PHYSICAL REVIEW D **86,** 063012 (2012)

An improved, "phase-relaxed" $\mathcal{F}$-statistic for gravitational-wave data analysis

Curt Cutler

Jet Propulsion Lab, 4800 Oak Grove Drive, Pasadena, California 91109, USA
Theoretical Astrophysics, California Institute of Technology, Pasadena, California 91125, USA
(Received 25 April 2011; published 25 September 2012)

Rapidly rotating, slightly nonaxisymmetric neutron stars emit nearly periodic gravitational waves (GWs), quite possibly at levels detectable by ground-based GW interferometers. We refer to these sources as "GW pulsars." For any given sky position and frequency evolution, the $\mathcal{F}$-statistic is the maximum likelihood statistic for the detection of GW pulsars. However, in "all-sky" searches for previously unknown GW pulsars, it would be computationally intractable to calculate the (fully coherent) $\mathcal{F}$-statistic at every point of (a suitably fine) grid covering the parameter space: the number of grid points is many orders of magnitude too large for that. Therefore, in practice some nonoptimal detection statistic is used for all-sky searches. Here we introduce a "phase-relaxed" $\mathcal{F}$-statistic, which we denote $\mathcal{F}_{\rm pr}$, for incoherently combining the results of fully coherent searches over short time intervals. We estimate (very roughly) that for realistic searches, our $\mathcal{F}_{\rm pr}$ is $\sim$10–15% more sensitive than the "semicoherent" $\mathcal{F}$-statistic that is currently used. Moreover, as a by-product of computing $\mathcal{F}_{\rm pr}$, one obtains a rough determination of the time-evolving phase offset between one's template and the true signal imbedded in the detector noise. Almost all the ingredients that go into calculating $\mathcal{F}_{\rm pr}$ are *already* implemented in the LIGO Algorithm Library, so we expect that relatively little additional effort would be required to develop a search code that uses $\mathcal{F}_{\rm pr}$.

DOI: 10.1103/PhysRevD.86.063012 PACS numbers: 95.55.Ym, 04.80.Nn, 95.75.Pq, 97.60.Gb

I. INTRODUCTION

The $\mathcal{F}$-statistic is the maximum likelihood statistic for the detection of nearly monochromatic gravitational waves (GWs) from a neutron star with known (or assumed) sky location and frequency evolution [1]. The basic idea behind the $\mathcal{F}$-statistic is simply this: for any given sky location and frequency evolution, the set of possible GW signals forms a four-dimensional (real) vector space [1,2]. The four basis vectors are the two quadratures (sin and cos) of each of the two polarization bases, $+$ and $\times$. Because the set is a vector space (not just a 4D manifold), it is computationally trivial to maximize the likelihood function over this set; $\mathcal{F}$ is the maximized log-likelihood. For Gaussian noise, the probability distribution function (pdf) of $\mathcal{F}$ is also particularly simple: a (perhaps noncentral) χ^2 distribution with 4 degrees of freedom. In the original paper by Jaranowski, Krolak, and Schutz [1] (hereinafter referred to as JKS), the $\mathcal{F}$-statistic was derived only for the case of a single GW detector and a single GW pulsar. (We trust that the reader understands that GW "pulsar" is a slight misnomer, since the GW emission is sinusoidal, not pulsed.) Cutler and Schutz [3] showed how the $\mathcal{F}$-statistic can be generalized in a straightforward manner to the cases of (1) a network of detectors noise curves, and (2) an entire collection of known sources. A proper Bayesian version of the $\mathcal{F}$-statistic is derived in Ref. [4].

In practice, searches for nearly monochromatic GWs are separated into a few different types, depending on how much is known about the source. The different types of searches can have vastly different computational requirements. While the search for a GW counterpart to a known radio pulsar is trivial in terms of computational burden, "all-sky" searches for GW pulsars with no known counterpart (and hence unknown frequency and frequency derivatives) are currently limited by the available computational power. That is, we could dig deeper into the existing data sets if we possessed either larger computational resources or more efficient algorithms. In this paper we demonstrate a way of significantly improving on the existing algorithms.

Currently, the most sensitive all-sky searches are based on the following idea [5]: For a GW pulsar with unknown frequency evolution (i.e., unknown f, $\dot{f}$, etc.), computational power required for a fully coherent search grows as a high-power of the total observation time, T. Therefore, in practice one divides T into some number N (typically of order 10^2) short intervals of duration $\Delta T = T/N$, performs a coherent search on each short interval, and then "adds up" the power from all the subintervals. More specifically, the current method is to calculate a semicoherent detection statistic $\mathcal{F}_{\rm sc}$, defined as the sum of the $\mathcal{F}$ values from each of the short intervals:

$$\mathcal{F}_{\rm sc} \equiv \sum_{i=1}^{N} \mathcal{F}_i. \tag{1.1}$$

Because calculating each $\mathcal{F}_i$ involves a maximization over four free parameters, the pdf for $2\mathcal{F}_{\rm sc}$ is a χ^2 distribution with $4N$ degrees of freedom (d.o.f.) But this is far more parameters than are actually needed to describe the physical system. Consider the very first interval. The imbedded GW signal is described by four parameters: (h_0, ι, ψ, Φ).

But the triplet (h_0, ι, ψ) are the same for all N intervals. All that changes from interval to interval is the overall phase Φ. Assuming maximum ignorance of the signal's phase evolution, one therefore needs $N - 1$ additional phases to fully describe the signal. So, the GW signal is fully described by only $N + 3$ parameters. We define $\mathcal{F}_{\rm pr}$ to be the maximized log-likelihood on this $N + 3$-dimensional space. Our main aims in this paper are (i) to demonstrate an efficient algorithm for calculating $\mathcal{F}_{\rm pr}$ and (ii) to illustrate its superiority as a detection statistic (superior in the sense of improved ROC curves, where ROC stands for receiver operating characteristics).

