All-sky search of NAUTILUS data

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2008 Class. Quantum Grav. 25 184012

(http://iopscience.iop.org/0264-9381/25/18/184012)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 137.138.139.20
The article was downloaded on 09/08/2012 at 11:09

Please note that terms and conditions apply.
All-sky search of NAUTILUS data

P Astone1, M Bassan2, P Bonifazi1, K M Borkowski1, R J Budzyński1, A Chincarini6, E Coccia3,7, S D’antonio2, M Di Paolo Emilio7,8, V Fafone9, S Frasca10, S Foffa11, G Giordano9, P Jaranowski12, W Kondrak13, A Królak13, M Maggiore11, A Marin9, Y Minenkov3, I Modena2, G Modesto6, A Moleti7, G V Pallottino10, C Palomba1, R Parodi6, M Pietka12, G Pizzella14, H J Pletsch15, L Quintieri9, F Ricci10, A Rocchi2, F Ronga9, R Sturani11, R Terenzi3, R Vaccarone6 and M Visco1

1 Istituto Nazionale di Fisica Nucleare INFN, Rome, Italy
2 University of Rome ‘Tor Vergata’ and INFN, Rome II, Italy
3 INFN, Laboratori Nazionali del Gran Sasso, Assergi, L’Aquila, Italy
4 IFSI-CNR and INFN, Rome, Italy
5 Centre of Astronomy, Nicolaus Copernicus University, Toruń, Poland
6 Department of Physics, Warsaw University, Warsaw, Poland
7 INFN, Laboratori Nazionali del Gran Sasso, Assergi, L’Aquila, Italy
8 Università dell’Aquila, Italy
9 Istituto Nazionale di Fisica Nucleare INFN, Genova, Italy
10 University of Rome ‘La Sapienza’ and INFN, Rome, Italy
11 Institute of Mathematics, Polish Academy of Sciences, Warsaw, Poland
12 Faculty of Physics, University of Białystok, Białystok, Poland
13 University of Rome ‘Tor Vergata’ and INFN, Frascati, Italy
14 Dep. de Phys. Théorique, Université de Genève, Genève, Switzerland
15 Albert-Einstein-Institut, Max-Planck-Institut f¨ur Gravitationsphysik, Hannover, Germany

E-mail: krolak@impan.gov.pl

Received 15 April 2008, in final form 22 June 2008
Published 2 September 2008
Online at stacks.iop.org/CQG/25/184012

Abstract
A search for periodic gravitational-wave signals from isolated neutron stars in the NAUTILUS detector data is presented. We have analyzed half a year of data over the frequency band (922.2; 923.2) Hz, the spindown range (−1.463 × 10−8; 0) Hz/s and over the entire sky. We have divided the data into two day stretches and we have analyzed each stretch coherently using matched filtering. We have imposed a low threshold for the optimal detection statistic to obtain a set of candidates that are further examined for coincidences among various data stretches. For some candidates we have also investigated the change of the signal-to-noise ratio when we increase the observation time from 2 to 4 days. Our analysis has not revealed any gravitational-wave signals. Therefore we have imposed upper limits on the dimensionless gravitational-wave amplitude over the parameter space that we have searched. Depending on frequency, our upper limit ranges from 3.4 × 10−23 to 1.3 × 10−22. We have
attempted a statistical verification of the hypotheses leading to our conclusions. We estimate that our upper limit is accurate to within 18%.

PACS numbers: 95.55.Ym, 04.80.Nn, 95.75.Pq, 97.60.Gb

(Some figures in this article are in colour only in the electronic version)

1. Introduction

We present results of the search of the NAUTILUS (a resonant bar detector [1]) data for periodic gravitational-wave (GW) signals. We search the band $(922.2; 923.2)$ Hz and the spindown range $(-1.463 \times 10^{-8}; 0)$ Hz/s over the entire sky. We analyze NAUTILUS data collected in the year 2001. We divide the data into stretches of 2 sidereal days. Each stretch of data is analyzed coherently using matched filtering in the form of the $F$-statistic [2, 3]. We have analyzed slightly more than half a year of data.

