Journal of Physics: Conference Series **716** (2016) 012003 doi:10.1088/1742-6596/716/1/012003

Distinguishing transient signals and instrumental disturbances in semi-coherent searches for continuous gravitational waves with line-robust statistics

David Keitel

Albert-Einstein-Institut, Callinstr. 38, 30167 Hannover, Germany

E-mail: `david.keitel@ligo.org`

Abstract. Non-axisymmetries in rotating neutron stars emit quasi-monochromatic gravitational waves. These long-duration 'continuous wave' signals are among the main search targets of ground-based interferometric detectors. However, standard detection methods are susceptible to false alarms from instrumental artefacts that resemble a continuous-wave signal. Past work [Keitel, Prix, Papa, Leaci and Siddiqi 2014, *Phys. Rev. D* **89** 064023] showed that a Bayesian approach, based on an explicit model of persistent single-detector disturbances, improves robustness against such artefacts. Since many strong outliers in semi-coherent searches of LIGO data are caused by transient disturbances that last only a few hours or days, I describe in a recent paper [Keitel D 2015, LIGO-P1500159] how to extend this approach to cover transient disturbances, and demonstrate increased sensitivity in realistic simulated data. Additionally, neutron stars could emit transient signals which, for a limited time, also follow the continuous-wave signal model. As a pragmatic alternative to specialized transient searches, I demonstrate how to make standard semi-coherent continuous-wave searches more sensitive to transient signals. Focusing on the time-scale of a single segment in the semi-coherent search, Bayesian model selection yields a simple detection statistic without a significant increase in computational cost. This proceedings contribution gives a brief overview of both works.

1. Introduction

Rotating neutron stars (NSs) with non-axisymmetric deformations emit long-term stable quasi-monochromatic gravitational radiation, called *continuous waves* (CWs). [1] These are searched for with data from terrestrial interferometers, such as LIGO and Virgo. A standard detection method is the matched-filter $\mathcal{F}$*-statistic* [2, 3]. Large-parameter space blind searches are most efficient with *semi-coherent* methods [4, 5, 6], splitting the data into shorter segments.

CW detection methods, such as the semi-coherent $\mathcal{F}$-statistic, are vulnerable to deviations from the standard Gaussian noise model that more closely resemble the quasi-monochromatic signal model. The data typically contains many narrow spectral artefacts, called *lines*, of instrumental or environmental origin. [7, 8, 9] These produce strong outliers in the $\mathcal{F}$-statistic, thus requiring some form of vetoing or follow-up procedure. [10, 11, 12, 13]

In previous work [14, 15], it was shown how a Bayesian model-comparison approach [16, 17] can be used to derive a *line-robust statistic* that achieves the same sensitivity as the $\mathcal{F}$-statistic in quiet, almost-Gaussian noise, but is significantly less affected by lines.

A recent extension of this approach [18] includes models for *transient* artefacts and transient signals. Transient line-like disturbances with typical durations of a few hours have been found in LIGO data [11, 12, 19] and cause similar problems as persistent lines. In addition, NSs can also emit transient signals that follow the CW model over time-scales of hours to days (*tCWs*, see [17, 20]). So a general detection statistic should be robust to both persistent and transient lines, while retaining or improving sensitivity to persistent and transient signals.

In Sec. 2, I give a brief review of the line-robust statistics of [14] for the case of persistent CW signals and persistent line disturbances, Then I sketch the approach of [18] to addressing transient disturbances and signals in Sec. 3, without any mathematical details. In both cases, I present results of injection studies with either archival LIGO S5 data or simulated data.

2. Line-robust statistics for persistent signals and disturbances

In the Bayesian hypothesis testing framework, the multi-detector $\mathcal{F}$-statistic [2, 3] is derived from the posterior odds between a CW signal hypothesis $\widehat{\mathcal{H}}_{\mathrm{S}} : \boldsymbol{x}(t) = \boldsymbol{n}(t) + \boldsymbol{h}(t; \mathcal{A}, \lambda)$ and a pure Gaussian noise hypothesis $\widehat{\mathcal{H}}_{\mathrm{G}} : \boldsymbol{x}(t) = \boldsymbol{n}(t)$. Here, $\boldsymbol{n}(t)$ is a multi-detector Gaussian time series and $\boldsymbol{h}(t; \mathcal{A}, \lambda)$ is a CW waveform with amplitude parameters $\mathcal{A}$ and phase-evolution parameters λ. As first found in [16] and extended to the semi-coherent case in [17], the result after marginalisation over $\mathcal{A}$ is $\widehat{O}_{\mathrm{S/G}}(\boldsymbol{x}) \equiv P\left(\widehat{\mathcal{H}}_{\mathrm{S}} \middle| \boldsymbol{x}, \mathcal{I}\right) \Big/ P\left(\widehat{\mathcal{H}}_{\mathrm{G}} \middle| \boldsymbol{x}, \mathcal{I}\right) \propto \mathrm{e}^{\widehat{\mathcal{F}}(\boldsymbol{x})}$.

