Incoherent strategies for the network detection of periodic gravitational waves

P. Astone, S. Frasca, C. Palomba

INFN, Sezione di Roma and Università "La Sapienza", Roma, Italia.
E-mail: cristiano.palomba@roma1.infn.it

In the Virgo Collaboration, a hierarchical procedure for the blind search of continuous gravitational signals has been developed. A brief description of the method with some bibliographic references and of the preliminary results obtained on the data of C6 and C7 Comissioning Runs can be found elsewhere in these Proceedings.[1] In this paper we focus attention on an important part of the analysis, consisting in doing coincidences among the candidates found in two or more data sets, which allows to strongly reduce the false alarm probability. A characteristic of the continuous signals search is that data sets can indifferently belong to a single or more detectors.

Keywords: Gravitational waves; Continuous sources; Virgo detector.

1. Need for coincidences

In the hierarchical procedure developed in Virgo for the search of continuous gravitational signals we select candidates in a given data set putting a threshold on the critical ratio (CR) of the Hough sky histograms, defined as $CR = \frac{n-\mu}{\sigma}$ where n is the number count in a given cell of the histogram, μ is the mean number count and σ the standard deviation. The value of the threshold is chosen as a compromise between the need to minimize the sensitivity loss and to have a manageble number of candidates. By doing coincidences among candidates of two or more data sets we strongly reduce the false alarm probability and, because the coherent "follow-up" is done only on the surviving candidates, also its computational load is lowered. Making coincidences means to check if the parameters of a pair of candidates are within a given coincidence window. To perform coincidences we need at least two data sets. We can choose them in different ways and, as we will see, not all the choices are equivalent.

- Distinct data sets of one detector

This is the most standard choice. 'Spurious' candidates can appear for each data set, and then also in the coincidences, if they cover a short time interval.[1] This not only affects the false alarm probability but also the accuracy with which the parameters of a source, especially the position, can be determined. In addition, a loss of accuracy may also rise in frequency because to do coincidences among candidates we must compare their frequencies referred to the same reference time and a larger coincidence window must be used if the initial times of the two data sets are very different.

- Twofold 'mixed' data sets of one detector

We can take the two original data sets (call them a_0 and b_0) and suitably mix them creating two new sets (a_1 and b_1). A simple choice would consist, for instance, in taking a_1 as the first half of a_0 plus the first half of b_0 and b_1 as the second half of a_0 plus the second half of b_0. In this way the difference among the initial times of the two is shorter than before and, moreover, the time interval covered by each of them is larger, thus reducing the number of spurious coincidences.

- N-fold 'mixed' data sets

We can generalize the previous point by mixing more pieces of the original data sets. A particularly convenient choice is to produce new sets with approximately the same sensitivity. If we call a_i and b_i, with $i = 1, ..n;\ n \geq 2$, the pieces, one new set could be done, e.g., as $a_1 + a_3 + ... + b_1 + b_3 + ...$ and the other one as $a_2 + a_4 + ... + b_2 + b_4 + ...$ If the noise does not change a lot from one pieces to the next one the two sets will have about the same overall noise level. In this way, as will be shown in the following, the sensitivity of the analysis is larger.

In the case of data sets belonging to different detectors we can still work as in one of the previous cases. A difference is that we may have 'parallel' sets, i.e. covering the same time interval. This would minimize the coincidence window in frequency. Moreover, noise is uncorrelated in the original data sets.

Let us show that we benefit from making coincidences among data sets with the same sensitivity. Let us assume to have two data sets with corresponding linear signal to noise ratio SNR_1 and SNR_2, for a unitary amplitude signal in arbitrary units. By re-organizing them in two new data sets with equal sensitivity, the resulting SNR for both is $SNR = \sqrt[4]{\frac{SNR_1^4 + SNR_2^4}{2}}$ assuming the incoherent step of the hieararchical procedure (the Hough transform) is done adaptively.[2] The number count distribution in the Hough histograms is a binomial which, for average value $\mu >> 1$, can be approximated by a gaussian with the same parameters. Within this assumption, the critical ratios for the original data sets are

$$CR_1 = G(0;1) + SNR_1^2 \cdot h_{gw}^2; \quad CR_2 = G(0;1) + SNR_2^2 \cdot h_{gw}^2$$

where $G(0;1)$ describes a standard Gaussian variable and h_{gw} is the amplitude of the gravitational signal. In the case of two data sets with the same sensitivity we have

$$CR_1 = G(0;1) + SNR^2 \cdot h_{gw}^2; \quad CR_2 = G(0;1) + SNR^2 \cdot h_{gw}^2$$

The CR for a coincidence is $CR_{coin} = min(CR_1, CR_2)$ where CR_1 and CR_2 refers to the two coincident candidates; then, given a threshold CR_{thr}, we can take the probability $P(CR_{coin} > CR_{thr})$ as a measure of 'effectiveness' which allows us to compare the two cases, see Fig.(1). For equal sensitivity data sets the probability is larger, that is we could choose a lower threshold for candidate selection with a lower sensitivity loss at fixed false alarm probability. Or, viceversa, taking fixed the threshold we have a lower false alarm probability.

Fig. 1. 'Effectiveness' of the coincidence method for the original data sets ('standard' case, with $SNR_2 = \frac{SNR_1}{2}$) and two 'equal sensitivity' sets obtained from them.

Let us generalize the previous result by showing that, at least in principle, starting from a given data chunk, the bigger is the number N of 'equal sensitivity' subsets extracted from it and the larger is the gain in false alarm probability at fixed detection probability. In each subset we have a sensitivity loss $\sqrt[4]{N}$ because the sensitivity of the incoherent step scales as $\sqrt[4]{\frac{T_{obs}}{T_{FFT}}}$. Given a threshold for candidate selection, making coincidences among the candidates of the N subsets we have a reduction of the false alarm probability $P_{fa} = \left(P^s_{fa}\right)^N$, where P^s_{fa} refers to the original data set. The detection probability decreases as $\sqrt{N}$. Then, we can reduce the threshold for candidate selection by the same quantity in order to compare the false alarm probabilities for the same detection probability. This is shown in Fig.(2), where we clearly see that the false alarm probability decreases with increasing N, independently of the threshold.

(a) (b)

Fig. 2. (a) False alarm probability, at equal detection probability, as a function of the threshold for candidate selection. The most upper curve is for the original data set $N = 1$); the lower curves are for the coincidences among $N =$ 2,3 or 4 subsets with equal sensitivity extracted from the original one. (b) Ratio $\frac{N_i}{N_1}$, $i = 2, 3, 4$, of the false alarm probabilities, at fixed detection probability.

References

1. F. Acernese *et al.*, *First coincidence search among gravitational wave periodic source candidates using Virgo data*, these Proceedings.
2. C. Palomba, P. Astone, S. Frasca, *Classical Quantum Gravity* **22**, S1255 (2005).