Previous analyses of bar detector data for periodic GW signals were: the search of the galactic center and the globular cluster 47 Tucanae with the ALEGRO detector [4], the search of the galactic center using the EXPLORER detector data [5] and an all-sky search using the EXPLORER data [6, 7]. LIGO (Laser Gravitational Wave Observatory) data were searched for known pulsars [8–10], over the entire sky [11] using the coherent method and over all the sky using incoherent methods [12, 13]. Currently LIGO data are analyzed over the entire sky by the Einstein@Home project [14].

In section 2 we present the data analysis methods used in our search. In section 3 we outline our search procedure. In section 4 we discuss the analysis of the candidates. This analysis consists of two parts: the first part is the search for coincidences between the candidates each obtained from a different stretch of data and the second part is an investigation of the increase of the signal-to-noise ratio of candidates when we increase the observation time from 2 to 4 days. In section 5, we impose upper limits on amplitudes of the gravitational waves in the parameter space that we have searched.

2. Data analysis methods

In order to search for gravitational waves from long-lived periodic sources we have used the maximum likelihood (ML) method. For the case of Gaussian noise the ML method consists of linearly filtering the data with a template matched to the signal that we are searching for. The main complication of the matched filtering is that the signal depends on several unknown parameters. This requires evaluation of the likelihood function over a large parameter space. In order to minimize the computation time we use several data analysis tools. Firstly we find the maximum likelihood estimators of some parameters (four in the case of a GW signal from a rotating neutron star that we are searching for) in a closed analytic form, thereby reducing the dimensionality of the parameter space that we have to search. The likelihood function over the reduced parameter space is called the $F$-statistic and it is derived in [2]. Secondly we analyze data of length equal to an integer multiple of a sidereal day. This leads to a considerable simplification of the $F$-statistic and consequently reduces the number of numerical operations to evaluate it. The $F$-statistic for observation time equal to an integer number of sidereal days is given in section III of [15]. Thirdly we use optimal numerical algorithms, in particular the fast Fourier transform (FFT) in order to calculate the $F$-statistic efficiently. Fourthly we
minimize the number of $F$-statistic calculations over the parameter space by solving a covering problem for this space \cite{16, 17}. Let us explain the latter two data analysis tools in more detail.

**Fast Fourier transform.** Estimates have shown \cite{20, 21} that for the bandwidth and the spindown range that we search we need to take into account in our templates only one spindown parameter in order to match the signal. Consequently the phase modulation function $\phi(t)$ of the waveform is given by

$$\phi(t) = \omega_0 t + \omega_1 t^2 + (\omega_0 + 2\omega_1) \frac{n_0 \cdot r_d(t)}{c},$$

where $\omega_0$ is an angular frequency and $\omega_1$ is the spindown parameter, $n_0$ is the constant unit vector in the direction of the star in the Solar System Barycenter (SSB) reference frame (it depends on the right ascension $\alpha$ and the declination $\delta$ of the source) and $r_d$ is the vector joining the SSB with the detector and $c$ is the speed of light. The detection statistic $F$ involves two integrals of the form

$$F = \int_{T_0}^{T} x(t) a(t) e^{-i \phi(t)} dt,$$

where $x(t)$ is the data stream and $a(t)$ is the amplitude modulation function that depends on $\delta$ and $\alpha$. The above integral is not a Fourier transform because the frequency $\omega_0$ in the phase multiplies the term $n_0 \cdot r_d(t)$ which is a nonlinear function of time. In order to convert the integral into a Fourier transform we introduce the following interpolation procedure. The phase $\phi(t)$ (equation (1)) can be written as

$$\phi(t) = \omega_0 [t + \phi_m(t)] + \phi_s(t),$$

where

$$\phi_m(t) := \frac{n_0 \cdot r_d(t)}{c},$$

$$\phi_s(t) := \omega_1 t^2 + 2 \frac{n_0 \cdot r_d(t)}{c} \omega_1 t.$$
The covering problem. The covering problem is to find the minimum number of templates in the parameter space \[17\], so that the fractional loss in signal-to-ratio is not less than \[1 − MM\] (MM is the minimal-match parameter introduced by Owen \[22\]). In order to solve the covering problem we introduce a useful approximate model of the gravitational-wave signal from a rotating neutron star. The model relies on (i) neglecting all spindowns in the phase modulation due to motion of the detector with respect to the SSB; and (ii) discarding the component of the vector \(\mathbf{r}_d\) (connecting the SSB and the detector) which is perpendicular to the ecliptic plane. These approximations lead to the following phase model of the signal:

