Data analysis of continuous gravitational wave: all-sky search and study of templates

D. C. Srivastava*† and S. K. Sahay*

Department of Physics, DDU Gorakhpur University, Gorakhpur-273009, UP, India

Accepted 2002 September 3. Received 2002 September 3; in original form 2001 November 23

ABSTRACT

We have studied the problem of an all-sky search in reference to a continuous gravitational wave particularly for such sources whose wave-form are known in advance. We have made an analysis of the number of templates required for matched-filter analysis as applicable to these sources. We have employed the concept of fitting factor (FF); treating the source location as the parameters of the signal manifold and have studied the matching of the signal with templates corresponding to different source locations. We have investigated the variation of FF with source location and have noticed a symmetry in template parameters, \( \theta \) and \( \phi \). It has been found that the two different template values in source location, each in \( \theta \) and \( \phi \), have the same FF. We have also computed the number of templates required assuming the noise power spectral density \( S_\text{n}(f) \) to be flat. It is observed that higher FF requires an exponentially increasing, large number of templates.

Key words: gravitational waves – methods: data analysis – pulsars: general.

1 INTRODUCTION

Gravitational wave (GW) laser interferometer antennas are essentially omni-directional with a response better than 50 per cent of the average over 75 per cent of the whole sky (Grishchuk et al. 2001). Hence the data analysis systems will have to carry out all-sky searches for its sources. We know that the amplitude of intense GW believed to be bathing the earth is very small, as compared to the sensitivity of GW detectors, and is further masked by the dominant noise. In these circumstances, continuous gravitational wave (CGW) sources are of prime importance because for such sources we can achieve enhanced signal-to-noise (S/N) ratios by investigating longer observation data sets. However, a long observation time introduces modulation effects, arising as a result of the relative motion of the detector and the source. As a consequence, the power in the forest of side bands is redistributed, resulting in the reduction of the expected power due to amplitude modulation (AM). The problem of all-sky search gains another dimension in view of the fact that there are reasons to believe the presence of intense GW sources whose locations and even frequencies are not known. Amongst such sources pulsars occupy an important position. In a similar way to an all-sky search, one will also have to do an all-frequency search. An all-sky, all-frequency search is the holy grail of gravitational pulsar astronomy. In this paper we confine ourselves to the problem of all-sky search.

 Searches of CGW without a priori knowledge appears to be quite computationally demanding even by the standard computers expected to be available at in the near future. For example, in the case of a bandwidth of \( 10^3 \) Hz, an observation time \( 10^7 \) s and a star’s minimum decay time of \( 100 \) yr, one would require a \( 10^{14} \) Tllops computer (Frasca 2000). Very fast computers and large memories with an ample amount of disc space seems inevitable. However, a choice of optimal data processing and a clever programming is also an integral part of a solution to this problem. Amongst these the pre-correction of time-series due to the Doppler modulation before the data is processed may be a method, which will reduce the computational requirements. In reference to this, Schutz (1991) has introduced the concept of patch in the sky as the region of space throughout which the required Doppler correction remains the same. He has also demonstrated that the number of patches required for \( 10^7 \) s observation data set and 1-KHz signal would be about \( 1.3 \times 10^3 \). However, the size of the patch would also depend on the data analysis technique being employed.

 Matched filtering is the most suitable technique for the detection of signals from sources viz., pulsars whose wave form is known. The wave forms are used to construct a bank of templates, which represent the expected signal wave form with all possible ranges of its parameters. The time of arrival, source location, frequency of the signal, ellipticity of the source and its spin down represent important parameters of GW emitted by a pulsar. For detection of GW we check if the cross correlation of the templates with the corresponding data set exceeds the preassigned threshold. We introduce in the next section the criterion of the fitting factor (FF) (Apostolatos 1995) applicable to such analysis. We consider the source location

*E-mail: dcs@gkpu.ernet.in (DCS); ssahay@iucaa.ernet.in (SKS)
†Visiting Associate, Inter University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune-411007, India.
as parameters of the signal manifold and investigate the matching of the waveforms corresponding to different source locations. In Section 3 we compute the number of templates required for all-sky search. A discussion of the results is provided in the Section 4.

2 MATCHED FILTER ANALYSIS: templates

The bank of templates is matched, in practice, to only a discrete set of signals from among the continuum of possible signals. Consequently, it is natural that all the signals will not get detected with equal efficiency. However, it is possible to choose the set of templates judiciously so that all the signals of a given amplitude are detected with a given minimum detection loss. The standard measure for deciding what class of wave form is good enough is the FF. It quantitatively describes the closeness of the true signals to the template manifold in terms of the reduction of S/N ratio arising due to the cross correlation of a signal outside the manifold with the best matching templates lying inside the manifold. If the FF of a template family is unity the signal lies in the manifold. If the FF is less than unity the signal lies outside manifold.

