A Bayesian method to set upper limits on the strength of a periodic gravitational wave signal from the remnant of SN1987A: possible applications in LIGO searches.

Richard Umstätter[1], Renate Meyer[1], Nelson Christensen[2]

[1]Department of Statistics, University of Auckland, Auckland, New Zealand
[2]Physics and Astronomy, Carleton College, Northfield, MN 55057, USA

Abstract. We present a method that assesses the theoretical detection limit of a Bayesian Markov chain Monte Carlo search for a periodic gravitational wave signal emitted by a neutron star. Inverse probability yields an upper limit estimate for the strength when a signal could not be detected in an observed data set. The proposed method is based on Bayesian model comparison that automatically quantifies Occam's Razor. It limits the complexity of a model by favoring the most parsimonious model that explains the data. By comparing the model with a signal from a pulsar to the null model that assumes solely noise, we derive the detection probability and an estimate for the upper limit that a search, for example, for a narrow-band emission for SN1987a, might yield on data at the sensitivity of LIGO data for an observation time of one year.

PACS numbers: 04.80.Nn, 02.70.Uu.

1. Introduction

Several mechanisms have been proposed that would cause rapidly rotating neutron stars to emit quasi-periodic gravitational waves [1, 2]. Interferometric gravitational wave detectors that are now operating in numerous locations around the world [3, 4, 5, 6] now allow for their verification and much work has gone into the development of dedicated search algorithms for these signals. Radio observations can provide the sky location, rotation frequency and spin-down rate of known pulsars. The frequency of the reported remnant of SN1987a for example is not known accurately [7] but Markov Chain Monte Carlo (MCMC) methods [8, 9, 10, 11, 12] are able to search a range of frequencies (and other physical parameters) in a reasonable time.

As in previous studies [13, 14] the signal under consideration is one that is expected from a non-precessing triaxial neutron star. The gravitational wave signal from such an object is at $f = 2f_{\rm r}$ twice its rotation frequency f_r, and we characterize the amplitudes of each polarization with overall strain factor, h_0. The measured gravitational wave signal will also depend on the antenna patterns of the detector for the 'cross' and 'plus' polarizations, $F_{\times,+}(t;\psi,\alpha,\delta)$, giving a signal $s(t) = F_+(t;\ldots)h_0(1+\cos^2\iota)\cos\Phi(t;\ldots)/2+$

$F_\times(t;\ldots)h_0\cos\iota\sin\Phi(t;\ldots)$ where ι is the inclination angle. The antenna pattern of the detector depends on time t, the polarization angle ψ and its location determined by right ascension α and declination angle δ. The location is assumed to be known from, for example, radio observations. A simple slowdown model [15] provides the phase evolution of the signal as

$$\Phi(t;\mathbf{n},f,\dot{f}) = \phi_0 + 2\pi\left[f(T_{(\alpha,\delta)}-T_0)+\dot{f}(T_{(\alpha,\delta)}-T_0)^2/2\right], \tag{1}$$

where

$$T_{(\alpha,\delta)} = t+\delta t = t+\frac{\mathbf{r}\cdot\mathbf{n}}{c}+\Delta T \tag{2}$$

is the time of arrival of the signal at the solar system barycenter when t is the time at the detector. Here, ϕ_0 is the phase of the signal at a fiducial time T_0, $\mathbf{r}$ is the position of the detector with respect to the solar system barycenter, $\mathbf{n}$ is a unit vector in the direction of the neutron star (depending on α and δ), c is the speed of light and ΔT contains the relativistic corrections to the arrival time [16].

If f, $\dot{f}$, and $\mathbf{n}$ are known from radio observations, for instance, the signal can be *heterodyned* by multiplying the data by $\exp[-i\Phi(t;\mathbf{n},f,\dot{f})]$, low-pass filtered and resampled, so that the only time varying quantity remaining is the antenna pattern of the interferometer. The reference sky location is also needed for the heterodyning process prior to the MCMC simulation. We are left with a simple model with four unknown parameters h_0, ψ, ϕ_0, and ι. If there is an uncertainty in the frequency and frequency derivative two additional parameters come into play, the differences between the signal and heterodyne frequency and frequency derivatives, Δf and $\Delta\dot{f}$. The unit vector $\mathbf{n}$ points to the right ascension α and declination δ of the purported neutron star.

A detailed description of the heterodyning procedure is presented elsewhere [13, 14]. The model of the heterodyned signal of a pulsar has form [14]

$$\begin{aligned} y(t_k;\boldsymbol{a}) = {} & F_+(t_k;\psi,\alpha,\delta)h_0(1+\cos^2\iota)e^{i\Delta\Phi(t_k;\alpha,\delta,\Delta f,\Delta\dot{f})}/4 \\ & - iF_\times(t_k;\psi,\alpha,\delta)h_0\cos\iota e^{i\Delta\Phi(t_k;\alpha,\delta,\Delta f,\Delta\dot{f})}/2, \end{aligned} \tag{3}$$

