm th}}$ and it is the same for all the SFTs (this is true even if the SFTs have different PSDs because the peaks are selected based on using the SFTs normalized according to their own PSDs). In the presence of a *small* signal $h(t)$, the peak selection probability to first order in λ_i is

$$\eta_i = \alpha \left(1 + \frac{1}{2}\rho_{\rm th}\lambda_i\right) \tag{13}$$

We have dropped the reference to frequency bins for convenience and we have defined

$$\lambda_i = \frac{4|\tilde{h}_i|^2}{S_n}\,. \tag{14}$$

Here $\tilde{h}_i$ is the Fourier transform of the signal $h_i(t)$ where

$$h_i(t) = F_+^{(i)} h_+(t) + F_\times^{(i)} h_\times(t)\,. \tag{15}$$

The dependence of $h(t)$ on i is only through the antenna pattern functions. It can then be shown that

$$|\tilde{h}_i(f_k)|^2 = \frac{T_{\rm coh}}{4}\left((F_+^{(i)})^2 A_+^2 + (F_\times^{(i)})^2 A_\times^2\right)\,. \tag{16}$$

(Note that the derivation of this equation as presented in [2] is incorrect though the final result is correct.) As usual, the amplitudes are

$$A_+ = \frac{1}{2}h_0(1+\cos^2\iota) \qquad \text{and} \qquad A_\times = h_0\cos\iota\,. \tag{17}$$

We average over the pulsar orientation $\cos\iota$ and the polarization angle ψ (appearing in the antenna patterns). For this, we use the relations

$$\langle (F_+^{(i)})^2\rangle_\psi = \langle (F_\times^{(i)})^2\rangle_\psi = \frac{a_i^2 + b_i^2}{2}\,. \tag{18}$$

Furthermore,

$$\langle A_+^2 + A_\times^2\rangle_\iota = \frac{4h_0^2}{5}\,. \tag{19}$$

Putting these facts together, and also taking into account the spread of the signal within a frequency bin, we obtain

$$\langle\lambda_i\rangle_{\iota,\psi} = 0.7737 \times \frac{2h_0^2 T_{\rm coh}(a_i^2+b_i^2)}{5S_n}\,. \tag{20}$$

The factor 0.7737 arises from considering the spread of the signal over a frequency bin.

To recall, the final statistic that we are considering is $n = \sum w_i n_i$. The mean and variance in the absence of a signal are

$$\bar{n} = A\alpha \qquad \text{and} \qquad \sigma_n^2 = ||\mathbf{w}||^2\alpha(1-\alpha)\,. \tag{21}$$

It is interesting to note that the taking weights different from unity always *increases* the variance. To see this, let us work in a normalization where the sum of the weights is N: $\sum_{i=1}^N w_i = N$. Let us define ϵ_i such that $w_i = 1+\epsilon_i$. Then $\sum_{i=1}^N \epsilon_i = 0$ and

$$\sum_{i=1}^N w_i^2 = \sum_{i=1}^N (1+\epsilon_i)^2 = N + \sum_{i=1}^N \epsilon_i^2 \geq N\,. \tag{22}$$

The increase in sensitivity with non-unit weigts is despite this increase in the variance, which serves to increase the threshold as we shall see below.

In the absence of a signal, let us approximate the distribution $p(n)$ of n by a Gaussian with mean and variance as above. Thus, for a number count threshold $n_{\rm th}$, the false alarm rate is

$$\alpha_H = \int_{n_{\rm th}}^{\infty} p(n)dn = \frac{1}{2}\text{erfc}\left(\frac{n_{\rm th}-\bar{n}}{\sqrt{2}\sigma_n}\right) . \tag{23}$$

From this we can solve for $n_{\rm th}$ and obtain (4) which is:

$$n_{\rm th} = \alpha A + \sqrt{2||\mathbf{w}||^2\alpha(1-\alpha)}\text{erfc}^{-1}(2\alpha_H) . \tag{24}$$

For a given set of weights and a peak selection threshold, this equation decides what number count threshold must be used to obtain a desired false alarm rate α_H. Note that when the weights are all unity as in the usual Hough transform, then we get $A = ||\mathbf{w}||^2 = N$.

Now let us consider the false dismissal rate. In the presence of a signal, the mean and variance of n become

$$\bar{n} = A\alpha + \frac{\alpha\rho_{\rm th}}{2}\sum_{i=1}^{N} w_i\lambda_i \qquad \text{and} \qquad \sigma_n^2 = \sum_{i=1}^{N} w_i^2\eta_i(1-\eta_i) . \tag{25}$$

Note: We have used the same symbols for the mean and variance as in the absence of a signal. This should hopefully not cause any confusion.