Even if the signal discrete templates lie within the template manifold it would be unlikely that any of the actual templates used would correspond to the signal. The parameters describing the search template (source location, ellipticity, etc.) can vary continuously through out a finite range of values. The set of templates characterized by the continuously varying parameters is of course infinite. However, in practice the interferometer output must be cross correlated with a finite subset of the templates whose parameter values vary in discrete steps from one template to the next. This subset (the discrete template family) has measure zero on the manifold of the full set of possible templates (the continuous template family), so the template which most closely match a signal will generally lie in between the signal and the nearest of the discrete template family. The mismatch between the signal and the nearest of the discrete templates will cause some reduction in S/N ratio. This would mean that the members of the discrete template family must be chosen so as to render acceptable loss of S/N ratio.

The study of templates in time-domain has been made by many research workers (Schutz 1991; Królak & Brady 1997; Brady et al. 1998; Brady & Creighton 2000; Jaranowski et al. 1998; Jaranowski & Królak 2000). However, the analysis in the frequency domain has the advantage of incorporating the spectral noise density of the interferometer. In order to determine the number of templates required for matched filtering analysis, we make use of the formula for FF as given by Apostolatos (1995) i.e.

\[
\text{FF} = \max_{\nu, \phi} \frac{\langle h(f) h_T(f; \theta_T, \phi_T) \rangle}{\sqrt{\langle h_T(f; \theta_T, \phi_T) \rangle^2 \langle h(f) \rangle^2}}
\]

(1)

where \(h(f)\) and \(h_T(f; \theta_T, \phi_T)\) represent respectively the FTs of the actual signal wave form and the templates.

In earlier papers we have made a study of continuous gravitational wave output through Fourier transform (FFT) (Srivastava & Sahay 2002a,b). These papers are referred in the sequel respectively as Papers I and II. It has been established there that the AM of CGW data output results into redistribution of power at four additional frequencies \(f \pm 2f_{\text{rot}}, f \pm f_{\text{rot}}\) in accordance with the frequency modulation (FM). Hence it is sufficient for the analysis of FF to consider only the frequency modulated FT. The results obtained in Paper II regarding FT of frequency modulated data output i.e. (Paper II equations 27 and 28) may be arranged, making use of the symmetry property of the Bessel functions, as follows.

\[
\begin{align*}
\tilde{h}(f) & \simeq \frac{v}{w_{\text{rot}}} \left[ \frac{J_0(Z) J_0(N)}{2v^2} \left[ (\sin(R + Q) - \sin(R + Q - v\xi_\nu)) + i (\cos(R + Q) - \cos(R + Q - v\xi_\nu)) \right] \\
& + J_0(Z) \sum_{n=1}^{\infty} \frac{J_0(N)}{v^2 - m^2} \left[ (\cos(R - \nu) - \cos(R - \nu - v\xi_\nu)) \right] \\
& + \sum_{k=1}^{\infty} \sum_{m=-\infty}^{\infty} e^{iA B (\ell - iD)} \right] ;
\end{align*}
\]

(2)

\[
\begin{align*}
X &= \sin(R + Q - m\tau(2)) \\
Y &= \cos(R + Q - m\tau(2)) \\
U &= \sin\nu_x \cos m(\xi_\nu - \delta) \\
&- \frac{\nu_x}{2} \left[ \sin\nu_x \sin m(\xi_\nu - \delta) - \sin m\delta \right] \\
V &= \cos\nu_x \cos m(\xi_\nu - \delta) \\
&+ \frac{\nu_x}{2} \sin\nu_x \sin m(\xi_\nu - \delta) - \cos m\delta
\end{align*}
\]

(3)

Now it is straightforward to compute FF. To understand the procedure let us consider a source at location \((\theta, \phi) = (25^\circ, 30^\circ)\) emitting frequency \(f_e = 0.5\) Hz. In order to evaluate the summation given in equation (2), let us note that the value of Bessel function decreases rapidly as its order exceed the argument. Accordingly, for the present case it is sufficient to take the ranges of \(k\) and \(m\) as 1 to 1610 and \(-4\) to 4, respectively. We wish to analyse the data set for \(T_e = 1\) sidereal d. To maximize FF over \(\theta\) and \(\phi\), we first maximize over \(\phi\) by fixing \(\phi_T\) to some arbitrarily selected value say, \(\phi_T = \phi = 30^\circ\). Having done this we maximize over \(\theta\) by varying \(\theta_T\) in discrete steps over its entire range i.e. \(0^\circ\) to \(180^\circ\). The results obtained are plotted in Fig. (1). It is remarked that in order to compute the inner product of two waveform \(h_1\) and \(h_2\) which is defined as

\[
\langle h_1 | h_2 \rangle = 2 \int_0^\infty \frac{\tilde{h}_1^*(f) \tilde{h}_2(f)}{S_n(f)} \, df
\]