The simple line model of [14] describes a single-detector line by $\widehat{\mathcal{H}}_{\mathrm{S}}$ restricted to that detector (indexed by X), $\widehat{\mathcal{H}}_{\mathrm{L}}^X : x^X(t) = n^X(t) + h^X(t; \mathcal{A}^X)$. A hypothesis test of 'CW signal vs. either Gaussian noise or a line' yields posterior odds

$$\widehat{O}_{\mathrm{S/GL}}(\boldsymbol{x}) = \frac{P\left(\widehat{\mathcal{H}}_{\mathrm{S}} \middle| \boldsymbol{x}, \mathcal{I}\right)}{P\left(\widehat{\mathcal{H}}_{\mathrm{G}} \middle| \boldsymbol{x}, \mathcal{I}\right) + P\left(\widehat{\mathcal{H}}_{\mathrm{L}} \middle| \boldsymbol{x}, \mathcal{I}\right)} \,. \tag{1}$$

Introducing the simplest example case of 2 detectors $X = 1, 2$ and 2 segments $\ell = 1, 2$, as well as a pictorial presentation of $\square$ for pure Gaussian noise in the data sub-set (X, ℓ) and $\blacksquare$ for a signal or disturbance, I can also represent the standard $\mathcal{F}$-statistic and (1) as

$$\mathrm{e}^{\widehat{\mathcal{F}}} \propto \frac{P\left(\begin{smallmatrix}\blacksquare\blacksquare\\ \blacksquare\blacksquare\end{smallmatrix} \,\middle|\, \boldsymbol{x}\right)}{P\left(\begin{smallmatrix}\square\square\\ \square\square\end{smallmatrix} \,\middle|\, \boldsymbol{x}\right)} \qquad \text{and} \qquad O_{\mathrm{S/GL}}(\boldsymbol{x}) = \frac{P\left(\begin{smallmatrix}\blacksquare\blacksquare\\ \blacksquare\blacksquare\end{smallmatrix} \,\middle|\, \boldsymbol{x}\right)}{P\left(\begin{smallmatrix}\square\square\\ \square\square\end{smallmatrix} \,\middle|\, \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\blacksquare\square\\ \blacksquare\square\end{smallmatrix} \,\middle|\, \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\square\blacksquare\\ \square\blacksquare\end{smallmatrix} \,\middle|\, \boldsymbol{x}\right)} \,. \tag{2}$$

See [14] for the derivation, tuning methods for the free parameters and detailed tests. As a detection statistic, one can always replace the posterior odds by a corresponding Bayes factor that factors out the prior odds, e.g. $B_{\mathrm{S/GL}}(\boldsymbol{x}, \mathcal{I}) \equiv O_{\mathrm{S/GL}}(\boldsymbol{x}, \mathcal{I}) / o_{\mathrm{S/GL}}(\mathcal{I})$.

Fig. 1, reproduces some of the main results from an injection study in [14], searching for simulated persistent CW signals in archival LIGO S5 data. These demonstrate that a properly tuned $B_{\mathrm{S/GL}}$ behaves just like $\widehat{\mathcal{F}}$ in quiet data (left panel), but improves over both $\widehat{\mathcal{F}}$ and the established *$\mathcal{F}$-stat consistency veto* [10, 11, 13] in the presence of lines (right panel).