We note that our work is rather similar in spirit to a recent paper by Dergachev [6], though we believe that we have advanced the idea considerably further; e.g., while Dergachev [6] essentially restricts to the case of a single polarization, we explicitly treat the realistic case of two polarizations, which in this context represents a significant complication. We note, too, a recent paper by Pletsch [7], which addresses the following, rather different weakness in the detection statistic Eq. (1.1). To understand the problem, let $(t_0, t_1, \cdots, t_N)$ be the boundary points of the N short intervals, and consider any one such boundary time, say t_{50}. In Eq. (1.1), data sampled only slightly earlier than t_{50} gets combined *incoherently* with data sampled only slightly later than t_{50}, which clearly exacts a price in sensitivity. Pletsch overcomes this problem (which stems from the rather arbitrary choice of boundary points) using his "sliding window" technique [7]. We suspect that relatively simple refinements of the detection statistic that we develop in this paper could also capitalize on Pletsch's basic insight, but we leave such refinements to future work.

The plan of this paper is as follows. In Sec. II, we briefly review the fully coherent $\mathcal{F}$-statistic, partly to establish notation. We generally try to align our notation with that of JKS, to ease comparison with their work. We also review the (currently used) semicoherent $\mathcal{F}$-statistic and its properties. As further motivation for our work, in Sec. IV we consider a "warm-up" problem that we can easily treat analytically and that qualitatively has much in common with our actual problem. In Sec. V, in order to illustrate the use of $\mathcal{F}_{\rm pr}$, we present numerical results for one example search. For the same search, we also investigate the relative power/sensitivity of $\mathcal{F}_{\rm pr}$ versus $\mathcal{F}_{\rm sc}$. Our conclusions are summarized in Sec. VI.

This paper represents our "first-cut" analysis of the $\mathcal{F}_{\rm pr}$ statistic. There remains significant follow-up work to better elucidate the properties of $\mathcal{F}_{\rm pr}$ and to implement it in realistic, hierarchical searches. This future work is also summarized in Sec. VI.

II. REVIEW OF SIGNAL PROCESSING FOR GW PULSARS

In this section we review the rudiments of signal processing that we will require, partly to fix notation. We also review both the coherent and semicoherent versions of the $\mathcal{F}$-statistic. For simplicity of exposition, in this paper we will restrict to the case of a single detector, and assume that the detector noise is stationary. The extension to the more realistic case of multiple detectors with slowly changing noise spectra is completely straightforward.

A. Mathematics of signal processing

We begin by reviewing the basic mathematics of signal processing. For more details, we refer the reader to Thorne, Finn and Chernoff, and/or Cutler and Flanagan [2,8,9].

Assuming that the noise is stationary and Gaussian, the noise spectral density $S_h(f)$ determines a natural inner product $(\ldots|\ldots)$ on the vector space of all detector outputs $x(t)$:

$$(\mathbf{x}|\mathbf{y}) \equiv 2\int_{-\infty}^{\infty} df \frac{\tilde{x}^*(f)\tilde{y}(f)}{S_h(f)}, \tag{2.1}$$

where $\tilde{x}(f)$ and $\tilde{y}(f)$ denote the Fourier transforms of $x(t)$ and $y(t)$, and $S_h(f)$ is the *single-sided* spectral density of the noise. In terms of this inner product, the pdf for the noise $\mathbf{n}(t)$ takes the form

$$\mathrm{pdf}[\mathbf{n}] = \mathcal{N} e^{-(\mathbf{n}|\mathbf{n})/2}, \tag{2.2}$$

where $\mathcal{N}$ is a normalization constant. Using $\langle\cdots\rangle$ to denote "expectation value" (over many realizations of the noise), it follows from Eq. (2.2) that

$$\langle(\mathbf{x}|\mathbf{n})(\mathbf{y}|\mathbf{n})\rangle = (\mathbf{x}|\mathbf{y}). \tag{2.3}$$

In this paper, we will be concerned with waveforms $h(t)$ that are nearly monochromatic [here meaning that their frequencies $f(t)$ are slowly varying]. In this case their inner product is equally simple in the time domain. Taking the measurement time interval to be 0 to T, by Parseval's theorem we have

$$(\mathbf{h_1}|\mathbf{h}_2) = 2\int_0^T \frac{\tilde{h}_1(t)\tilde{h}_2(t)}{S_h(f(t))} dt. \tag{2.4}$$

B. The fully coherent $\mathcal{F}$-statistic

Next we briefly review the use of the coherent $\mathcal{F}$-statistic in GW pulsar searches. For more details we refer the reader to Cutler and Schutz [3]. Consider a nearly monochromatic GW signal from an individual source with known sky location and known frequency evolution $f(t)$. The GW signal is then characterized by four remaining unknowns: an overall amplitude A (equivalent to the combination $h_0 \sin\zeta \sin^2\theta$ in the notation of JKS), two angles ι and ψ that characterize the waves' polarization (equivalent to determining the direction of the neutron star's spin axis), and an overall phase Φ.

The GW signal $h(t)$ registered by the detector depends nonlinearly on ι, ψ, Φ, but, crucially, one can make a simple change of variables–to $(\lambda^1, \lambda^2, \lambda^3, \lambda^4)$–such that the dependence of $h(t)$ is linear in these new variables:

$$h(t) = \sum_{a=1}^{4} \lambda^a h_a(t), \tag{2.5}$$

where the four basic waveforms $h_a(t)$ are defined by

$$\begin{aligned} h_1(t) &= F_+(t)\cos\Phi(t), \qquad h_2(t) = F_\times(t)\cos\Phi(t),\\ h_3(t) &= F_+(t)\sin\Phi(t), \qquad h_4(t) = F_\times(t)\sin\Phi(t). \end{aligned} \tag{2.6}$$

Here $\Phi(t)$ is the waveform phase at the detector:

$$\Phi(t) \approx 2\pi \int^t f(t')dt', \tag{2.7}$$

where $f(t')$ is the measured GW frequency at the detector at time t'. The measured frequency includes the Doppler effect from the detector's motion relative to the source, as well as the Einstein and Shapiro delays associated with the Earth's orbit around the Sun. When the GW pulsar is in a binary, then $f(t')$ also includes the Roemer, Einstein, and Shapiro delays associated with that binary orbit. (We emphasize that the known-pulsar searches described here do *not* require that the GW pulsar be isolated, but just that there exists an accurate timing model for the emitted waves.) The $F_+(t)$ and $F_\times(t)$ terms in Eq. (2.5) are the beam-pattern functions that describe the detector's response to the $+$ and $\times$ polarizations, respectively. We note that the exact form of $F_+(t)$ and $F_\times(t)$ depends on one's convention for decomposing the waveform into "plus" and "cross" polarizations; a one-parameter family of choices is possible, corresponding to the freedom to rotate the axes around the line of sight. JKS follow the conventions of Bonazzola and Gourgoulhon [10].