\[
\phi_{\text{lin}}(t) = \omega_0 t + \omega_1 t^2 + \alpha_1 \mu_1(t) + \alpha_2 \mu_2(t),
\]

where \(\alpha_1\) and \(\alpha_2\) are new constant parameters,

\[
\alpha_1 := \omega_0 (\sin \alpha \cos \delta \cos \epsilon + \sin \delta \sin \epsilon),
\]

\[
\alpha_2 := \omega_0 \cos \alpha \cos \delta,
\]

where \(\epsilon\) is the obliquity of the ecliptic and where \(\mu_1(t)\) and \(\mu_2(t)\) are known functions of time,

\[
\mu_1(t) := R_{ES}^x(t) + R_{E}^x(t) \cos \epsilon,
\]

\[
\mu_2(t) := R_{ES}^y(t) + R_{E}^y(t).
\]

\(R_{ES}^x\) is the \(x\)-component of the vector joining the center of Earth and the SSB and \(R_{E}^x\) is the \(x\)-component of the vector joining the center of Earth and the detector. \(R_{ES}^y(t)\) and \(R_{E}^y(t)\) are the corresponding \(y\)-components. We also neglect the slowly varying modulation of the signal’s amplitude, so finally we approximate the whole signal \(h(t)\) by

\[
h(t) = A_0 \cos(\phi_{\text{lin}}(t) + \phi_0),
\]

where \(A_0\) and \(\phi_0\) are the constant amplitude and initial phase, respectively. The above signal model is called *linear* because it has the property that its phase given by equation \(9\) is a linear function of the parameters. We have shown \[19\] that the above model is a good approximation to the accurate response of the detector to the GW signal in the sense that the Fisher matrix for the linear model reproduces well the Fisher matrix for the accurate model. Thus, whenever a Fisher matrix is needed we can use the Fisher matrix for the linear model as an approximation to the Fisher matrix for the accurate model. The great advantage of the linear model is that components of its Fisher matrix are constant, independent of the values of the parameters. In order to solve the covering problem for the parameter space we use the Fisher matrix as a metric on the parameter space. Because the components of the Fisher matrix are constant the grid is uniform what greatly simplifies its construction. In our search, as a grid we use the hypercubic lattice \[16\]. However we have an additional constraint. In order to apply the FFT algorithm the nodes of the grid must coincide with the Fourier frequencies. We have constructed a suitable grid by performing rotations and dilatations of the original hypercubic lattice. We construct the grid in the parameters \(\omega_0, \omega_1, \alpha_1, \alpha_2\) and then transform it to parameters \(\omega_0, \omega_1, \delta, \alpha\) for which the \(F\)-statistic is calculated.

The linear parametrization has one more application. We use it in order to calculate the threshold for the \(F\)-statistic corresponding to a certain false alarm probability. Namely, using the linear parametrization we divide the parameter space into cells as explained in \[2, 20\]. All the cells are exactly the same and their number \(N_c\) is easily calculated using the Fisher matrix (see section IIIB of \[20\]). The false alarm probability \(\alpha\) is the probability that \(F\) exceeds threshold \(F_o\) in one or more cells and is given by

\[
\alpha = 1 - [1 - P_F(F_o)]^{N_c},
\]

where \(P_F\) is the false alarm probability for a single cell.
3. Search procedure