where t_k is the time of the k^{th} bin and $\boldsymbol{a} = (h_0,\cos\iota,\phi_0,\psi,\Delta f,\Delta\dot{f})$ is a vector of the unknown parameters. $\Delta\Phi(t;\alpha,\delta,\Delta f,\Delta\dot{f})$ represents the residual phase evolution of the signal, equaling $\phi_0+2\pi[\Delta f(T_{(\alpha,\delta)}-T_0)+\Delta\dot{f}(T_{(\alpha,\delta)}-T_0)^2/2]$, where $T_{(\alpha,\delta)}$ (Eq. (2)) depends on the known sky location of the pulsar. Note, that the gravitational wave oscillates at twice the rotation frequency of the pulsar's rotation frequency. Therefore, the frequency in Eq. 3 refers to the gravitational wave frequency. The objective is to fit this model to the data $B_k = y(t_k;\boldsymbol{a})+\epsilon_k$,where ϵ_k is assumed to be normally distributed noise with a mean of zero and known variance σ_k^2. Assuming statistical independence of the binned data points, B_k, the joint likelihood that these data $\boldsymbol{d}=\{B_k\}$ arise from a model with a certain parameter vector $\boldsymbol{a}$ is [14]

$$p(\boldsymbol{d}|\boldsymbol{a}) \propto \prod_k \exp\left[-\left|(B_k-y(t_k;\boldsymbol{a}))/\sigma_k\right|^2\right]/2 = \exp\left[-\chi^2(\boldsymbol{a})/2\right], \tag{4}$$

where

$$\chi^2(\boldsymbol{a}) = \sum_k |B_k-y(t_k;\boldsymbol{a})|^2/\sigma_k^2. \tag{5}$$

In order to draw any inference on the unknown parameter vector $\boldsymbol{a}$ the posterior probability of $\boldsymbol{a}$ given $\boldsymbol{d}$ is needed, which can be obtained from the likelihood via an application of Bayes' theorem. The unnormalized posterior density $p(\boldsymbol{a}|\boldsymbol{d}) \propto p(\boldsymbol{a})p(\boldsymbol{d}|\boldsymbol{a})$ is the product of the prior density of $\boldsymbol{a}$, $p(\boldsymbol{a})$, and the joint likelihood, $p(\boldsymbol{d}|\boldsymbol{a})$. In this study uniform priors distributions are used with prior ranges $[0, 2\pi]$, $[-\pi/4, \pi/4]$ and $[-1, 1]$ for the angle parameters ϕ_0, ψ and $\cos\iota$ respectively.

For h_0, a uniform prior is specified with boundary $[0, 10^{-20}]$. For the frequency and spin down uncertainty, suitable uniform priors are used with ranges of $[-\frac{1}{120}, \frac{1}{120}]$ Hz and $[-10^{-9}, 10^{-9}]\,\mathrm{Hz\,s^{-1}}$ for Δf and $\Delta\dot{f}$, respectively, as applied in [10]. The normalized posterior density $p(\boldsymbol{a}|\boldsymbol{d}) = p(\boldsymbol{a})p(\boldsymbol{d}|\boldsymbol{a})/p(\boldsymbol{d})$ cannot be evaluated analytically, therefore Monte Carlo methods are used here to explore $p(\boldsymbol{a}|\boldsymbol{d})$, as described in [10].

When the signal-to-noise ratio (SNR) and hence the signal's evidence declines, it becomes increasingly difficult to sample efficiently from the posterior distribution using MCMC. The major problem lies in the frequency parameters Δf and $\Delta\dot{f}$. Long integration periods yield narrow posterior modes and when the SNR is small, their occurrence is also negligible with most of the posterior probability mass spread over the entire parameter space determined by the prior distribution. The sampling process of an MCMC sampler becomes inefficient in covering that part of the parameter space where the signal is concentrated. The question that will be addressed in this paper is the threshold of the SNR for which MCMC sampling becomes ineffective and below which no signal parameters can be retrieved.

2. The detection of weak signals

The presence of a signal within the data can be assessed by a formal Bayesian model comparison of the model the contains a signal with the null model that contains no signal. Bayes factors could be applied but they require a properly converged MCMC output. Without the need of MCMC samples, this paper aims to give theoretical detection probabilities dependent on signal-to-noise ratios.

2.1. Derivation of a theoretical detection probability

For the Bayesian Information Criterion (BIC), also called the Schwarz criterion, there is no particular need for the MCMC output samples. The BIC is defined [17] as

$$\mathrm{BIC} = -2\log(\text{maximum likelihood}) + \mathcal{P}, \tag{6}$$

where the penalty term $\mathcal{P} = d\log n$ brings in the number of $d = 6$ independent parameters that describe the model, and the number n of data samples. The penalty term penalizes the number of parameters in a model in order give preference to simpler models and meet the principle of Occam's Razor.

The objective is to derive a theoretical limit for the detection of a signal within a data set observed during a determined observation period at a certain noise level. This

section is dedicated to find a distribution of the BIC depending on the noise, conditioned on the parameters of a potential pulsar.

The observation period is a vector $\mathrm{OP} = (t_1, \ldots, t_n)'$ of n time points t_k with $k \in \{1, \ldots, n\}$ during which the data has been collected starting from t_{start} and ending at t_{end}. The noise vector is a vector $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)'$ for the n data bins. Given the *true* parameter vector of the pulsar from which the signal arises, the full information needed for a detection is determined by the vector $\boldsymbol{a}_* = (h_0^*, \cos\iota^*, \psi^*, \alpha^*, \delta^*, \Delta f^*, \Delta \dot{f}^*, \boldsymbol{\sigma}, \mathrm{OP})'$. Although some parameters like the sky location are expected to be known, they are essential factors for the detection probability in connection with the observation period and the noise. These are essential parts of the parameter vector as the detection depends significantly on them.