(4)

\[
\begin{align*}
\nu_e &= \text{One day} \\
T_e &= 0.5 \text{ Hz} \\
\phi &= \phi_T = 30^\circ \\
\theta &= 30^\circ
\end{align*}
\]

Figure 1. Variation of FF with \(\theta_T\).
where $^*$ denotes complex conjugation, $\tilde{a}(f)$ denotes the Fourier transform of the quantity underneath ($\tilde{a}(f) = \int_{-\infty}^{\infty} a(t) \exp(-2\pi if t) \, dt$) and $S_n(f)$ is the spectral density of the detector noise. One would require to integrate the expression over the bandwidth of Doppler modulated signal. The bandwidth may be determined by computing the maximum value of the Doppler shift. In accordance with the present case we consider the bandwidth equal to 0.002 Hz. In a similar manner one may require to integrate the expression over the bandwidth of Doppler shift. In accordance with equation (23) in Paper I, the Doppler shift is $\sim 10^{-4} f_s$. However, for the present case we consider the bandwidth equal to 0.002 Hz. In a similar manner one may require to integrate the expression over the bandwidth of Doppler shift. In accordance with equation (23) in Paper I, the Doppler shift is $\sim 10^{-4} f_s$. However, for the present case we consider the bandwidth equal to 0.002 Hz.

The following points in reference to these plots may be noted.

(i) The FF is unity for $T_\theta = 25^\circ$, $155^\circ$ (Fig. 1) for $T_\phi = 35^\circ$, $145^\circ$ (Fig. 2) and for $T_\phi = 220^\circ$, $320^\circ$ (Fig. 3).

(ii) It is found that the dependence of FF on the template variables $T_\theta$ and $T_\phi$ may be expressed via

$$ FF = e^{-0.00788T_\theta \theta} $$

$$ FF = e^{-0.01778T_\phi \phi} $$

(iii) The oscillatory behaviour is more or less typical in waveform that match well or bad depending on their parameters. However, we are content with the technique we have employed as the region of such artificial facets fall in the region of the $FF < 0.25$.

Finally, we conclude this section by noting the symmetry property of the FF with template parameters. A closer look of the graphs and the remark (i) above reveal the following symmetry property. The FF is symmetrical under following transformations.

$$ T_\theta \rightarrow \pi - T_\theta \quad 0 \leq T_\theta \leq \pi $$

$$ T_\phi \rightarrow \pi - T_\phi \quad 0 \leq T_\phi \leq \pi $$

$$ T_\phi \rightarrow 3\pi - T_\phi \quad \pi \leq T_\phi \leq 2\pi $$

Let us note that these symmetry properties are based on our results for 1-d observation time. The generic nature of the symmetries may be established only after studying the variation of FF with $T_\phi$. Unfortunately, we could not make this analysis because of our limitations on the computational facilities.

3 NUMBER OF TEMPLATES

It is important to study the problem of number of templates for all-sky search in the light of FF. The results of the previous section reveal that the grid spacing $\delta \theta$ in the $\theta$-parameter of templates may be expressed symbolically as a function of FF, $f_s$, $T_\theta$, $\theta$ and $\phi$ as

$$ \delta \theta = F(F, f_s, \theta, \phi, T_\theta). $$

Similarly, we have

$$ \delta \phi = G(F, f_s, \theta, \phi, T_\theta). $$

In the literature it is reported that the choice of grid spacing $\delta \phi$ depends insignificantly on $\phi$, whereas $\delta \theta$ depends as $\sin 2\theta$ (Brady & Creighton 2000). However, in order to arrive at such conclusions one may have to make analyses for different values of $\theta$ and $\phi$. Unfortunately, due to the limited memory and the efficiency of the computer, we could not make a study of this aspect.

In view of this, equations (5) and (6) may be equivalently expressed as

$$ F(F, 0.5, 25^\circ, 30^\circ, T) = \left[-(0.00788)^{-1} \ln(F)\right]^{1/2} $$

$$ G(F, 25, 1^\circ, 35^\circ, T) = \left[-(0.01778)^{-1} \ln(F)\right]^{1/2} $$

For any chosen value of FF one can determine $\delta \theta$ and $\delta \phi$. However, there is no unique choice for it. Our interest would be in the assignment of $\delta \theta$ and $\delta \phi$ such that the spacing is maximum, resulting in the least number of templates. As we mentioned earlier, there is the stringent requirement of reducing computer time. Accordingly, there is a serious need to adopt some procedure/formalism to achieve this. For example, one may adopt the method of a hierarchical search given by Mohanty & Dhurandhar (1996) and Mohanty (1998). This search is carried out in two steps. At the first level, one would start with a template bank that had coarse spacing in the parameter space but with a lower threshold. At the next level, a more finely spaced set...
of templates and a higher threshold would be used but only around those templates of the previous level which exceeded the previous threshold.