It will be useful to expand σ_n^2 upto first order in the λ_i:

$$\sigma_n^2 = \alpha(1-\alpha)||\mathbf{w}||^2\left(1 + \frac{\rho_{\rm th}}{2||\mathbf{w}||^2}\frac{1-2\alpha}{1-\alpha}\sum_{i=1}^{N} w_i^2\lambda_i\right) . \tag{26}$$

To get the false dismissal rate, we approximate the number count distribution $p(n|h)$ by a Gaussian with the mean and variance as above. Then the false dismissal rate is

$$\beta_H = \int_{-\infty}^{n_{\rm th}} p(n|h)dn = \frac{1}{2}\text{erfc}\left(\frac{\bar{n}-n_{\rm th}}{\sqrt{2}\sigma_n}\right) . \tag{27}$$

Expanding again in powers of λ_i and keeping only the first order terms, we get

$$\text{erfc}^{-1}(2\beta_H) + \text{erfc}^{-1}(2\alpha_H) = \sqrt{\frac{\alpha\rho_{\rm th}^2}{8(1-\alpha)}}\frac{\sum_{i=1}^{N} w_i\lambda_i}{||\mathbf{w}||} + \frac{\rho_{\rm th}}{4}\frac{1-2\alpha}{1-\alpha}\frac{\sum_{i=1}^{N} w_i\lambda_i}{||\mathbf{w}||^2}\text{erfc}^{-1}(2\alpha) . \tag{28}$$

The first term on the right hand side of this equation is proportional to $\sqrt{N}$ while the second term does not grow with N when N is large (the easiest way to see this is by taking $w_i \propto 1/N$). Thus the first term dominates and we get

$$\mathcal{S} = \sqrt{\frac{\alpha\rho_{\rm th}^2}{8(1-\alpha)}}\frac{\sum_{i=1}^{N} w_i\lambda_i}{||\mathbf{w}||} . \tag{29}$$

The peak selection threshold is chosen to minimize β_H (or equivalently maximize $\mathcal{S}$ for fized α_H). From (29), this leads us to

$$\frac{d}{d\rho_{\rm th}}\sqrt{\frac{\alpha\rho_{\rm th}^2}{8(1-\alpha)}} = 0 . \tag{30}$$

This is independent of the weights and is exactly what we had obtained in the standard Hough transform with unit weights. The solution to this equation is then the same as before: $\rho_{\rm th} = 1.6$ which leads to $\alpha = e^{-\rho_{\rm th}} = 0.2$.

From (29), we can get the formula for the sensitivity. First we average over ι and ψ using (20) to obtain

$$\mathcal{S} = \sqrt{\frac{\alpha\rho_{\text{th}}^2}{8(1-\alpha)}\frac{\sum_{i=1}^N w_i\langle\lambda_i\rangle_{\iota,\psi}}{||\mathbf{w}||}} = \sqrt{\frac{\alpha\rho_{\text{th}}^2}{8(1-\alpha)}\times 0.7737\times\frac{2h_0^2T_{\text{coh}}}{5S_n}\frac{\sum_{i=1}^N w_i(a_i^2+b_i^2)}{||\mathbf{w}||}}\,. \tag{31}$$

Substituting $\rho_{\text{th}} = 1.6$ and defining $X_i = a_i^2 + b_i^2$, we can write the above equation as

$$h_0 = 3.38\mathcal{S}^{1/2}\left(\frac{||\mathbf{w}||}{\mathbf{w}\cdot\mathbf{X}}\right)^{1/2}\sqrt{\frac{S_n}{T_{\text{coh}}}} \tag{32}$$

This is equation (7) in the case when the $S_n^{(i)}$ are taken as constant, and is what we wanted to show.