A signal *detection* depends on the actual evidence of the model that assumes the presence of a signal from a pulsar within an arbitrary data set when compared to the null model of mere noise. Each potential data set under consideration is based on the *true* parameters of a potential pulsar. Therefore each model comparison is conditioned on a data set $\boldsymbol{d}_*$ that is conditioned on the parameter vector $\boldsymbol{a}_*$. This fact can be used to obtain, for large sample sizes, an approximation for the maximum likelihood value since the maximum likelihood estimate (MLE) is asymptotically consistent and efficient under certain regularity conditions that are generally satisfied [18]. Thus the estimates converge to the true values for large samples sizes. The sample sizes that we expect are in fact in the range of tens of thousands.

A potential data set $\boldsymbol{d}_*$ from a pulsar, based on a *true* parameter vector $\boldsymbol{a}_*$ is modeled by $\mathcal{M}_* : \boldsymbol{d}_*^{(k)} = y(t_k; \boldsymbol{a}_*) + \epsilon_k$ with noise vector ϵ_k. Due to the fact that $\boldsymbol{d}_*$ is conditioned on $\boldsymbol{a}_*$, an approximate maximum log-likelihood under model $\mathcal{M}_1$ is

$$\log \mathrm{ML}_{\boldsymbol{d}_*, \boldsymbol{a}_*, \mathcal{M}_1} \approx -\chi^2_{\boldsymbol{d}_*, \boldsymbol{a}_*, \mathcal{M}_1}(\boldsymbol{a}_*)/2 = -\sum_k \frac{|\epsilon_k|^2}{2\sigma_k^2}, \tag{7}$$

This term comprises the sum of the squared residuals as the model is fitted by the *true* parameter vector. On the other hand, under model $\mathcal{M}_0$ that encompasses no parameters, the log-likelihood has a constant value and therefore its maximum is

$$\log \mathrm{ML}_{\boldsymbol{d}_*, \boldsymbol{a}_*, \mathcal{M}_0} = -\chi^2_{\boldsymbol{d}_*, \boldsymbol{a}_*, \mathcal{M}_0}/2 = -\sum_k \frac{|y_k(t_k; \boldsymbol{a}_*) + \epsilon_k|^2}{2\sigma_k^2}, \tag{8}$$

where the summation term contains the true and given parameter vector of the signal. It is clear that $\log \mathrm{ML}_{\boldsymbol{d}_*, \boldsymbol{a}_*, \mathcal{M}_1} \geq \log \mathrm{ML}_{\boldsymbol{d}_*, \boldsymbol{a}_*, \mathcal{M}_0} \forall \boldsymbol{a}_*$. As a result of this, naturally model $\mathcal{M}_1$ has to be preferred at all times. This, however, does not take into account the penalty term that comes into play due to the principle of Occam's razor. Equality of Eq. 8 and 7 can only be achieved for a zero amplitude h_0^* in parameter vector $\boldsymbol{a}_*$. But how large do we have to choose this amplitude, also considering other influential parameters, in order to justify model $\mathcal{M}_1$ with its many more parameters? This is the essential idea behind this model comparison approach and the penalty terms play a key role in it.

We aim to compare model $\mathcal{M}_0$ and $\mathcal{M}_1$ conditioned on the data set $\boldsymbol{d}_*$, conditioned on a potential pulsar characterized by the *true* parameter vector $\boldsymbol{a}_*$. By substituting Eq. (7) and Eq. (8) into Eq. (6), we obtain

$$\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*,\mathcal{M}_0} = -2\log \mathrm{ML}_{\boldsymbol{d}_*,\boldsymbol{a}_*,\mathcal{M}_0} \tag{9}$$

as model $\mathcal{M}_0$ has $d = 0$ parameters and

$$\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*,\mathcal{M}_1} = -2\log \mathrm{ML}_{\boldsymbol{d}_*,\boldsymbol{a}_*,\mathcal{M}_1} + \mathcal{P}. \tag{10}$$

With respect to $\mathcal{M}_0$ and $\mathcal{M}_1$, a probability for model $\mathcal{M}_1$ can be derived by

$$p(\mathcal{M}_1|\boldsymbol{d}_*,\boldsymbol{a}_*) = \left(1 + e^{\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*}/2 - \log p(\mathcal{M}_1) + \log p(\mathcal{M}_0)}\right)^{-1} \tag{11}$$