We have searched the data collected by the NAUTILUS detector in the year 2001. The bandwidth of $[922.2; 923.2]$ Hz, where the detector is most sensitive, has been analyzed. We have divided the data into stretches which span two sidereal days. We have assumed a minimum pulsar spindown age $\tau_{\text{min}}$ equal to 1000 years and so we have searched the negative frequency time derivatives in the range of $(-1.463 \times 10^{-8}; 0)$ Hz/s. For this $\tau_{\text{min}}$ and 2 days of the observation time it is sufficient to include only one spindown in the phase $[21, 20]$. Each two day sequence was analyzed coherently using the $F$-statistic. We have used the constrained hypercubic grid as explained in the previous section. For the grid construction we have assumed the minimal-match parameter $MM = \sqrt{3}/2$ [22]. Using this minimal-match value our grid consists of around $3.1 \times 10^{13}$ grid points ($2^{19}$ frequency bins, $\sim 10^3$ spindowns, $\sim 6 \times 10^4$ sky positions). The threshold on $2F$ corresponding to 1% false alarm probability has been calculated using equation (15) and is around 72. In order to compensate the loss of signal-to-noise ratio (SNR) due to the discreteness of the grid, imperfect templates and numerical approximation in evaluation of the $F$-statistic (resampling procedure) we have adopted two lower thresholds on $2F$ equal to 40 and 50. We have registered parameters of all templates which crossed the threshold of 40. For threshold crossings of 50 we have performed a verification procedure. The verification procedure consisted of calculating the $F$-statistic for the template parameters of the candidate using a 4 day stretch of data involving the original two day stretch. For a true gravitational-wave signal by this procedure one would expect an increase of signal-to-noise ratio by $\sqrt{2}$. In total, we have analyzed 93 two day data stretches. In figure 1 we have presented the two-sided amplitude spectrum of the NAUTILUS
detector data that we have analyzed. The spectrum was obtained in the following way. We have estimated the power spectrum density in each of the 93 two day data sequences and then we have taken the square root of the average of the 93 power spectra. We see that the best sensitivity is around $5 \times 10^{-22}$ Hz$^{-1/2}$. Moreover, we have obtained the rms error of our power spectrum estimate by calculating the variance from the estimates of the spectra of in each of the 93 data segments. The relative $1\sigma$ error in the amplitude power spectrum is around 18%.

During the search we have obtained 537 665 380 candidates above the $2F$-threshold of 40 and 9 038 817 above the $2F$-threshold of 50.

4. Analysis of the candidates

4.1. Signal-to-noise ratio of the candidates

In figure 2 we have plotted a histogram of the frequencies of all the candidates above the $2F$ threshold of 50. The histogram shows an excess of candidates in the frequency band of (922.4; 922.6) Hz. This excess is a result of the presence of a periodic interference in the data that appears as a series of harmonics in the bandwidth of the detector. One of the harmonics is located in the subband (922.2; 923.2) Hz. The effect of the harmonic is visible in our estimate of the spectrum (figure 1) and appears as a bump in the band (922.5; 922.6) Hz.

As a first step in the candidate analysis we have calculated the increase in signal-to-noise ratio when we increase the observation time from 2 to 4 days. This has been done for all the candidates above the threshold $2F = 50$. Figure 3 shows the highest increase in SNR for candidates when going from a two day data stretch to the four day one. The maximum is calculated for each of the 93 data stretches analyzed. We see that typically the highest gain in the signal-to-noise is 1.2. This should be compared with the theoretical gain of $\sqrt{2}$.
of SNR when we increase the observation time by a factor of 2. The periodic interference present in the data to which we attribute these maximum SNR increases does not give a higher increase of the SNR because its frequency changes erratically over the observation time of days and it cannot reproduce the Doppler shift of a real GW signal modulated by detector motion with respect to the SSB. Assuming that the two day sequence is independent of the four day sequence we could perform the $F$-test that consists of calculation of the ratio $F$ of the $F$-statistic for 4 days observation time and the $F$-statistic for 2 days observation time. Taking as the null hypothesis for the test that data is only Gaussian noise the 2 $F$-statistic has the central $\chi^2$ distribution with 4 degrees of freedom and the ratio $F$ has Fisher–Snedecor distribution $F(4, 4)$. The typical highest value of $F$ for a given data segment is around 1.5. The probability of $F$ crossing the threshold 1.5 is around 37%. This would give a high confidence that data is noise only. Unfortunately this is only a crude approximation because the two day sequence is contained in the four day one and the assumption of independence of the two $F$-statistic is not fulfilled.