Next we define the 4×4 matrix Γ_{ab} by

$$\Gamma_{ab} \equiv \left(\frac{\partial \mathbf{h}}{\partial \lambda^a} \Big| \frac{\partial \mathbf{h}}{\partial \lambda^b}\right) = (\mathbf{h_a}|\mathbf{h_b}). \tag{2.8}$$

Because both the observation time and 1 day [the time scale on which the $F_{+,\times}(t)$ vary] are vastly larger than the period of the sought-for GWs (typically 10^{-2}–10^{-3} s), we can replace $\cos^2\Phi(t)$, $\sin^2\Phi(t)$, and $\cos\Phi(t)\sin\Phi(t)$ by their time averages: $\cos^2\Phi(t)$, $\sin^2\Phi(t) \rightarrow \frac{1}{2}$, while $\cos\Phi(t)\sin\Phi(t) \rightarrow 0$. Then we have

$$\begin{aligned} \Gamma_{11} &\approx \int F_+(t)F_+(t)S_h^{-1}(f(t))dt\\ \Gamma_{12} &\approx \int F_+(t)F_\times(t)S_h^{-1}(f(t))dt\\ \Gamma_{22} &\approx \int F_\times(t)F_\times(t)S_h^{-1}(f(t))dt; \end{aligned} \tag{2.9}$$

additionally, $\Gamma_{33} \approx \Gamma_{11}$, $\Gamma_{34} \approx \Gamma_{12}$, $\Gamma_{44} \approx \Gamma_{22}$, and $\Gamma_{13} \approx \Gamma_{14} \approx \Gamma_{23} \approx \Gamma_{24} \approx 0$.

The best-fit values of λ^a satisfy

$$\frac{\partial}{\partial \lambda^a}\left(\mathbf{x} - \sum_b \lambda^b \mathbf{h_b} | x - \sum_c \lambda^c \mathbf{h_c}\right) = 0, \tag{2.10}$$

implying that

$$\lambda^a = \sum_b (\Gamma^{-1})^{ab}(\mathbf{x}|\mathbf{h_b}). \tag{2.11}$$

Then $2\mathcal{F}$, which is defined to be twice the log of the maximized likelihood ratio, is just

$$\begin{aligned} 2\mathcal{F} &= (\mathbf{x}|\mathbf{x}) - \left(\mathbf{x} - \sum_b \lambda^b \mathbf{h_b} | \mathbf{x} - \sum_c \lambda^c \mathbf{h_c}\right)\\ &= \sum_{a,d} (\Gamma^{-1})^{ad}(\mathbf{x}|\mathbf{h_a})(\mathbf{x}|\mathbf{h_d}). \end{aligned} \tag{2.12}$$

Using $2\mathcal{F}$ as one's detection statistic satisfies the Neyman-Pearson criterion for an optimum test: it minimizes the false dismissal (FD) probability for any given false alarm (FA) probability.

Writing $\mathbf{x} = \mathbf{n} + \mathbf{h}$, and plugging into Eq. (2.12), we find

$$\langle 2\mathcal{F}\rangle = 4 + (\mathbf{h}|\mathbf{h}), \tag{2.13}$$

where we have used Eq. (2.3) and the fact that $\langle(\mathbf{h}|\mathbf{n})\rangle = 0$. More generally, it is easy to show that $y \equiv 2\mathcal{F}$ follows a χ^2 distribution with 4 d.o.f. and noncentrality parameter $\rho^2 \equiv (\mathbf{h}|\mathbf{h})$:

$$P(y) = \chi^2(y|4;\rho^2). \tag{2.14}$$

As pointed out by JKS, if we use the following complexified variables,

$$2F_a \equiv (\mathbf{x}|\mathbf{h_1} - i\mathbf{h_3}), \qquad 2F_b \equiv (\mathbf{x}|\mathbf{h_2} - i\mathbf{h_4}), \tag{2.15}$$

then the expression (2.12) for $2\mathcal{F}$ can be rewritten in a particularly simple form: Eq. (2.12) becomes

$$2\mathcal{F} = \frac{8}{D}[B|F_a|^2 + A|F_b|^2 - 2C\,\Re(F_a F_b^*)], \tag{2.16}$$

where

$$A \equiv (\mathbf{h_1}|\mathbf{h_1}), \quad B \equiv (\mathbf{h_2}|\mathbf{h_2}), \quad C \equiv (\mathbf{h_1}|\mathbf{h_2}), \tag{2.17}$$

and $D \equiv AB - C^2$. (Note that the A,B,C terms defined here are, in the single-detector case, larger than the A, B, C terms in JKS by a factor of the observation time T.)

C. The semicoherent $\mathcal{F}$-statistic

As mentioned above, the current method of incoherently combining the coherent results from successive intervals is just to sum the $\mathcal{F}$-statistics from all the intervals:

$$2\mathcal{F}_{\rm sc} \equiv \sum_{i=1}^{N} 2\mathcal{F}_i. \tag{2.18}$$

It also easy to show that $y \equiv 2\mathcal{F}_{\rm sc}$ follows a χ^2 distribution with $4N$ degrees of freedom:

$$P(y) = \chi^2(y|4N;\rho^2), \tag{2.19}$$

where the noncentrality parameter $\rho^2 = \sum_{i=1}^{N}\rho_i^2$.