Here, $p(\mathcal{M}_0)$ and $p(\mathcal{M}_1)$ are prior probabilities for $\mathcal{M}_0$ and $\mathcal{M}_1$ respectively. The interested reader is referred [19] for a more detailed derivation. We will address different prior scenarios later but for now, we choose equal probabilities $p(\mathcal{M}_0) = p(\mathcal{M}_1) = 0.5$ for the models as a natural choice when there is no prior information about the possible existence of a signal. This yields $p(\mathcal{M}_1|\boldsymbol{d}_*,\boldsymbol{a}_*) = \left(1 + e^{\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*}/2}\right)^{-1}$ where $\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*} := \mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*,\mathcal{M}_1} - \mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*,\mathcal{M}_0}$. It represents the probability that the data $\boldsymbol{d}_*$ from a potential pulsar with given parameter vector $\boldsymbol{a}_*$ is better modeled by $\mathcal{M}_1$ (a signal) rather than $\mathcal{M}_0$ (no signal). In other words it is the probability for the existence of a signal in the data that is emitted by a pulsar with parameter vector $\boldsymbol{a}_*$. It is merely the difference of the two BIC values under consideration that is responsible for a signal detection. A difference of zero for example would yield a 50% probability for both models. A probability conditioned on data $\boldsymbol{d}_*$ from the vector $\boldsymbol{a}_*$, can be expressed as

$$p(\mathcal{M}_1|\boldsymbol{a}_*) = \mathrm{E}\left[p(\mathcal{M}_1|\boldsymbol{d}_*,\boldsymbol{a}_*)|\boldsymbol{a}_*\right] = \mathrm{E}\left[(1 + e^{\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*}/2})^{-1}|\boldsymbol{a}_*\right]. \tag{12}$$

There is no simple way to solve this expression analytically and although feasible, a Monte Carlo sampling process would be lengthly. From a physical perspective, phase ϕ_0 and the frequency parameters Δf, $\Delta\dot{f}$ should have no impact on the actual signal detection as the SNR mainly depends on the amplitude h_0^*, inclination $\cos\iota^*$, noise $\boldsymbol{\sigma}$, and observation time OP. To a smaller extent the SNR is also influenced by the course of the antenna pattern over the observation time OP with parameters ψ^*, α^*, and δ^*. We assume the sky location to be known and condition on α^* and δ^*.

The probability $p(\mathcal{M}_1|\boldsymbol{a}_*)$ is determined by the distribution of $\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*}$. Thus the characteristics of $\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*}$ will be derived below. By using equations Eq. (7), Eq. (8), Eq. (9), Eq. (10) we obtain

$$\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*} \approx \sum_k |\epsilon_k|^2/\sigma_k^2 + \mathcal{P} - \sum_k |y_k(\boldsymbol{a}_*) + \epsilon_k|^2/\sigma_k^2 \tag{13}$$

In [14], white Gaussian noise $\epsilon_{k,\mathrm{re}}, \epsilon_{k,\mathrm{im}} \sim N(0,\sigma_k^2)$ is assumed where the σ_k^2 are estimated for each bin from the noise floor in a 4 Hz band of data around the signal frequency. By substituting $y(t_k;\boldsymbol{a}_*)$ of Eq. 3 and defining some abbreviations, $F_{+,\times}(t_k;\psi^*,\alpha^*,\delta^*) := F_k^{+,\times}$, $e^{i\Delta\Phi(t_k;\alpha^*,\delta^*,\Delta f^*,\Delta\dot{f}^*)} := e^{i\Delta\Phi_k} = \cos(\Delta\Phi_k) + i\sin(\Delta\Phi_k)$,

$\frac{1}{4}h_0^*(1+\cos^2\iota^*) =: A^+$, and $\frac{1}{2}h_0^*\cos\iota^* =: A^\times$ we can rewrite Eq. (13) as

$$\begin{aligned}\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*} \approx \mathcal{P} &- \sum_k \left[(A^+F_k^+)^2/\sigma_k^2 + (A^\times F_k^\times)^2/\sigma_k^2\right] \\ &- 2\sum_k \left(\left[A^+F_k^+\cos(\Delta\Phi_k) + A^\times F_k^\times \sin(\Delta\Phi_k)\right]/\sigma_k\right)\epsilon_{k,\mathrm{re}}/\sigma_k \\ &- 2\sum_k \left(\left[A^+F_k^+\sin(\Delta\Phi_k) - A^\times F_k^\times \cos(\Delta\Phi_k)\right]/\sigma_k\right)\epsilon_{k,\mathrm{im}}/\sigma_k. \end{aligned} \tag{14}$$

The quadratic noise terms cancel out and we are left with normally distributed terms. Given a pulsar with parameter vector $\boldsymbol{a}_*$, the $\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*}$ is thus normally distributed. The terms that contain the phase evolution canceled out as well and Eq. (14) is thus independent of the parameters ϕ_0, Δf, and $\Delta\dot{f}$. With $\epsilon_{k,\mathrm{re}}, \epsilon_{k,\mathrm{im}} \sim N(0,\sigma_k^2)$ we have $\mathrm{E}(\epsilon_{k,\mathrm{re}}/\sigma_k) = \mathrm{E}(\epsilon_{k,\mathrm{im}}/\sigma_k) = 0$ and the expected value of Eq. (14) has the form

$$\mu_{\boldsymbol{a}_*} := \mathrm{E}(\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*}) = \mathcal{P} - \sum_k \sigma_k^{-2}\left[(A^+F_k^+)^2 + (A^\times F_k^\times)^2\right]. \tag{15}$$

Eq. (15) yet allows some insight as it tells us that for a given arbitrary parameter vector $\boldsymbol{a}_*$, model $\mathcal{M}_0$ would be preferred over $\mathcal{M}_1$, if $\mu_{\boldsymbol{a}_*} > 0$. Given a parameter vector $\boldsymbol{a}_*$, the variance of Eq. (14) is

$$\sigma^2_{\boldsymbol{a}_*} = \mathrm{Var}(\Delta\mathrm{BIC}_{\boldsymbol{d}_*,\boldsymbol{a}_*}) = 4(\mathcal{P} - \mu_{\boldsymbol{a}_*}). \tag{16}$$

Both expressions Eq. (15) and Eq. (16) only depend on the five parameters h_0^*, $\cos\iota^*$, ψ^*, α^*, and δ^*. The parameters ψ^*, α^*, and δ^* only enter in the plus and cross polarization terms F_k^+ and $F_k^\times$ of the antenna pattern which depends on the orientation sweep of the interferometer towards the pulsar and the polarization angle of the gravitational wave that it emits.

