Hidden Markov model tracking of continuous gravitational waves from young supernova remnants

L. Sun, A. Melatos, S. Suvorova, W. Moran and R. J. Evans

1 ozGrav, University of Melbourne, Parkville, Victoria 3010, Australia
2 School of Electrical and Computer Engineering, RMIT University, Melbourne, Victoria 3000, Australia
3 Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Victoria 3010, Australia

(Dated: September 3, 2018)

Abstract: Searches for persistent gravitational radiation from nonpulsating neutron stars in young supernova remnants (SNRs) are computationally challenging because of rapid stellar braking. We describe a practical, efficient, semi-coherent search based on a hidden Markov model (HMM) tracking scheme, solved by the Viterbi algorithm, combined with a maximum likelihood matched filter, the $F$-statistic. The scheme is well suited to analyzing data from advanced detectors like the Advanced Laser Interferometer Gravitational Wave Observatory (Advanced LIGO). It can track rapid phase evolution from secular stellar braking and stochastic timing noise torques simultaneously without searching second- and higher-order derivatives of the signal frequency, providing an economical alternative to stack-slide-based semi-coherent algorithms. One implementation tracks the signal frequency alone. A second implementation tracks the signal frequency and its first time derivative. It improves the sensitivity by a factor of a few upon the first implementation, but the cost increases by two to three orders of magnitude.

PACS numbers: 95.85.Sz, 97.60.Jd

I. INTRODUCTION

Rotating neutron stars in young supernova remnants (SNRs) are plausible sources of quasi-monochromatic gravitational radiation detectable by ground-based interferometers such as the Laser Interferometer Gravitational Wave Observatory (LIGO) and the Virgo detector [1–3]. The emission is predicted to occur at a frequency proportional to the star's spin frequency $f_s$. A thermoelastic [4, 5] or magnetic [6, 8] mass quadrupole emits at $f_s$, and/or a $2f_s$, an r-mode current quadrupole emits at roughly $4f_s/3$, perturbed by an equation-of-state correction [9–12], and a current quadrupole produced by nonaxisymmetric circulation in a superfluid pinned to the crust emits at $f_s$ [13–16]. There are several reasons to devote attention to this source class. First, a young object has spent less time settling down since its birth; slow, diffusive processes like ohmic [17] or thermal [18] relaxation are still in the early stages of erasing nonaxisymmetries inherited at birth [11, 13, 19]. Second, the rapid spin down of a differentially rotating young neutron star excites high-Reynolds-number turbulence, which produces a time-varying current quadrupole moment [13, 20, 21]. Third, young radio pulsars are known to undergo glitches [22, 25], which are ascribed to differential rotation [26, 28] or starquakes [29] and can also lead to quadrupole moment variations.

Initial LIGO achieved its design sensitivity over a wide frequency band during Science Run 5 (S5) [30] and exceeded it during Science Run 6 (S6) [2]. Several continuous-wave searches targeting young SNRs have been carried out using Initial LIGO data. A directed, radiometer search for SNR 1987A was conducted in S5, yielding a 90% confidence strain upper limit $h_0^{90\%} = 7 \times 10^{-25}$ for a circularly polarized signal in the most sensitive band, near 160 Hz [31]. A semi-coherent cross-correlation search for the same target in S5 data improved the upper limit to $h_0^{90\%} = 3.8 \times 10^{-25}$ [32]. A coherent search for Cassiopeia A (Cas A) was conducted on a 12-day stretch of S5 data in the band 100–300 Hz, yielding $h_0^{90\%} = 10^{-24}$ in the range 0.7–1.2 [33]. The upper limit for Cas A was improved by approximately a factor of two by a semi-coherent Einstein@Home search in S6. The search was conducted in a broad frequency band 50–1000 Hz, yielding the best $h_0^{90\%} \approx 2.9 \times 10^{-25}$ at 170 Hz [34]. In S6, directed searches were conducted for nine nonpulsating X-ray sources (central compact objects) in SNRs with the maximum likelihood matched filter called the $F$-statistic. The searches combine multi-detector data coherently over 5.3 to 25.3 days, yielding strain upper limits in the range $3.7 \times 10^{-25} \leq h_0^{95\%} \leq 6.4 \times 10^{-25}$ [35]. The foregoing upper limits are slated to improve in the future. The strain noise in the first observation run (O1) of Advanced LIGO is 3–4 times lower than in S6 across the most sensitive band, between 100 Hz and 300 Hz, and $\sim 10^2$ times lower around 50 Hz [36]. The sensitivity is expected to improve roughly two-fold relative to O1 after further upgrades [37].