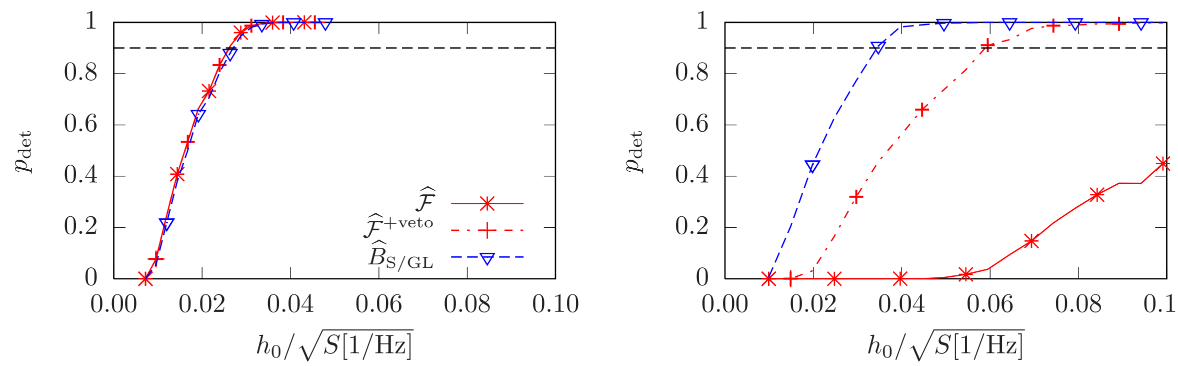

Figure 1. Detection efficiency p_{det} for persistent CW signals in LIGO S5 data, as a function of PSD-scaled signal amplitude $h_0/\sqrt{S}$, for the following semi-coherent statistics: $\widehat{\mathcal{F}}$ [2, 3], $\widehat{\mathcal{F}}^{+\mathrm{veto}}$ with consistency veto [10, 11], $\widehat{B}_{\mathrm{S/GL}}$ from Eq. (1) [14]. Left panel: a quiet 50 mHz band starting at 54.20 Hz, right panel: a disturbed 50 mHz band starting at 69.70 Hz. The dashed horizontal line marks $p_{\mathrm{det}} = 90\%$. Statistical uncertainties are smaller than the plot markers; some markers are omitted to make curves more discernible. Figures reproduced with data from [14].

3. Addressing transient disturbances and signals

To address transient effects, [18] adds a new 'transient-line hypothesis' for a quasi-harmonic disturbance in a single segment ℓ and single detector X: $\widetilde{\mathcal{H}}_{\mathrm{tL}}^{X\ell} : x^{X\ell}(t) = n^{X\ell}(t) + h^{X\ell}(t; \mathcal{A}^{X\ell})$, and a multi-detector-coherent transient-signal hypothesis $\widetilde{\mathcal{H}}_{\mathrm{tS}}^{\ell} : \boldsymbol{x}^{\ell}(t) = \boldsymbol{n}^{\ell}(t) + \boldsymbol{h}^{\ell}(t; \mathcal{A}^{\ell})$.

Testing persistent CWs against all three types of noise (Gaussian, persistent line, transient line) then leads to posterior odds

$$\widehat{O}_{\mathrm{S/GLtL}}(\boldsymbol{x}) = \frac{P\left(\widehat{\mathcal{H}}_{\mathrm{S}} \middle| \boldsymbol{x}, \mathcal{I}\right)}{P\left(\widehat{\mathcal{H}}_{\mathrm{G}} \middle| \boldsymbol{x}, \mathcal{I}\right) + P\left(\widehat{\mathcal{H}}_{\mathrm{L}} \middle| \boldsymbol{x}, \mathcal{I}\right) + P\left(\widehat{\mathcal{H}}_{\mathrm{tL}} \middle| \boldsymbol{x}, \mathcal{I}\right)} \,. \tag{3}$$

or, pictorially in the simplest example case,

$$\widehat{O}_{\mathrm{S/GLtL}}(\boldsymbol{x}) = \frac{P\left(\begin{smallmatrix}\blacksquare&\blacksquare\\\blacksquare&\blacksquare\end{smallmatrix} \mid \boldsymbol{x}\right)}{P\left(\begin{smallmatrix}\square&\square\\\square&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\blacksquare&\square\\\blacksquare&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\square&\blacksquare\\\square&\blacksquare\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\blacksquare&\square\\\square&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\square&\blacksquare\\\square&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + \ldots} \,. \tag{4}$$

An even more general test for the possibility of either type of signal (persistent CW or transient tCW) against the same combined noise model yields posterior odds

$$\widehat{O}_{\mathrm{StS/GLtL}}(\boldsymbol{x}) = \frac{P\left(\widehat{\mathcal{H}}_{\mathrm{S}} \middle| \boldsymbol{x}, \mathcal{I}\right) + P\left(\widehat{\mathcal{H}}_{\mathrm{tS}} \middle| \boldsymbol{x}, \mathcal{I}\right)}{P\left(\widehat{\mathcal{H}}_{\mathrm{G}} \middle| \boldsymbol{x}, \mathcal{I}\right) + P\left(\widehat{\mathcal{H}}_{\mathrm{L}} \middle| \boldsymbol{x}, \mathcal{I}\right) + P\left(\widehat{\mathcal{H}}_{\mathrm{tL}} \middle| \boldsymbol{x}, \mathcal{I}\right)} \tag{5}$$