We are left with the random variable $\Delta\mathrm{BIC}|\boldsymbol{a}_* \sim \mathrm{N}(\mu_{\boldsymbol{a}_*}, \sigma^2_{\boldsymbol{a}_*})$ that depends on five parameters of the pulsar plus noise $\boldsymbol{\sigma}$ and observation period OP. If we assume constant noise σ over time, we can combine h_0^* and σ to a more handy SNR h_0^*/σ parameter. We define a new vector $\boldsymbol{a}_\bullet = (h_0^*/\sigma, \cos\iota^*, \psi^*, \alpha^*, \delta^*, \mathrm{OP})'$ with observation period $\mathrm{OP} = (t_1, \ldots, t_n)'$. Explicitly, the difference in the BIC values with respect to models $\mathcal{M}_0$ and $\mathcal{M}_1$, for arbitrary data sets, conditioned on $\boldsymbol{a}_\bullet$ follow the distribution $\Delta\mathrm{BIC}|\boldsymbol{a}_\bullet \sim \mathrm{N}(\mu_{\mathrm{BIC},\boldsymbol{a}_\bullet}, \sigma^2_{\mathrm{BIC},\boldsymbol{a}_\bullet})$ with

$$\mu_{\mathrm{BIC},\boldsymbol{a}_\bullet} = \mathcal{P} - \left(\frac{h_0^*}{\sigma}\right)^2\left(\left[\frac{1}{4}(1+\cos^2\iota^*)\right]^2\sum_k\left(F_k^+\right)^2 + \left[\frac{1}{2}\cos\iota^*\right]^2\sum_k\left(F_k^\times\right)^2\right) \tag{17}$$

and

$$\sigma^2_{\mathrm{BIC},\boldsymbol{a}_\bullet} = 4\left(\mathcal{P} - \mu_{\mathrm{BIC},\boldsymbol{a}_\bullet}\right). \tag{18}$$

Using these information, Monte Carlo methods can be used to estimate Eq. (12).

As an example, we consider a data set that we can expect was taken over one year at the three LIGO interferometers Hanford (4km, 2km) and Livingston (4km) with three different noise levels at the three interferometers. A sensible heterodyning frequency for the SN1987a remnant is $f = 2f_\mathrm{r} = 935\mathrm{Hz}$ [7]. For the purpose of illustrating an example we will assume rough noise levels that are likely to be close

to LIGO's S5 values [20] (LIGO Document G060293-01) at the frequency in question using the relation to the noise power spectral density as applied in [12]. Thus, as an example, we choose the noise levels 8×10^{-24} (Hanford 4km), 1.5×10^{-23} (Hanford 2km), and 9×10^{-24} (Livingston 4km) at that frequency and an observation period OP of one year of S5 data that would be heterodyned to a potential source at $\alpha^* = 5^{\mathrm{h}}\,35^{\mathrm{m}}\,28.03^{\mathrm{s}}$ and $\delta^* = -69^\circ\,16'\,11.79''$ (SN1987a) with 525600 bins at one sample per minute. The data are analyzed for each interferometer separately and also combined by the sum of the log-likelihoods, as we assume independence. The parameter vector encompasses $\boldsymbol{a}_\bullet = \left(h_0^*/\sigma, \cos\iota^*, \psi^*, \alpha^* = 5^{\mathrm{h}}\,35^{\mathrm{m}}\,28.03^{\mathrm{s}}, \delta^* = -69^\circ\,16'\,11.79'', \mathrm{OP}\right)$ in which the values of h_0^*/σ and $\cos\iota^*$ and ψ^* are unknown. In order to derive a probability conditioned on h_0^*/σ, we need to marginalize $p(\mathcal{M}_1|\boldsymbol{a}_\bullet)$ over $\cos\iota^*$ and ψ and obtain $p(\mathcal{M}_1|h_0^*/\sigma, \alpha^*, \delta^*, \mathrm{OP}) = \int p(\mathcal{M}_1|\boldsymbol{a}_\bullet) dp_{\cos\iota}(\cos\iota^*) dp_\psi(\psi^*)$.

Fig. 1 displays the probability of a signal detection as a function of the amplitude. Two different prior probabilities on the signal existence are chosen. The natural choice is $p(\mathcal{M}_1) = 0.5$ when there is no information available. However, we know that we focus only a 1/60Hz band and the probability of an existence needs to be split on the frequency bands in which we expect a signal. In addition, we do not know whether there is a neutron star at all which lowers the probability further. For this reason, we chose a rather arbitrary and extremely small probability of $p(\mathcal{M}_1) = 10^{-9}$ in order to asses the impact of that prior probability. We obtain the graph shown in Fig. 1.