4.2. Coincidences among the candidates from different data stretches

Candidates from different data stretches are considered coincident if they cluster closely together in the four-dimensional parameter space $(\omega_0, \omega_1, \delta, \alpha)$. We employ the clustering method described in [14], which uses a grid of ‘coincidence cells’. This method will reliably detect strong signals which would produce candidates with closely-matched parameters in many of the different data stretches.

In a first step, the frequency value of each candidate above the threshold of $2F = 40$ is shifted to the same fiducial time: the GPS start time of the earliest ($j = 1$) stretch, $t_{\text{fiducial}} = t_1 = 662\,547\,735.998\,8098$ s. Defining $T_0$ to be the time span of two sidereal days,
the frequencies of the candidates are shifted to \( t_{\text{fiducial}} \) via
\[
\omega_0(t_{\text{fiducial}}) = \omega_0(t_j) + (j - 1)2\omega_1 T_0,
\]
where \( t_j \) is the starting time of the \( j \)th data stretch, given by \( t_j = t_{\text{fiducial}} + (j - 1)T_0 \).

To find coincidences, a grid of cells is constructed such that the cells are rectangular in the coordinates \((\omega_0, \omega_1, \delta, \alpha \cos \delta)\). The dimensions of the cells are adapted to the parameter space search. Thus, the cells are constructed to be as small as possible to reduce the probability of coincidences due to false alarms. However, since each of the 93 different data stretches uses a slightly different parameter space grid, the coincidence cells must be chosen to be large enough that the candidates from a source (which would appear at slightly different points in parameter space in each of the 93 data stretches) would still lie in the same coincidence cell. As a conservative choice we use cell sizes in \( \omega_0 \) of \( 5.8 \times 10^{-4} \) Hz, in \( \omega_1 \) of \( 2.08 \times 10^{-11} \) Hz s\(^{-1}\) and an isotropic cell grid in the sky with equatorial spacing of 0.028 rad. Each candidate event is assigned to a particular cell. In cases where two or more candidate events from the same data stretch \( j \) fall into the same cell, only the candidate having the largest value of \( 2F \) is retained in the cell. Then the number of candidate events per cell coming from distinct data stretches is counted.

From the 93 different data stretches, this coincidence method found that we get candidates which appear consistently in not more than 4 data stretches uniformly over the search bandwidth, where there are no instrumental interferences. This is the background of the number of coincidences. We would like to test the null hypothesis that the coincidences are result of the noise only. Let us assume that the parameter space is divided into \( N_{\text{cell}} \) independent coincidence cells, the candidate events are independent and the probability for a candidate event to fall into any given coincidence cell is \( \frac{1}{N_{\text{cell}}} \). Thus probability \( \epsilon \) that a given coincidence cell is populated with one or more candidate events is given by
\[
\epsilon = 1 - \left(1 - \frac{1}{N_{\text{cell}}}\right) \epsilon_{\text{seg}},
\]
where \( \epsilon_{\text{seg}} \) is the number of candidate events per data segment. The probability \( p_F \) that any given coincidence cell contains candidate events from \( C_{\text{max}} \) or more distinct data segments is given by a binomial distribution
\[
p_F = \sum_{n=C_{\text{max}}}^{N_{\text{seg}}} \binom{N_{\text{seg}}}{n} \epsilon^n (1 - \epsilon)^{N_{\text{seg}} - n}.
\]
Finally the probability \( P_F \) that there is \( C_{\text{max}} \) or more coincidences in one or more of the \( N_{\text{cell}} \) cells is
\[
P_F = 1 - (1 - p_F)^{N_{\text{cell}}}.
\]
The expected number of cells with \( C_{\text{max}} \) or more coincidences is given by
\[
N_F = N_{\text{cell}} p_F.
\]
In our case the number of cells is given by \( N_{\text{cell}} = 5.9 \times 10^{10} \), the number of data segments is \( N_{\text{seg}} = 93 \), and the number of candidates per data segment is \( \epsilon_{\text{seg}} = 5.8 \times 10^6 \). From equation (19) we find that the probability of finding \( C_{\text{max}} = 4 \) or more coincident candidates is almost one. Thus for the background coincidences we can accept the null hypothesis that they are from noise only with a high confidence. Over the bandwidth \( \langle 922.4; 922.6 \rangle \) Hz we find an excess of coincidences with the maximum of 8 coincidences. By equation (19), the false alarm probability associated with 8 or more coincidences is of the order of \( 10^{-11} \) and thus they cannot be attributed to noise. We consider these coincidences to be due to the periodic interference present in the data.
5. Upper limits