$$= \frac{P\left(\begin{smallmatrix}\blacksquare&\blacksquare\\\blacksquare&\blacksquare\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\blacksquare&\blacksquare\\\square&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\square&\square\\\blacksquare&\blacksquare\end{smallmatrix} \mid \boldsymbol{x}\right)}{P\left(\begin{smallmatrix}\square&\square\\\square&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\blacksquare&\square\\\blacksquare&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\square&\blacksquare\\\square&\blacksquare\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\blacksquare&\square\\\square&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\square&\blacksquare\\\square&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\square&\square\\\blacksquare&\square\end{smallmatrix} \mid \boldsymbol{x}\right) + P\left(\begin{smallmatrix}\square&\square\\\square&\blacksquare\end{smallmatrix} \mid \boldsymbol{x}\right)} \,.$$

The corresponding Bayes factors $\widehat{B}_{\mathrm{S/GLtL}}$ and $\widehat{B}_{\mathrm{StS/GLtL}}$ are again obtained by factoring out the prior odds. For full expressions of these detection statistics, a detailed derivation and a description of the injection studies, the reader is referred to [18].

Fig. 2 shows results of an injection study similar to those in [14] and Fig. 1, using fully simulated data: Gaussian noise, a single-segment disturbance in one detector, and either persistent CW signals (left panel) or tCW signals that last for exactly one segment (right panel).

These tests demonstrate that, with appropriate tuning, the detection statistic $\widehat{B}_{\mathrm{StS/GLtL}}$ is indeed both *robust against transient or persistent disturbances* and *sensitive to transient or*

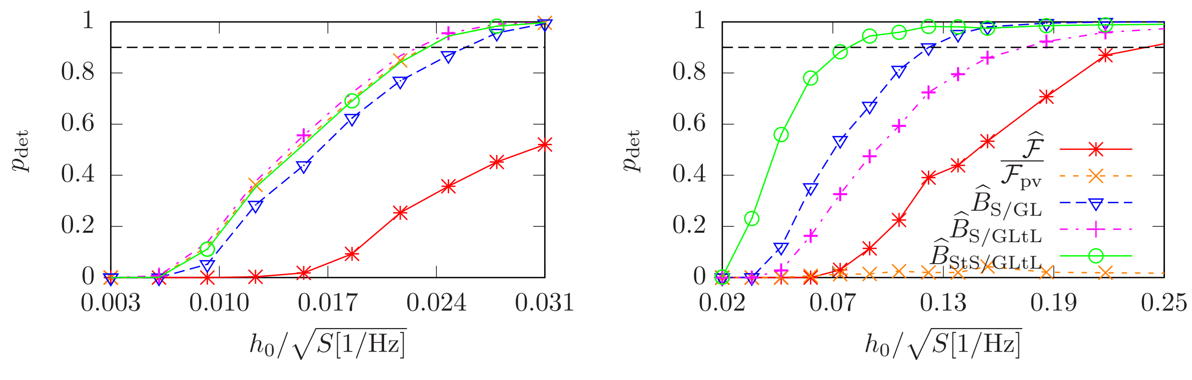

Figure 2. Detection efficiency p_{det} for different signal types in simulated data, for the following statistics: $\widehat{\mathcal{F}}$ [2, 3], $\overline{\mathcal{F}_{\mathrm{pv}}}$ with permanence veto [12], $\widehat{B}_{\mathrm{S/GL}}$ from Eq. (1) [14], $\widehat{B}_{\mathrm{S/GLtL}}$ from Eq. (3) [18], $\widehat{B}_{\mathrm{StS/GLtL}}$ from Eq. (5) [18]. In both panels, the underlying noise is Gaussian with a transient single-segment disturbance. Left panel: persistent CW signals, right panel: transient tCw signals with duration $T_{\mathrm{inj}} = T_{\mathrm{seg}} = 60\mathrm{h}$ in random segments. Plot properties as in Fig. 1.

persistent CW-like signals. In the first aspect, it reproduces the performance of a previously-introduced *permanence veto* [12], while in the latter it avoids the near-total dismissal of tCWs by that veto and improves over existing semi-coherent detection statistics.