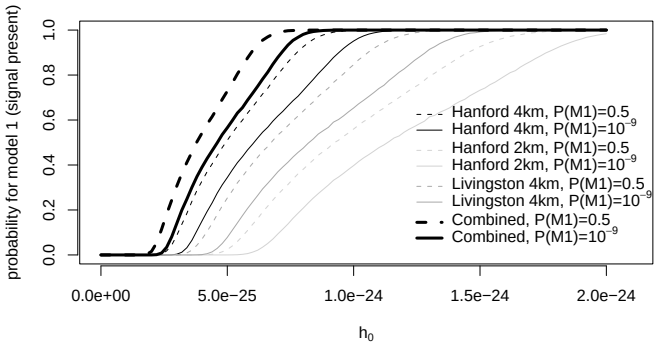

Figure 1. Signal detection probability for the three different interferometers and two different prior probabilities for model $\mathcal{M}_1$ as a function of the amplitude h_0^* for one year of S5 data. The curves of two prior probabilities $p(\mathcal{M}_1) = 0.5$ (dashed lines) and $p(\mathcal{M}_1) = 10^{-9}$ (solid lines) are shown.

A larger amplitude h_0^* is required for a successful detection when we doubt the existence of a signal. Hence, the data must speak more clearly for a signal in order to overcome the low prior probability but since the observation period of one year is rather long, the effect of the prior probability is fairly small.

All graphs compiled so far are showing a signal detection probability given a particular scenario but the question we aim to answer in the next section is how strong a signal still can be even if a signal can't be seen.

2.2. Performance of the Bayesian MCMC search in setting an upper limit using S5 data

The upper limit estimate for a Bayesian MCMC search involves testing the hypothesis $h_0^* < \mathrm{UL}$ vs. $h_0^* \geq \mathrm{UL}$ under the assumption $\mathcal{M}_0$ that there is no signal in the data. The derivation of the probability $p(h_0^* < \mathrm{UL}|\mathcal{M}_0)$ will shed light on this matter. We condition on noise, observation period, and location and after integrating over the prior distributions of $\cos\iota^*$ and ψ^*, $p(\mathcal{M}_0|h_0^*, \mathrm{OP}, \boldsymbol{\sigma}, \alpha^*, \delta^*) := \int\int p(\mathcal{M}_0|\boldsymbol{a}_\bullet) dp_{\cos\iota}(\cos\iota^*) dp_\psi(\psi^*)$, we obtain

$$p(h_0^* < \mathrm{UL}|\mathcal{M}_0, \mathrm{OP}, \boldsymbol{\sigma}, \alpha^*, \delta^*) = \frac{\int_0^{\mathrm{UL}} p(\mathcal{M}_0|h_0^*, \mathrm{OP}, \boldsymbol{\sigma}, \alpha^*, \delta^*) p(h_0^*) dh_0^*}{\int_0^{\infty} p(\mathcal{M}_0|h_0^*, \mathrm{OP}, \boldsymbol{\sigma}, \alpha^*, \delta^*) p(h_0^*) dh_0^*}. \quad (19)$$

In order to derive Eq. (19) we need to find a suitable prior for $p(h_0^*)$. One choice could be to put a uniform prior on h_0 with large boundary $[0, 10^{-20}]$. The upper boundary of the prior range has negligible impact on the results of Eq. (19) as long as this boundary is significantly larger then the upper limit estimate. Fig. 2 displays Eq. (19) for two different prior probabilities on whether we expect a signal at SN1987a.

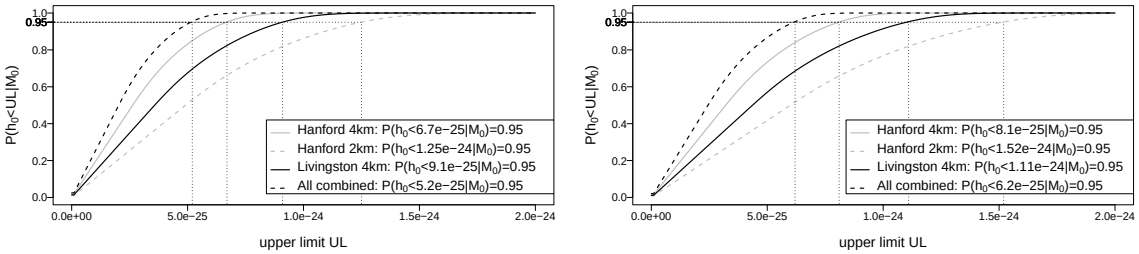

Figure 2. Estimated sensitivity of the Bayesian method described in this paper, assuming one year of data with the noise level likely to be close to the LIGO interferometers during their S5 run [20]. The model prior probabilities are $p(\mathcal{M}_1) = 0.5$ (left) and $p(\mathcal{M}_1) = 10^{-9}$ (right). The prior for h_0 is $h_0 \sim \mathrm{Unif}(0, 10^{-20})$. The assumed noise levels are $\sigma_{\mathrm{H}_1} = 8 \times 10^{-24}$, $\sigma_{\mathrm{H}_2} = 1.5 \times 10^{-23}$, and $\sigma_{\mathrm{L}_1} = 9 \times 10^{-24}$.