The genaralization to the case whan the different SFT noise floors are different is trivial. The starting point is again (20) which we modify by replacing S_n by $S_n^{(i)}$ to obtain

$$\langle\lambda_i\rangle_{\iota,\psi} = 0.7737\times\frac{2h_0^2T_{\text{coh}}(a_i^2+b_i^2)}{5S_n^{(i)}}\,. \tag{33}$$

Then, in equation (31), we cannot take the $S_n^{(i)}$ outside the sum on the right hand side and we obtain instead

$$\mathcal{S} = \sqrt{\frac{\alpha\rho_{\text{th}}^2}{8(1-\alpha)}\frac{\sum_{i=1}^N w_i\langle\lambda_i\rangle_{\iota,\psi}}{||\mathbf{w}||}} = \sqrt{\frac{\alpha\rho_{\text{th}}^2}{8(1-\alpha)}\times 0.7737\times\frac{2h_0^2T_{\text{coh}}}{5\langle S_n^{(i)}\rangle}\frac{1}{||\mathbf{w}||}\sum_{i=1}^N w_i(a_i^2+b_i^2)\frac{\langle S_n^{(i)}\rangle}{S_n^{(i)}}}\,. \tag{34}$$

where we have multiplied and divided by $\langle S_n^{(i)}\rangle$ which is the mean of all the $S_n^{(i)}$ (it could also be the harmonic mean, the median etc.). This then shows that we just have to take $X_i \propto (a_i^2+b_i^2)/S_n^{(i)}$ to obtain the final formula for the sensitivity.

3 Modifications to the LAL Hough routines

The existing Hough routines have been modified to implement the statistic described above, and the changes are committed to cvs. The basic idea of the changes is the following: we keep the LUT construction routines untouched and only modify how the different partial hough map derivatives are combined. Also, we want to make sure that the existing Hough algorithm (when all the weights are unity) still works easily. Thus, the existing functions are not essentially modified, only new functions are added. The S2 Hough driver has been modified to use this new functionality.

The changes in more detail:

- An extra field has been added to the PHMD structure which is a REAL4 weight.
- The typedefs HoughDT and HoughTT which define the types of the hough derivatives and total hough maps, are now REAL4. Of course, they can be changed back to make everything same as before.
- Most of the new functions are in the module `DriveHough.c` and `HoughMap.c`. There are functions to calculate the weights and to fill up the weight entry in the PHMD structure.
- In `HoughMap.c`, there is a function `LALHOUGHAddPHMD2HD-W` which is basically the same as `LALHOUGHAddPHMD2HD`; it is meant to add a PHMD to a total hough derivative. The only difference is that instead of incrementing by ± 1, it adds or subtracts the weight of the PHMD (also for the first column correction). Thus, the PHMD is multiplied by the weight before being added to the total hough map derivative.

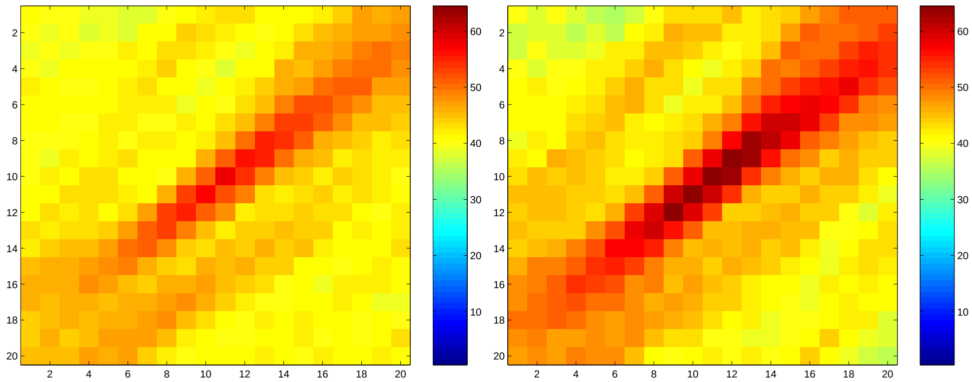

Figure 1: Comparison between equal weights (left figure) and weights which take into account the amplitude modulation (right figure). The signal is clearly stronger in th eright figure (for some reason Matlab only gives color bars from 0 to 60, so the number counts in the figure have been scaled and do not go from 0 to 2000.)

- In `DriveHough.c`, there is a new function `LALHOUGHConstructHMT-W` which is essentially the same as `LALHOUGHConstructHMT`, except that it calls `LALHOUGHAddPHMD2HD-W` instead of `LALHOUGHAddPHMD2HD`.
- In the Hough driver `DriveHough-v3.c` in `lalapps`, there are extra options to turn on or turn off these new functions.

4 Tests and results

The first test is whether if we set all the weights to unity, then we get back the results with the old routines with the only difference being that the number counts are now REAL4 instead of integer. This has been tested with 200 fake noise SFTs with no signal created using `lalapps-makefakedata`. A frequency band of 2Hz from 255 to 257 Hz was analyzed and it was verified that that all the Hough maps, statistics and histograms were identical.