In the cases of interest to us, $4N$ will generally be large, and then the χ^2 distribution with $4N$ d.o.f. can often be approximated as a Gaussian. Let $y \equiv 2\mathcal{F}$, and let $\rho_{\rm tot}^2 \equiv \sum_{i=1}^{N} \rho_i^2$. Then

$$P(y) = \chi^2(y|4N;\rho_{\rm tot}^2) \approx (16\pi N)^{-1/2} e^{-(y-<y>)^2/(16N)}, \tag{2.20}$$

where $\langle y\rangle = 4N + \rho_{\rm tot}^2$. For example, using this approximation [and the fact that for a Gaussian $P(y)$, events 2.326σ above the mean occur 1% of the time], we see that the threshold value $y_{\rm th}$ that yields a 1% FA probability is

$$y_{\rm th} \approx 4N + 2.326\sqrt{8N} \qquad (\text{large } N). \tag{2.21}$$

III. THE "PHASE-RELAXED" $\mathcal{F}$-STATISTIC

We are now ready to define $\mathcal{F}_{\rm pr}$. Basically, $\mathcal{F}_{\rm pr}$ coincides with the full matched-filtering SNR^2, under the assumption that the manifold of waveforms is $N+3$-dimensional (i.e., four parameters for the first segment, and $N-1$ for the relative phase offsets of the remaining segments). What makes $\mathcal{F}_{\rm pr}$ useful in practice is that we have also found a simple and efficient method for calculating it.

A. Motivation and definition

We begin by defining complex basis functions H_+ and $H_\times$ by

$$H_+ \equiv h_1 - ih_3, \qquad H_\times \equiv h_2 - ih_4. \tag{3.1}$$

This complex representation is especially convenient for our purposes because H_+ and $H_\times$ both transform very simply under an overall phase shift in $\Phi(t)$: under $\Phi(t) \to \Phi(t) + \delta$, H_+ and $H_\times$ transform as $H_+(t) \to e^{-i\delta}H_+(t)$ and $H_\times(t) \to e^{-i\delta}H_\times(t)$. [Note that the minus sign in the exponent in the term $e^{-i\delta}$ stems from the minus signs in the definitions of H_+ and $H_\times$ in Eq. (3.1)]. For these complexified signals, our usual inner product becomes a Hermitian one; for nearly monochromatic signals near frequency f, this Hermitian inner product is given simply by

$$(\mathbf{x}|\mathbf{y}) = \frac{2}{S_h(f)} \int_0^T x^*(t) y(t) dt. \tag{3.2}$$

Clearly $(\mathbf{x}|\mathbf{y}) = (\mathbf{y}|\mathbf{x})^*$.

Next we define $\Gamma_{\alpha\beta}$ by

$$\Gamma_{\alpha\beta} \equiv (H_\alpha|H_\beta) = \Gamma^*_{\beta\alpha}, \tag{3.3}$$

where α and β run over $+, \times$. It follows immediately that for GW data $x(t)$, (twice the) $\mathcal{F}$-statistic is given by

$$2\mathcal{F} = (\Gamma^{-1})^{\alpha\beta}(H_\alpha|\mathbf{x})(\mathbf{x}|H_\beta). \tag{3.4}$$

Now imagine breaking up the full integration time T into N intervals of duration ΔT_i, for $i = 1, 2, \cdots, N$. (We expect that in practice the ΔT_i will generally be of approximately the same length, but this is not required.) Next define $x^i(t)$ to be the restriction of $x(t)$ to the ith interval; i.e., $x^i(t) = x(t)$ for t in the ith interval, and $x^i(t) = 0$ for t outside the ith interval. Then clearly we have

$$2\mathcal{F} = (\Gamma^{-1})^{\alpha\beta}\Big(\sum_i (H_\alpha|\mathbf{x^i}\Big)\Big(\sum_j (\mathbf{x^j}|H_\beta)\Big). \tag{3.5}$$

To motivate our definition of $\mathcal{F}_{\rm pr}$, recall that if we had practically limitless computer power at our disposal, then the most sensitive search would be a coherent matched-filter search over a fine grid covering the entire GW-pulsar parameter space. However for "blind" GW pulsar searches (i.e., searches for GW pulsars whose sky location and/or time-changing frequency are unknown), maintaining phase coherence between the template signal and true imbedded signal, over time scales of months to years, would require an extremely fine grid on parameter space, and (one easily shows) many of orders of magnitude more computing power than is realistic [11].

The basic idea behind semicoherent searches is to employ a detection statistic that is less sensitive to phase decoherence across the whole observation time, which allows one to use a much coarser grid on parameter space. In effect, one sacrifices some sensitivity in the interest of computational practicality. For our phase-relaxed $\mathcal{F}$-statistic, the idea is that the search-template signal should remain approximately in phase with the true, imbedded signal in each interval ΔT_i–up to some constant phase offset" δ_i–but that the δ_i should be allowed to vary from interval to interval. That is, we replace

$$\begin{aligned}&(\Gamma^{-1})^{\alpha\beta}\Big(\sum_i (H_\alpha|\mathbf{x^i})\Big)\Big(\sum_j (\mathbf{x^j}|H_\beta)\Big)\\ &\quad\to (\Gamma^{-1})^{\alpha\beta}\Big(\sum_i (H_\alpha|\mathbf{x^i})e^{i\delta_i}\Big)\Big(\sum_j (\mathbf{x^j}|H_\beta)e^{-i\delta_j}\Big).\end{aligned} \tag{3.6}$$

Finally, we define (twice) $\mathcal{F}_{\rm pr}$ to be the right-hand side of (3.6), maximized over all phase-offsets δ_i:

$$\begin{aligned}2\mathcal{F}_{\rm pr} = \max_{\delta_1,\cdots,\delta_n}\Big\{&(\Gamma^{-1})^{\alpha\beta}\Big(\sum_i (H_\alpha|\mathbf{x^i})e^{i\delta_i}\Big)\\ &\times\Big(\sum_j (\mathbf{x^j}|H_\beta)e^{-i\delta_j}\Big)\Big\}.\end{aligned} \tag{3.7}$$

While there are N phase angles δ_i, only $N-1$ of them are actually independent; i.e., it is easy to check that $\mathcal{F}_{\rm pr}$ is invariant under $\delta_i \to \delta_i + c$, where c is any constant.

B. Maximizing over the phase offsets δ_i

The whole point of developing alternatives to the fully coherent $\mathcal{F}$-statistic is to save on computational cost, so for our phase-relaxed $\mathcal{F}$-statistic to be useful, we need a reasonably efficient way of maximizing over the δ_i. In this section we demonstrate one efficient method. We demonstrate only the simplest version of this method, which we regard as basically an "existence proof" that efficient methods do exist. It should be clear by the end of this section that there are many variations on our basic method by which one might attempt to improve its efficiency, but we defer such improvements to later work.