Our verification procedure consisting of coincidences among the candidates from distinct data segments and an analysis of the increase of signal-to-noise ratio presented in section 4 did not produce convincing evidence of a gravitational-wave signal. We therefore proceeded to estimate the upper limits for the amplitudes of the gravitational-wave signals in the parameter space that we have searched. Detection of a signal is signified by a large value of the $F$-statistic that is unlikely to arise from the noise-only distribution. If instead the value of $F$ is consistent with pure noise with high probability we can place an upper limit on the strength of the signal. One way of doing this is to take the loudest event obtained in the search and solve the equation

$$P = P_D(\rho_{ul}, F_{\text{loudest}})$$

for signal-to-noise ratio $\rho_{ul}$, where $P_D$ is the detection probability, $F_{\text{loudest}}$ is the value of the $F$-statistic corresponding to the loudest event and $P$ is a chosen confidence. Then $\rho_{ul}$ is the desired upper limit with confidence $P$. We can also obtain an upper limit $\rho_{ul}$ with confidence $P$ for several independent searches from the equation

$$P = 1 - \prod_{s=1}^{L} [1 - P_D(\rho_{ul}, F_{\text{loudest},s})],$$

where $F_{\text{loudest},s}$ is the threshold corresponding to the loudest event in $s$th search and $L$ is the number of searches. Here $P$ is the probability that a signal with signal-to-noise ratio $\rho_{ul}$ crosses the threshold $F_{\text{loudest},s}$ in at least one of the $L$ independent searches. To calculate
\( \rho_{ul} \text{ we assume that the data have a Gaussian distribution and consequently the probability of detection } P_D \text{ has a non-central } \chi^2 \text{ distribution with 4 degrees of freedom and the noncentrality parameter equal to } \rho_{ul}^2. \) We have investigated this assumption by obtaining histograms of the 2F-statistic values of the candidates and comparing them to the central \( \chi^2 \) distribution with 4 degrees of freedom. The result is shown in figure 4. There is an overall qualitative agreement of candidate’s distributions with the theoretical one. However, the candidate’s distributions do not pass a goodness-of-fit test for a \( \chi^2 \) distribution at the significance level of 5%.

In order to translate our upper limit on the SNR into the upper limit on the gravitational-wave amplitude, we use equation (93) of \cite{2} for signal-to-noise ratio of a GW signal from a spinning neutron star averaged over the source position and orientation. Thus \( h_{ul} \) and \( \rho_{ul} \) are related by the following formula:

\[
 h_{ul}(f) = \frac{5}{2} \sqrt{\frac{S(f)}{T_0}} \rho_{ul}, 
\]

where \( S(f) \) is one-sided spectral density at frequency \( f \). We have used equations (22) and (23) to obtain upper limits in 0.1 Hz bands over the bandwidth \( 922.2 \text{ Hz} \) that we have searched. The upper limit results are presented in figure 5. Assumed Gaussian noise, we have chosen the confidence \( \mathcal{P} = 90\% \) and we denote the upper limits by \( h_{ul}^{90\%} \). Our best upper limit is equal to \( 3.4 \times 10^{-23} \) at a frequency of 922.55 Hz. Using our 1σ rms error of the amplitude power spectrum estimate we reckon that our upper limit has likewise an error of 18%.
Acknowledgments

We would like to thank Maria Alessandra Papa for making available to us the large computing resources of the Albert-Einstein-Institut, enabling us to complete the data analysis for this paper in a reasonable time. We would also like to thank the Interdisciplinary Center for Mathematical and Computational Modelling of Warsaw University for computing time. AK would like to acknowledge the financial support of Istituto Nazionale di Fisica Nucleare (INFN), the INFN Section in University of Rome ‘La Sapienza’ and the Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut). HP would like to acknowledge the support of the IMPRS on Gravitational Wave Astronomy.

References

[17] Prix R 2007 Class. Quantum Grav. 24 5481–90