As long as transient disturbances or signals are short enough to have a large part of their effect within a single segment, $\widehat{B}_{\mathrm{StS/GLtL}}$ is well approximated by considering only the maximum term in all sums over segments. This means that the search application only needs to store the *loudest* single-segment $\mathcal{F}^{\ell}$, $\mathcal{F}^{X\ell}$. Hence, it is easy to modify existing semi-coherent CW searches, and the extra computational cost and memory requirements will be low. This approach is thus promising for a cheap transient search as an 'add-on' to existing search pipelines, such as Einstein@Home [21]; an option that should be all the more valuable since no dedicated transient-CW search has been performed on real LIGO data so far.

As an end note to this proceeding, let me point out another approach to the detection of transient GWs on similar time-scales, based not on an explicit CW-like model, but on the 'radiometer' search [22]. It should be interesting to compare the sensitivity of both approaches, and of the transient-optimized matched-filter statistic of [17], on realistic data and with various signal characteristics.

Acknowledgments
I thank Reinhard Prix for valuable advice throughout this project; B. Behnke, M.A. Papa, O. Piccini, I. Di Palma and H.-B. Eggenstein for their data-analysis work on LIGO S5 and S6 data that inspired this study; the LIGO Scientific Collaboration for providing those data sets; G. Woan, I.S. Heng and E. Thrane for inspiring discussions about transient signals, and S. Walsh for advice on injection procedures. The injection studies were carried out on the ATLAS cluster at AEI Hannover. This paper has been assigned LIGO document number LIGO-P1500160.

References
[1] Prix R (for the LSC) 2009 Gravitational waves from spinning neutron stars *Neutron Stars and Pulsars* (*Astrophysics and Space Science Library* vol 357) ed Becker W (Springer Berlin Heidelberg) p 651 ISBN 978-3-540-76964-4 (LIGO-P060039-v3) URL `https://dcc.ligo.org/LIGO-P060039/public`
[2] Jaranowski P, Królak A and Schutz B F 1998 *Phys. Rev. D* **58** 063001 (*Preprint* `gr-qc/9804014`)
[3] Cutler C and Schutz B F 2005 *Phys. Rev. D* **72** 063006 (*Preprint* `gr-qc/0504011`)
[4] Brady P R and Creighton T 2000 *Phys. Rev. D* **61** 082001 (*Preprint* `gr-qc/9812014`)
[5] Cutler C, Gholami I and Krishnan B 2005 *Phys. Rev. D* **72** 042004 (*Preprint* `gr-qc/0505082`)
[6] Prix R and Shaltev M 2012 *Phys. Rev. D* **85** 084010 (*Preprint* `1201.4321`)
[7] Aasi J *et al.* (LVC) 2012 *Class. Quant. Grav.* **29** 155002 (*Preprint* `1203.5613`)
[8] Accadia T *et al.* 2012 *J. Phys. Conf. Ser.* **363** 012037
[9] Aasi J *et al.* (LVC) 2015 *Class. Quant. Grav.* **32** 115012 (*Preprint* `1410.7764`)
[10] Aasi J *et al.* (LVC) 2013 *Phys. Rev. D* **87** 042001 (*Preprint* `1207.7176`)
[11] Aasi J *et al.* (LVC) 2013 *Phys. Rev. D* **88** 102002 (*Preprint* `1309.6221`)
[12] Behnke B, Papa M A and Prix R 2015 *Phys. Rev. D* **91** 064007 (*Preprint* `1410.5997`)
[13] Leaci P 2015 *Phys. Scripta* **90** 125001
[14] Keitel D, Prix R, Papa M A, Leaci P and Siddiqi M 2014 *Phys. Rev. D* **89** 064023 (*Preprint* `1311.5738`)
[15] Keitel D and Prix R 2015 *Class. Quant. Grav.* **32** 035004 (*Preprint* `1409.2696`)
[16] Prix R and Krishnan B 2009 *Class. Quant. Grav.* **26** 204013 (*Preprint* `0907.2569`)
[17] Prix R, Giampanis S and Messenger C 2011 *Phys. Rev. D* **84** 023007 (*Preprint* `1104.1704`)
[18] Keitel D 2015 *ArXiv e-prints* Tech. rep. LIGO-P1500159 (*Preprint* `1509.02398`)
[19] Piccinni O 2014 *Mitigation of transient disturbances in wide parameter space searches for continuous gravitational wave signals* Master's thesis Università di Roma La Sapienza
[20] Santiago Prieto R I 2014 *Transient Gravitational Waves at r-mode Frequencies from Neutron Stars* Ph.D. thesis University of Glasgow
[21] Allen B *et al.* Einstein@Home distributed computing project URL `http://einsteinathome.org/`
[22] Thrane E, Mandic V and Christensen N 2015 *Phys. Rev. D* **91** 104021 (*Preprint* `1501.06648`)