Our method is as follows. We can simplify the appearance of the equations by defining

$$K^{ji} \equiv (\Gamma^{-1})^{\alpha\beta}(\mathbf{x^j}|H_\beta)(H_\alpha|\mathbf{x^i}) \tag{3.8}$$

and defining v to be the following N-dimensional vector formed out of the phase offsets:

$$\mathbf{v} \equiv (e^{i\delta_1}, e^{i\delta_2}, \cdots e^{i\delta_n}). \tag{3.9}$$

Note that K^{ji} is Hermitian (i.e., $K^{ij} = K^{ji*}$), and that we can now rewrite Eq. (3.6) as

$$2\mathcal{F}_{\text{pr}} = \max_{\delta_1,\cdots,\delta_n} v_j^* K^{ji} v_i. \tag{3.10}$$

Of course, K^{ji} is completely determined by the two complex templates H_α and their inner products with the data, while our goal is to find the v_i that maximize $v_j^* K^{ji} v_i$, subject to the N constraints that $v_i v_i^* = 1\ \forall\ i$. (We emphasize that i is not summed over in these constraints.) Put another way: $\mathbf{v}$ lies on the unit N-torus (i.e., the unit circle cross itself N times.) Naturally, we employ the method of Lagrange multipliers to maximize $v_j^* K^{ji} v_i$ on this constraint surface. Since there are N constraints, we obtain N equations with N (real) Lagrange multipliers λ_j:

$$K^{ji} v_i = \lambda_j v_j \quad \forall\ j. \tag{3.11}$$

We emphasize that Eq. (3.11) is not an eigenvalue equation, since in general the N values λ_j will all be different.

Next we find it convenient to introduce a projection operator P operating on $\mathbb{C}^N$. Let $\mathbf{w} = (c_1 e^{i\delta_1}, c_2 e^{i\delta_2}, \cdots, c_N e^{i\delta_N})$, where the c_i are all real. Then P is defined by

$$P\mathbf{w} \equiv (e^{i\delta_1}, e^{i\delta_2}, \cdots, e^{i\delta_N}). \tag{3.12}$$

That is, the operator P takes any vector in $\mathbb{C}^N$ and projects it down onto the unit torus. (Note that P is not a linear operator, but it is true that $P^2 = P$.) Then Eq. (3.11) is clearly equivalent to the requirement that

$$PK\mathbf{v} = \mathbf{v}. \tag{3.13}$$

Hence the solution $\mathbf{v}$ is a fixed point of the operator PK. In fact, numerical experience shows that it is an attractive fixed point. That is, let $\mathbf{v_0}$ be some initial guess, and then operate on it repeatedly with PK. Define $(PK)^2 \equiv (PK)(PK)$, $(PK)^3 \equiv (PK)(PK)(PK)$, etc. Then for $\mathbf{v_0}$ sufficiently close to the true solution $\mathbf{v}$, we find that

$$(PK)^m \mathbf{v_0} \rightarrow \mathbf{v} \tag{3.14}$$

as m increases. In practice, we find that the convergence is quite rapid, and that the initial guess $\mathbf{v_0}$ need not be particularly close to the solution $\mathbf{v}$. In numerical experiments (in many thousands of cases, and covering a large range of N), we found that the following initial guess always led to converge of the iterated sequence. For each segment ΔT_i, it is trivial to calculate the fully coherent $\mathcal{F}_i$ and the corresponding best-fit parameters for that segment alone: $(A_i, \iota_i, \psi_i \Phi_i)$. Then we take as our initial guess

$$\mathbf{v_0} = (e^{i\Phi_1}, e^{i\Phi_2}, \cdots e^{i\Phi_n}); \tag{3.15}$$

i.e., the initial guess for the phase offset in each segment is the best-fit offset for that segment by itself.

IV. ANALYTIC RESULTS FOR A RELATED, WARM-UP PROBLEM

It is common sense that when one goes to solve some problem numerically, it is useful to have analytical results with which to compare it–ideally for a special case of the true problem, or, failing that, for some qualitatively similar problem. In this section we derive analytic results for the following case: Consider a vector space of waveforms that is completely described by 2 parameters per interval–so $2N$ parameters in all, where N is large–and consider two different searches: one search that maximizes the fit over those $2N$ parameters, and another, less efficient search, that begins with a $4N$-dimensional vector space (in which the true, $2N$-dim vector space lies), whose detection statistic is the maximized log-likelihood on the $4N$-dimensional space. That is, our two detection statistics are the $2N$- and $4N$-dimensional $\mathcal{F}$-statistics, which in this section we will denote $\mathcal{F}_{2N}$ and $\mathcal{F}_{4N}$.

For each search, there is a threshold value ρ_{th} such that the signal is detectable with $FA = 0.01$ and $FD = 0.5$. We can solve both problems at the same time, by considering

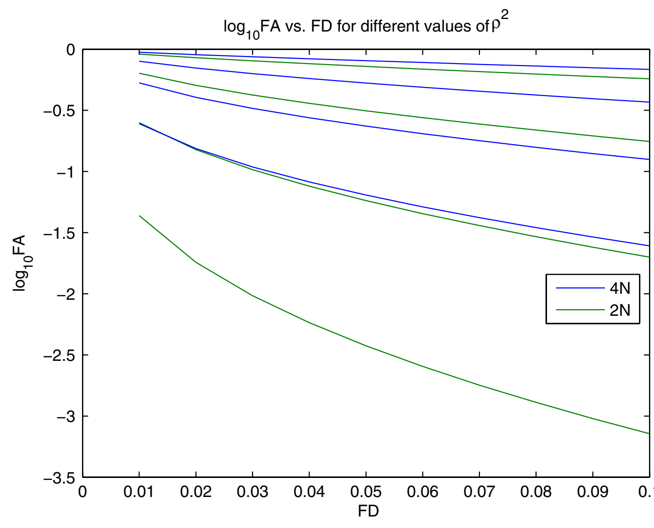

FIG. 1 (color online). Compares the FA probabilities for the two detection statistics, $\mathcal{F}_{4N}$ and $\mathcal{F}_{2N}$, as a function of FD probability, for $N = 100$ and for several values of the total squared signal strength, $\rho^2 \equiv \sum_{i=1}^N \rho_i^2$. The blue (dark) curves are for $\mathcal{F}_{4N}$ and the green (light) for $\mathcal{F}_{2N}$. From upper to lower, the squared signal strengths are $\rho^2 = 25, 50, 75, 100$.