The test of whether the sensitivity is actually improved. 2000 fake SFTs were created with a signal injected at $\alpha = 4.4880$ and $\delta = 0.6435$, at a frequency of 255 Hz and no spindown. Also, we had $\psi = \Phi_0 = 0$ and $a_+ = a_\times$. The amplitude was chosen to be large enough that the Hough maps show the signal clearly but does not saturate the number counts. A square skypatch of 0.2 rad was analyzed using the Hough transform with either all the weights being unity, or with the amplitude modulation taken into account (the noise weighing does not have any effect here). The Hough maps at the closest template points are shown in figure1. The amplitude weighing clearly has a significant effect and the maximum number count goes from about 580 to about 700 compared to a noise floor of a little more than 400. Similar results were observed for other sky locations and more tests are underway to quantify this better. However, it is clear that the improvement is significant.

4.1 Monte-Carlo results

To verify "experimentally" that the weighing scheme does indeed improve sensitivity, a set of Monte-Carlo tests is described in this section. The tests only consider the weighing doe to the amplitude mod-

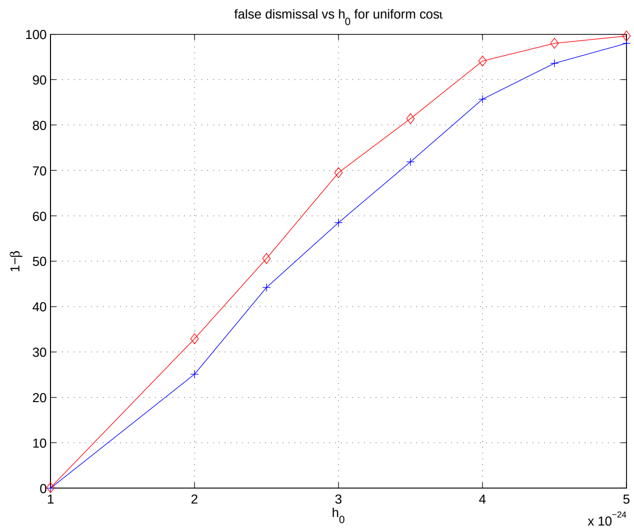

Figure 2: Results of the Monte-Carlo runs. The figures plots $1 - \beta_H$ as a function of the signal amplitude h_0 assuming a threshold given by a false alarm rate of $\alpha_H = 1 \times 10^{-10}$. The blue curve is with all weights being unity while the red curve has weights which take the amplitude modulation into account. The figure is for a uniform distribution of the pulsar orientation $\cos\iota$ between -1 and +1. We see that including the noise weighting improves the sensitivity, and the increase in the sensitivity for h_0 depends on teh false dismissal level.

ulation and not the one arising from non-stationarities in the SFT noise floor. The code for performing these tests is in `lalapps`, it is called `HoughValidateAM.c` and it is located in the directory `src/pulsar/hough/src2`.

Random Gaussian pure noise SFTs were generated between 250 and 260 Hz using `lalapps-makefakedata`. A pulsar signal was injected into these SFTs with a certain amplitude h_0 and two number counts were calculated: with all the weights being unity and with the amplitude modulation taken into account. These number counts were calculated using a template which was perfectly matched in frequency and sky location. In all these tests, the spindown was set to zero. The above steps were repeated 1000 times for different pulsar signals of the same amplitude and frequency, but randomly chosen sky-locations and pulsar orientations determined by $\cos\iota$. For each of these injections, apart from the number counts, we also calculate the threshold $n_{\rm th}$ for a false alarm rate of $\alpha_H = 1 \times 10^{-10}$. The false dismissal rate is calculated according to the fraction of times that the number count crosses this threshold. Note that the thresholds are different depending on whether the weights are used or not.

There are three plots. The first plot is shown in figure 2 which shows $1 - \beta_H$ as a function of the signal amplitude h_0 for a population of signals which have a uniform distribution in the sky and in $\cos\iota$. The plots in figure 3 are the same, but in these plots the value of $\cos\iota$ has been fixed to 0 and 1. In all cases, the amplitude weighing increases the sensitivity significantly.

Finally, let us investigate how large the size of the skypatch can be, such that there is an appreciable gain in sensitivity bu weighing. We consider a particular case presented earlier, namely, we inject a signal with amplitude $h_0 = 3.5 \times 10^{-24}$ with the sky-locations distributed uniformly over the sky, and $\cos\iota$ uniformly distributed between -1 and 1. The difference here is that the weights are not calculated at the precise sky-location of the signal, but rather at a mismatched point. The gain in sensitivity should be largest when the weights are calculated at the perfectly matched sky-point and should decrease as the mismatch is increased.