Since we focus our search on a possible pulsar in SN1987a, we can tailor a prior distribution for h_0 as we know the age of SN1987a and its distance. In [21] it is assumed that a newly formed neutron star spins at high rate and gravitational radiation slows it down. Two different prior scenarios are conceived here. According to [21], it is

$$h_0(f) = r^{-1}\sqrt{(5GI_{zz})/(8c^3\tau_{\mathrm{gw}}(f))}, \quad (20)$$

where $\tau_{\mathrm{gw}}(f)$ is the time for the gravitational wave frequency to drift down to frequency f from its original spin rate. In case of SN1987a it is 20 years. Here, G is Newton's constant, c the speed of light, r the distance to the neutron star, and I_{zz} the principal moment of inertia about the rotation axis. In order to derive a prior distribution for h_0 we need to determine prior distributions for r and I_{zz}. We assume the distance estimated in [22] with 50.9 ± 1.8kpc with $r \sim \mathrm{N}(50.9, 1.8^2)$kpc for accounting the uncertainty in the distance. For the moment of inertia, we choose a uniform prior within the range

$[10^{38}, 3 \times 10^{38}]$ as applied in [23]. These considerations yield a prior for h_0 as shown later in Fig. 3.

A totally different approach for obtaining a prior distribution for h_0 is by [13]

$$h_0 = 4\pi^2 G I_{zz} f^2 \epsilon / (c^4 r) \tag{21}$$

for a general pulsar expected at SN1987a. Here, f is the pulsar's rotation frequency, and ϵ its ellipticity. We heterodyne to a frequency of $f = 2f_{\rm r} = 935$Hz, and assume the gravitational wave frequency to have this value within a 1/60Hz frequency band for a particular search. An uncertainty beyond this needs to be accounted for in the prior $p(\mathcal{M}_1)$ for the existence of the signal within the 1/60Hz band around f because the signal is not seen outside that band after the heterodyning process.

We use the same uniform prior for I_{zz} as above but we have to find a suitable prior for the ellipticity. In [24], the ellipticity is assumed to have an exponential distribution (maximum entropy prior) with cut-off at a maximum ellipticity threshold. Although in [24] more pessimistic mean and maximum values are used, our choice is more optimistic in order to account for the fact that we know that a possible neutron star in SN1987a is very young. We choose a cut-off according to [21] at $\epsilon_{\rm max} \approx 9 \times 10^{-5}$ based on the idea of a hybrid neutron star with a mixed quark and baryon core and a normal neutron star in the outer part. For the mean of the exponential prior distribution we use an optimistic choice of $\epsilon_{\rm mean} = 5 \times 10^{-5}$. Both prior distributions for h_0 as discussed above are displayed in Fig. 3 along with their resulting upper limit estimates.

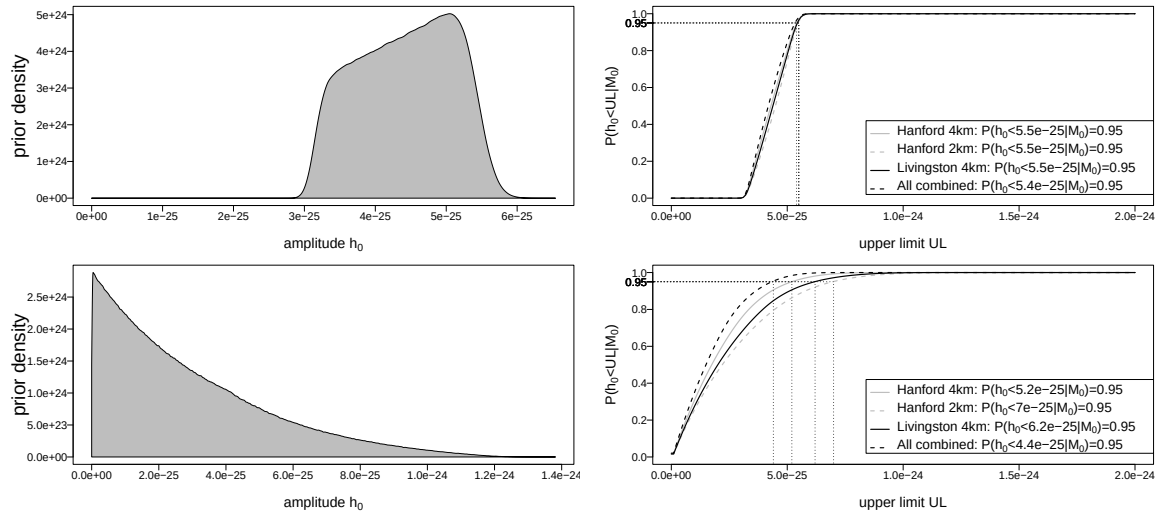

Figure 3. Two different prior distributions (left column) for h_0 for a possible neutron star in SN1987a and the corresponding curves for the Bayesian MCMC upper limit estimates (right column). The upper row corresponds to a prior subject to Eq. (20) whereas the lower row is based on Eq. (21). For the model selection, the prior probability is chosen $p(\mathcal{M}_1) = 10^{-9}$.