the general M-dimensional search. Then the expectation value of $\mathcal{F}_M$ is $\langle \mathcal{F}_M \rangle = M + \rho^2$ and its standard deviation is $\sigma_M = (2M + 4\rho^2)^{1/2}$. For large M, the χ^2 function approaches a Gaussian, so we will approximate the pdf of $\mathcal{F}_M$ as a Gaussian with this mean and standard deviation. Then the threshold for detection with $FA = 0.01$ is

$$\rho_M^{\rm th} = \langle \mathcal{F}_M \rangle + \sqrt{2}\sigma_M {\rm erfc}^{-1}(2FA) \tag{4.1}$$

$$= M + 3.29 M^{1/2}, \tag{4.2}$$

where in Eq. (4.1) both $\mathcal{F}_M$ and σ_M are to be evaluated at $\rho = 0$. Therefore $\rho^{{\rm th}_M = \sqrt{3.29 M^{1/2}} = 1.814 M^{1/4}}$, and we have

$$\rho_{4N}^{\rm th} / \rho_{2N}^{\rm th} = 2^{1/4} = 1.189. \tag{4.3}$$

Therefore using the correct statistic allows one to see sources $\approx 19\%$ farther away.

For comparison with results in the next section, we also plot in Fig. 1 the FA vs FD curves for the two statistics, for $N = 100$ and a range of ρ values.

V. NUMERICAL RESULTS FOR ONE EXAMPLE SEARCH

To illustrate the utility of our $\mathcal{F}_{\rm pr}$ statistic, in this section we present results for a simple, one-parameter family of examples (where the varied parameter is the strength of the embedded GW signal), and we compare the effectiveness of $\mathcal{F}_{\rm pr}$ and $\mathcal{F}_{\rm sc}$.

We fix the number of intervals at $N = 100$, and evaluate search effectiveness for signals with a range of total $\rho^2 \equiv \sum_{i=1}^N \rho_i^2$. We will eventually consider a range of ρ, but for now imagine ρ as fixed. For simplicity, in this example we will consider a case where the ρ_i are the same for all i, so $\rho_i^2 = \rho^2/N$, and where the Γ^i_{ab} matrices are also the same for all i: $\Gamma^i_{++} = 3$, $\Gamma^i_{+\times} = 1 = \Gamma^i_{\times +}$ and $\Gamma^i_{\times\times} = 1$ for all i.

We decompose the measured signal x^i into waveform plus noise,

$$x^i = h^i + n^i. \tag{5.1}$$

For each i, filtering the data with H_+ and H_x produces two complex numbers: $c^i_+ \equiv (\mathbf{x^i}|H_+)$ and $c^i_\times \equiv (\mathbf{x^i}|H_\times)$. Clearly, the measured signals c^i_α can be decomposed as

$$c^i_\alpha = (\mathbf{h^i}|H_\alpha) + (\mathbf{n^i}|H_\alpha) \tag{5.2}$$

$$\equiv g^i_\alpha + m^i_\alpha. \tag{5.3}$$

We simulate the noise piece m^i_α by taking random draws of (pairs of) complex numbers from a Gaussian distribution with covariance matrix

$$\langle m^{i*}_\alpha m^i_\beta \rangle \equiv \langle (H_\alpha|\mathbf{n^i})(\mathbf{n^i}|H_\beta) \rangle = (H_\alpha|H_\beta) = \Gamma_{\alpha\beta}. \tag{5.4}$$

Again for simplicity, we will consider a case where the g_i^α are the same for each i, modulo a random, complex phase factor. Our particular (and rather arbitrary) choice is

$$[g_i^+, g_i^\times] = (2N)^{-1/2}\rho[2 + \sqrt{6}, \sqrt{6}]e^{i\varphi^i}, \tag{5.5}$$

where the φ^i are random phases drawn uniformly from $(0, 2\pi)$. One easily checks that $\sum_{i=1}^N (h^i|h^i) = \rho^2$. The inclusion of the $e^{i\varphi^i}$ terms reflects our goal of modeling a case where frequency evolutions of the template and the true signal are so mismatched that their relative phases jump significantly and randomly from one interval to the next. Choosing the φ^i randomly corresponds to the worst-case scenario," where the true-versus-template phase offsets show no pattern. In practice, we expect that the situation will often be much more favorable for searches: i.e., the phase offsets might very often be well fit by some low-order polynomial in time. In a later paper we plan to investigate the extent to which the time evolution of the offsets can be fit by a few parameters, and how that information can be exploited to speed up other parts of the search.

Given one simulated data set c^i_α (200 complex numbers), we compute $2\mathcal{F}_{\rm pr}$. We repeat for 10000 data sets to determine the distribution of $2\mathcal{F}_{\rm pr}$, and calculate its mean $\langle 2\mathcal{F}_{\rm pr}(\rho)\rangle$, standard deviation $\sigma_{\rm pr}(\rho)$, skewness, and kurtosis. In practice, we find that the skewness and kurtosis are relatively small (as might be expected, since our N is large), so for the rest of this section we will approximate $p(2\mathcal{F}_{\rm pr}; \rho^2)$ as simply a Gaussian with our measured mean and standard deviation. Given these distributions, it is completely straightforward to determine the false alarm probability FA for any threshold value $2\mathcal{F}^{{\rm th}_{\rm pr}}$ for any pair of $2\mathcal{F}^{{\rm th}_{\rm pr}}$.

We expect that, in practical searches, $\mathcal{F}_{\rm pr}$ will find its main use in hierarchical search algorithms, in which a very coarse search at relatively low threshold identifies candidates for further examination, and these are winnowed down in successive stages [12,13]. In this context, one generally wants a fairly small FD rate ($< 1\%$, say), so as not to lose any events, and strongly prefers a very low FA probability, to reduce the computational cost of follow-ups. With this application in mind, in Figs. 2 and 3. we plot FA as a function of FD, for several values of ρ^2.