However, an important issue related to the problem of the number of templates regards the study of the behaviour of the number of templates with FF for different values of $f_0$ and $T_0$. We have made an investigation of this aspect. We assume a source location $(\theta, \phi) = (1', 30')$. We choose some value of FF, say 0.995. Now we maximize FF over $\phi$ by introducing spacing $\delta \phi$ so as to yield the selected FF. In the case under investigation, $\delta \phi$ is found to equal $4\delta \theta$ to equal 45$\times$10$^{-5}$. Thereafter we maximize over $\phi$ by introducing spacing $\delta \phi$ in the obtained bank of templates and determine the resulting FF. The results obtained may be expressed in the form of a graph, as shown in Figs (4) and (5). Interestingly, the nature of these curves are similar. We have obtained a best fitting graphs obtained for the number of templates assuming the noise power spectrum to be $f_0 T_0$ at, which is justified because the bandwidth of current interest and remains unsolved.

We have noticed marked symmetries in all-sky searches in both $\theta$ and $\phi$ space for 1-d observation time. It has been found that the two different template values in source location, each in $\theta_0$ and $\phi_0$, have the same FF. Accordingly, the computation burden will be reduced by a factor of 4. However, it is not clear whether this symmetry property can be established analytically as well. The source location, because of these symmetries, is uncertain and some other analysis is to be adopted for getting the exact location. We have computed the number of templates assuming the noise power spectral density $S_n(f)$ to be flat, which is justified because the bandwidth is extremely narrow.

The issues of optimum template parametrization and placement, and the related computational burden, have been discussed in the literature by several authors; notably by Sathyaprakash & Dhurandhar (1991), Dhurandhar & Sathyaprakash (1994), Owen (1996), Apostolatos (1995, 1996), Mohanty & Dhurandhar (1994), Owen (1996), Apostolatos (1995, 1996), Mohanty & Dhurandhar (1996), Mohanty (1998) and Owen & Sathyaprakash (1999). The question of possible efficient interpolated representation of the correlators is a problem of current interest and remains unsolved.

4 DISCUSSION

In view of the complexity of the FT, which contains trigonometric functions and the related computational burden, one has to be careful in computing the FF. We have found useful to employ the Romberg integration using Padé approximations. We have used (i) QROMO instead of QROMB, as the former takes care of singularities, and (ii) RATINT routine for Padé approximation (Press et al. 1986).

We have noticed marked symmetries in all-sky searches in both $\theta$ and $\phi$ space for 1-d observation time. It has been found that the two different template values in source location, each in $\theta_0$ and $\phi_0$, have the same FF. Accordingly, the computation burden will be reduced by a factor of 4. However, it is not clear whether this symmetry property can be established analytically as well. The source location, because of these symmetries, is uncertain and some other analysis is to be adopted for getting the exact location. We have computed the number of templates assuming the noise power spectral density $S_n(f)$ to be flat, which is justified because the bandwidth is extremely narrow.

The issues of optimum template parametrization and placement, and the related computational burden, have been discussed in the literature by several authors; notably by Sathyaprakash & Dhurandhar (1991), Dhurandhar & Sathyaprakash (1994), Owen (1996), Apostolatos (1995, 1996), Mohanty & Dhurandhar (1996), Mohanty (1998) and Owen & Sathyaprakash (1999). The question of possible efficient interpolated representation of the correlators is a problem of current interest and remains unsolved.

ACKNOWLEDGMENTS

The authors are thankful to Professor S. Dhurandhar, IUCAA, Pune, for stimulating discussions. The authors are thankful to the...
anonymous referee for making the detailed suggestions which resulted in the improvement of the paper. The authors are also thankful to IUCAA for providing hospitality and computational facilities where a major part of the work was carried out. This work is supported through research scheme vide grant number SP/S2/0-15/93 by DST, New Delhi.

REFERENCES

Apostolatos T. A., 1995, Phys. Rev. D, 52, 605
Apostolatos T. A., 1996, Phys. Rev. D, 54, 2421
Brady P. R., Creighton T., 2000, Phys. Rev. D, 61, 082001
Brady P. R., Creighton T., Cutler C., Schutz B. F., 1998, Phys. Rev. D, 57, 2101
Dhurandhar S. V., Sathyaprakash B. S., 1994, Phys. Rev. D, 49, 1707
Frasca S., 2000, Int. J Mod. Phys. D, 9, 369
Grishchuk L. P., Lipunov V. M., Postnov K. A., Prokhorov M. E., Sathyaprakash B. S., 2001, Phys. Usp. 44, 1

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.