The use of such priors changes the results for upper limit estimates compared to those in Fig. 2 (which were based on a uniform prior). In essence, a uniform prior

on the amplitude recovers the detection ability of an interferometer. For example, in case of combined data sets, an upper limit estimate based on a uniform prior requires an amplitude of at least 6.2×10^{-25}. The use of prior distributions based on Eq. (20) and Eq. (21), however, only have 0.001% and 11.4% probability mass above that limit, respectively. This inevitably yields values for the upper limit estimates dominated by the prior of h_0. This is obvious especially in case of the prior based on Eq. (20) and can be seen in Fig. 3.

3. Conclusions

The Bayesian MCMC methods work well when the SNR is sufficiently large but they struggle when the signal is too weak and the parameters that affect the phase evolution are not known. The fact that we integrate over very long observation periods requires an almost exact match of the phase evolution and almost all mass of the posterior distribution is highly concentrated around one point in the parameter space when the SNR is large. Finding this posterior peak with Bayesian MCMC methods is time consuming but once found, the sampling process is easy and efficient. With decreasing SNR, however, the sampler is forced to also sample from other areas of the parameter space determined by the prior. This requires multiple retrievals of the narrow peak and it requires extremely long runs to gain insight into the actual shape of the posterior distribution. The sampling speed depends on observation length and number of Markov chains involved when using parallel tempering. For one year of data, each single chain samples about 150 000 samples per week and chain on a 2.8 GHz machine. At low SNRs at least 10 chains are needed [19]. No sensible inference can be drawn from an MCMC output if no frequency parameters can be retrieved. In those cases, the method derived here, based on model comparison, provides an excellent means for estimating an upper limit for the amplitude of a signal when using Bayesian MCMC methods, given the observation period and noise. In practice this method could be used to estimate the sensitivity of the Bayesian MCMC method on actual S5 data. For long observation periods, the impact of prior information about the presence of a signal is rather small. The influence of the amplitude's prior only becomes significant when the sensitivity, with respect to the obtained data, is too small for the expected amplitudes. Consequently, when we expect amplitudes below the detection limit then the upper limit estimate is determined mainly by the prior distribution of the amplitudes.

Acknowledgments

This work was supported by the Marsden Fund Council from Government funding administered by the Royal Society of New Zealand (Grant UOA-204), and the National Science Foundation grant PHY-0553422.

References

[1] C. Cutler. *Physical Review D*, 66(8):084025, 2002.

[2] L. Bildsten. *Astrophysical Journal*, 501(1):L89–L93, 1998.

[3] J. Hough et al. In Tsubono et al. [25], pages 175–182.

[4] B. C. Barish et al. In Tsubono et al. [25], pages 155–161.

[5] A. Brillet et al. In Tsubono et al. [25], pages 163–173.

[6] K. Tsubono et al. In Tsubono et al. [25], pages 183–191.

[7] J. Middleditch et al. *New Astronomy*, 5(5):243–283, 2000.

[8] N. Christensen, R. Meyer, and A. Libson. *Classical and Quantum Gravity*, 21:317–330, 2004.

[9] N. Christensen, R. J. Dupuis, G. Woan, and R. Meyer. *Physical Review D*, 70(2):022001–1, 2004.

[10] R. Umstätter, R. Meyer, R.J. Dupuis, J. Veitch, G. Woan, and N. Christensen. *Classical and Quantum Gravity*, 21:S1655–S1665, 2004.

[11] R. Umstätter, R. Meyer, R.J. Dupuis, J. Veitch, G. Woan, and N. Christensen. In *AIP Conference Proceedings - Bayesian inference and Maximum Entropy Methods in Science and Engineering: 24th*, volume 735, pages 336–343. American Institute of Physics, 2004. International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering.

[12] J. Veitch, R. Umstätter, R. Meyer, N. Christensen, and G. Woan. *C.Q.G.*, 22:S995–S1001, 9 2005.

[13] B. Abbott et al. *Physical Review D*, 69(8):082004–1–082004–16, 2004.

[14] R. J. Dupuis and G. Woan. *Physical Review D*, 72(10):102002, November 2005.

[15] P. Jaranowski, A. Krolak, and B. F. Schutz. *Physical Review D*, 58(6):063001, 1998.

[16] J. H. Taylor. *Reviews of Modern Physics*, 66(3):711–719, 1994.

[17] R. E. Kass and A. E. Raftery. *Journal of the American Stat. Association*, 90(430):773–795, 1995.

[18] G. Casella and R. L. Berger. *Statistical Inference*. Duxbury, Pacific Grove, CA, 2nd edition, 2002.

[19] R. Umstätter. PhD thesis, University of Auckland, 2006.

[20] Ligo Document G060293-01, 2006.

[21] B. Abbott et al. *preprint*, May 2006. arXiv:gr-qc/0605028 v2.

[22] N. Panagia et al. In *Bulletin of the American Astronomical Society*, page 1243, December 1997.

[23] B. Abbott et al. *preprint*, February 2007. arXiv:gr-qc/0702039 v1.

[24] C. Polomba. *Classical and Quantum Gravity*, 22:S1027–S1039, 2005.

[25] K. Tsubono, M.-K. Fujimoto, and K. Kurodo, editors. Tokyo, 1997. Universal Academic Press.