How much does employing $\mathcal{F}_{\rm pr}$ increase the sensitivity of a blind GW pulsar search? Obtaining a useful and accurate answer to this question is much more complicated than it might initially seem, since the most sensitive known search algorithms for GW pulsars are hierarchical searches [12,13]. These searches involve several stages, with successive stages ruling out" an ever-increasing fraction of the parameter space. We imagine that the most sensitive search–at some fixed, realistic computational cost–might use $\mathcal{F}_{\rm pr}$ for only some of its stages. And while calculating $\mathcal{F}_{\rm pr}$ clearly requires more floating point operations than calculating $\mathcal{F}_{\rm sc}$, it is premature to compare these costs in a detailed way, since (i) there has as yet been no attempt to speed up our iterative relaxation scheme, and (ii) the extra information that comes with the phase offsets δ_i can

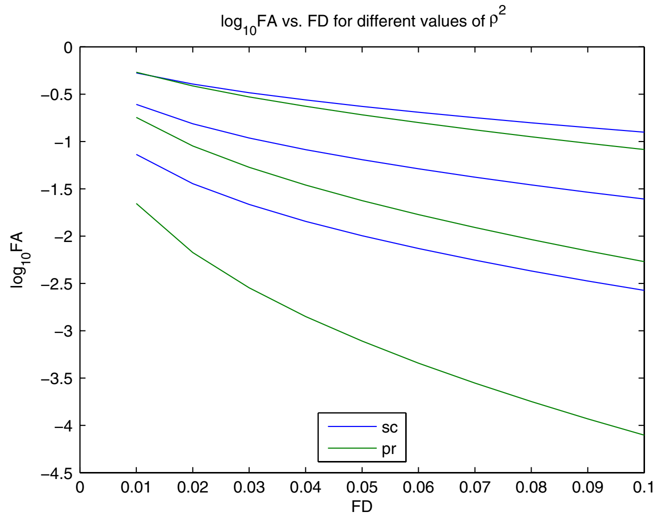

FIG. 2 (color online). Compares the FA probabilities for the two detection statistics, $\mathcal{F}_{\rm sc}$ and $\mathcal{F}_{\rm pr}$, as a function of FD probability, for several values of the (square of) the signal strength, ρ^2. The blue (dark) curves are for $\mathcal{F}_{\rm sc}$ and the green (light) for $\mathcal{F}_{\rm pr}$. From upper to lower, the signal strengths are $\rho^2 = 75$, 100, 125.

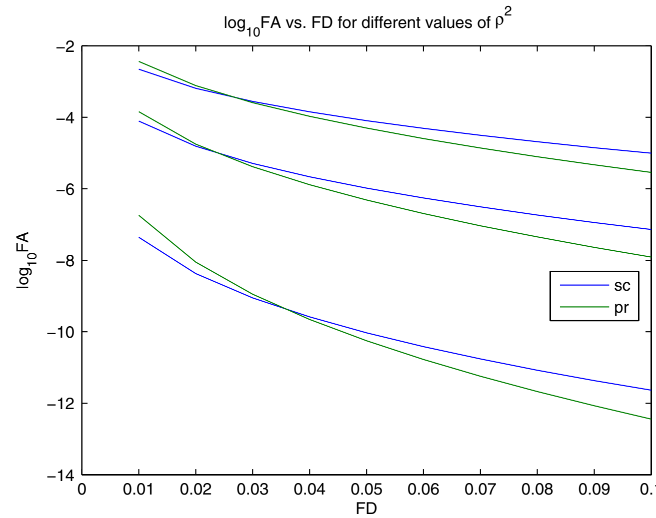

FIG. 4 (color online). Similar to Figs. 2 and 3 except that here the sensitivity of the $\mathcal{F}_{\rm sc}$ statistic is compared to that of $\mathcal{F}_{\rm pr}$ at *lower* ρ^2. From upper to lower, the signal strengths for $\mathcal{F}_{\rm sc}$ are for $\rho^2 = 140$, 160, 190.

presumably be used to speed up other parts of the search. Despite these difficulties we can obtain a rough estimate of the sensitivity improvement from using $\rho^2_{\rm pr}$, as follows. We have seen that given some detection statistic $\mathcal{S}$, one naturally obtains a map from the total ρ^2 to curves in the FA-FD plane. Let $\mathcal{C}(\mathcal{S}; \rho^2)$ denote that curve. Then we can look for pairs $\rho^2_{\rm pr}$ and $\rho^2_{\rm sc}$ such that $\mathcal{C}(\mathcal{F}_{\rm pr}; \rho^2_{\rm pr})$ lies close to $\mathcal{C}(\mathcal{F}_{\rm sc}; \rho^2_{\rm sc})$. Three such pairs are shown in Fig. 4. We see that, in the most relevant portion of $FD - FA$ plane, $\mathcal{C}(\mathcal{F}_{\rm pr}; 140)$ lies close to $\mathcal{C}(\mathcal{F}_{\rm sc}; 170)$, $\mathcal{C}(\mathcal{F}_{\rm pr}; 160)$ lies close to $\mathcal{C}(\mathcal{F}_{\rm sc}; 200)$, and $\mathcal{C}(\mathcal{F}_{pr}; 190)$ lies close to $\mathcal{C}(\mathcal{F}_{\rm sc}; 250)$. Thus, based on this example, we might estimate that using $\mathcal{F}_{\rm pr}$ rather than $\mathcal{F}_{\rm sc}$ affords an increase in sensitivity of ~20–30% in ρ^2, or ~10–15% in ρ. Clearly, to obtain a more reliable estimate we should perform a Monte Carlo simulation (based on random locations of the source on the sky, and random orientations of the GW pulsar's spin axis). We plan to do this in follow-up work.