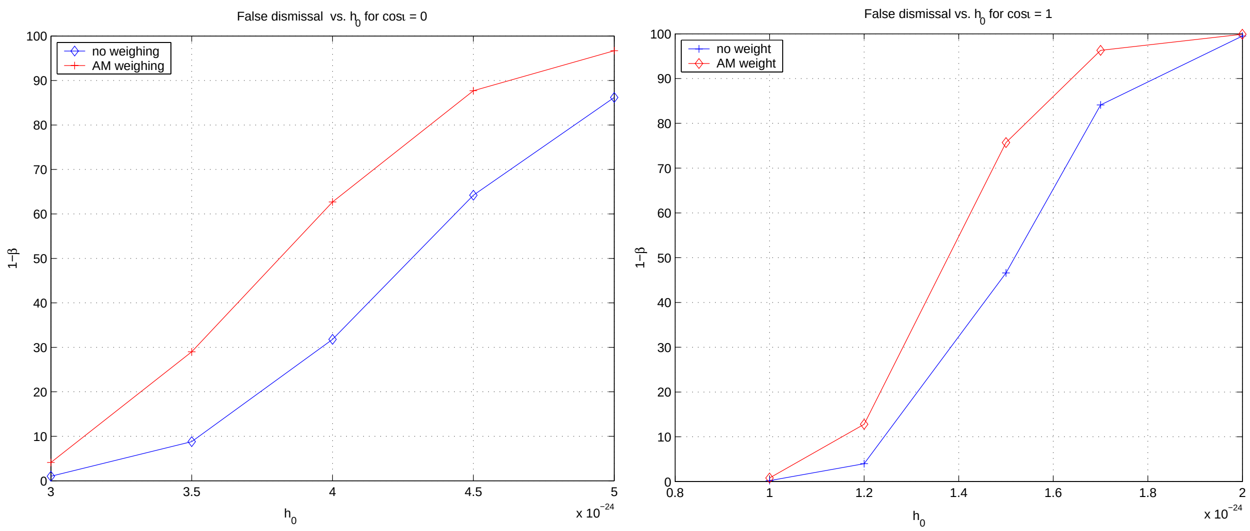

Figure 3: Same as fig. 2 but for $\cos\iota = 0$ (left figure) and $\cos\iota = 1$ (right figure).

We can decide how much gain we would like, and set the size of the sky-patch accordingly. The plot in figure 4 shows the detection probability as a function of the mismatch in sky position.

References

[1] C. Palomba, P. Astonem, and S. Frasca, “Adaptive Hough transform for the search of periodic sources”, GWDAW proceedings from Annecy.

[2] B. Krishnan, A.M. Sintes, M.A. Papa, B.F. Schutz, S. Frasca, and C. Palomba, Phys. Rev. D **70**, 082001 (2004).

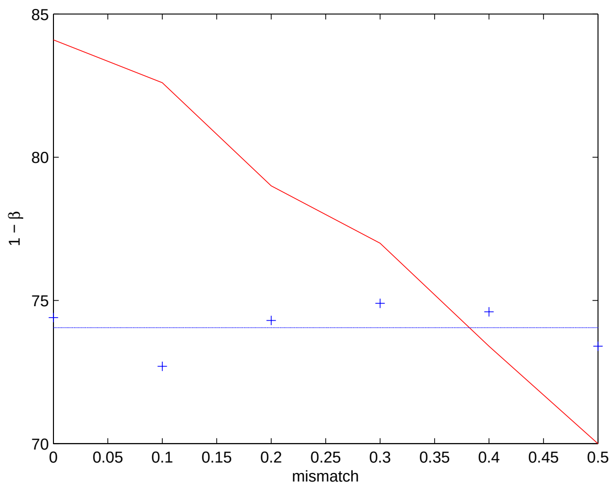

Figure 4: Plot of the detection probability as a function of the mismatch in sky-location in radians. The blue curve is the detection probability without the weighing while the red curve is with the weighing taken into account. As expected, the blue curve fluctuates about a mean value while the red curve decreases monotonically with increasing mismatch. Note that the x-axis is the mismatch in either α or δ which are both taken to be equal. The true mismatch is thus roughly $\sqrt{2}$ times the value on the x-axis. The corresponding skypatch size will be twice the value of the x-axis.