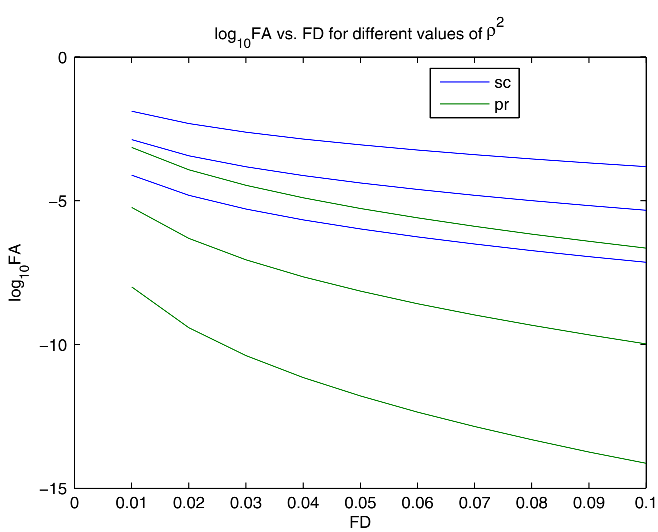

FIG. 3 (color online). Same as in Fig. 2, except that, from the upper to lower curves, the signal strengths are $\rho^2 = 150$, 175, 200.

Comparing Fig. 1 to Fig. 2, we see that the sensitivity gain from replacing the detection statistic $\mathcal{F}_{\rm sc}$ with $\mathcal{F}_{\rm pr}$ is qualitatively similar to the gain from $\mathcal{F}_{4N} \rightarrow \mathcal{F}_{2N}$, but that the latter gain is greater (at least for our one numerical example, and for total SNR ~10). That may seem surprising, since in the former case we are eliminating $3N - 3$ redundant parameters, while in the latter case we are eliminating only $2N$ redundant parameters. We conjecture that the main reason that the replacement $\mathcal{F}_{\rm sc} \rightarrow \mathcal{F}_{\rm pr}$ buys us less" in sensitivity is the following: In calculating $\mathcal{F}_{2N}$, the noise contributions from different intervals i still get combined incoherently. However in $\mathcal{F}_{\rm pr}$, the maximization over the phase offsets δ_i allows the noise contributions to combine coherently. Indeed for $\rho_i \lesssim 1$, the offsets δ_i are determined more by the noise than by the imbedded waveform. However we leave a thorough investigation of this effect to future work.

VI. SUMMARY AND FUTURE WORK

In this paper we defined a phase-relaxed $\mathcal{F}$-statistic, denoted $\mathcal{F}_{\rm pr}$, to be used in cases where the total observation time is sufficiently long that straightforward calculation of the fully coherent $\mathcal{F}$-statistic over the relevant parameter space is computationally intractable. The calculation of $\mathcal{F}_{\rm pr}$ takes as input the results from coherent searches over N shorter time intervals. Our $\mathcal{F}_{\rm pr}$ coincides with the fully coherent $\mathcal{F}$-statistic under the approximation that the phase offsets between template and imbedded signal are treated as an additional $N-1$ independent parameters. We also demonstrated one efficient, iterative method for calculating $\mathcal{F}_{\rm pr}$. We regard our iterative method as an existence proof'' for efficient algorithms. In future work we intend to explore variations on our basic method that we suspect would lead to substantial improvements in computational cost. We illustrated the use of $\mathcal{F}_{\rm pr}$ in one simple family of examples, in which the sensitivity improvement (compared to $\mathcal{F}_{\rm sc}$) was shown to be $\sim$10–15%.

Our example was based on the worst-case'' assumption that the phase offsets δ_i are completely random. In follow-on work we intend to examine the more realistic case where the δ_i are a (reasonably low-order) polynomial in time, and we plan to calculate the increased sensitivity based on a very large number of cases, in Monte Carlo fashion.

Other follow-on projects that we intend to work on include (i) the development of new vetoes [14] for instrumental artifacts, since, e.g., for true GW pulsars the parameters (h_0, ι, ψ) calculated from the first half of the observation should be consistent with those calculated from the second half; (ii) the use of the δ_i to quickly converge on improved estimates for the GW pulsar's frequency and spin-down parameters, and (iii) the optimal use of $\mathcal{F}_{\rm pr}$ in multistage, hierarchical searches.

Finally, while our primary interest in this paper has been the application of the $\mathcal{F}_{\rm pr}$-statistic to GW pulsar searches, it has not escaped our notice that the same idea and formalism can be applied, with only trivial modifications, to searches for (quasicircular, nonprecessing) inspiraling binaries. With binaries, we expect $\mathcal{F}_{\rm pr}$ to be most useful in those cases where the observed GW signal has a very large number of cycles; e.g., searches for neutron-star binaries by proposed GW detectors that have reasonable sensitivity at $\sim$1 Hz, such as the Einstein Telescope, Decigo, or the Big Bang Observer.

ACKNOWLEDGMENTS

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract to the National Aeronautics and Space Administration. We gratefully acknowledge support from NSF Grant No. PHY-0601459. We also thank Michele Vallisneri, Holger Pletsch, Reinhard Prix, Badri Krishnan, and Bruce Allen for helpful discussions.

[1] P. Jaranowski, A. Krolak, and B. F. Schutz, Phys. Rev. D **58**, 063001 (1998).
[2] C. Cutler and E. E. Flanagan, Phys. Rev. D **49**, 2658 (1994).
[3] C. Cutler and B. F. Schutz, Phys. Rev. D **72**, 063006 (2005).
[4] R. Prix and B. Krishnan, Classical Quantum Gravity **26**, 204013 (2009).
[5] B. Abbott *et al.*, Phys. Rev. D **80**, 042003 (2009).
[6] V. Dergachev, Classical Quantum Gravity **27**, 205017 (2010).
[7] H. J. Pletsch, Phys. Rev. D **83**, 122003 (2011).
[8] K. S. Thorne, in *300 Years of Gravitation*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1987), pp. 330–458.
[9] L. S. Finn and D. F. Chernoff, Phys. Rev. D **47**, 2198 (1993).
[10] S. Bonazzola and E. Gourgoulhon, Astron. Astrophys. **312**, 675 (1996).
[11] P. Brady, T. Creighton, C. Cutler, and B. F. Schutz, Phys. Rev. D **57**, 2101 (1998).
[12] P. Brady and T. Creighton, Phys. Rev. D **61**, 082001 (2000).
[13] C. Cutler, I. Gholami, and B. Krishnan, Phys. Rev. D **72**, 042004 (2005).
[14] Y. Itoh, M. A. Papa, B. Krishnan, and X. Siemens, Classical Quantum Gravity **21**, S1667 (2004).