The rapid spin down of young neutron stars is a serious challenge for searches of the above kind. A large
number of matched filters (i.e. templates) is required to track a signal with rapid evolving phase, when a radio ephemeris is unavailable. For example, the $\mathcal{F}$-statistic search in Ref. [36] is restricted to $\leq 25.3\, \text{d}$ in order to keep the number of matched filters manageable. Semi-coherent methods have been developed [38–40], based on the stack-slide algorithm [41], to sum the signal power in multiple coherent segments after sliding the segments in the frequency domain to account for the phase evolution of the source. However, stack-slide searches are still computationally challenging, when high-order derivatives of the signal frequency enter the phase model. Intrinsic, stochastic $f_s$ wandering (‘timing noise’) also degrades the sensitivity of these searches.

In this paper, we introduce an approach based on a hidden Markov model (HMM) [42] to tackle the above challenge. A HMM tracks signals with time-varying, unobservable parameters by modeling them as hidden states in a Markov chain. The HMM relates the observed data to the hidden states through a likelihood statistic and infers the most probable sequence of hidden states. It incoherently combines the coherent matched filter outputs from data blocks of duration $T_{\text{drift}}$ (analogous to $T_{\text{span}}$ in Ref. [36]), during which the signal parameters are assumed to remain constant. The sensitivity scales approximately proportional to $\sim (T_{\text{obs}} T_{\text{drift}})^{-1/4}$, where $T_{\text{obs}}$ is the whole observation time. The Viterbi algorithm [43] provides a computationally efficient HMM solution. A HMM was applied to search for continuous gravitational radiation from the most luminous low-mass X-ray binary, Scorpius X-1, in O1 data, taking into account the effects of spin wandering caused by the fluctuating accretion torque [44].

The structure of the paper is as follows. In Section II, we describe the signal model, introduce the $\mathcal{F}$-statistic, and discuss the search parameter ranges. In Section III, we formulate the HMM tracking problem with one hidden state variable, describe how to choose $T_{\text{drift}}$, and discuss the impact of timing noise. In Section IV, we conduct Monte-Carlo simulations in Gaussian noise, present search examples, and estimate the sensitivity. In Section V, we introduce an alternative HMM formulation with two hidden state variables, and present abridged simulation examples. In Section VI, we discuss the trade-off between computing cost and sensitivity. We also discuss the special case of young objects, whose current spin frequencies are close to the value at birth. A summary of the conclusions is provided in Section VII.

II. COHERENT MATCHED FILTER

In this section we start by describing the signal model in Section II.A. We then review the maximum likelihood matched filter corresponding to the signal model, called the $\mathcal{F}$-statistic, in Section II.B, and discuss the signal phase parameter ranges in Section II.C.

A. Signal model

We consider a continuous gravitational wave signal from a rotating neutron star modeled as a biaxial rotor. The Doppler modulation of the observed signal frequency due to the motion of the Earth with respect to the solar system barycentre (SSB) is taken into consideration. The signal phase observed at the detector is then given by [45]

$$\Phi(t) = 2\pi \sum_{k=0}^{s} \frac{f_0^{(k)} t^{k+1}}{(k+1)!} + \frac{2\pi}{c} \hat{\nu} \cdot \vec{r}(t) \sum_{k=0}^{s} \frac{f_0^{(k)} t^k}{k!}, \tag{1}$$

where $f_0^{(k)}$ is the $k$-th time derivative of the signal frequency at $t = 0$, $\hat{\nu}$ is the unit vector pointing from the SSB to the neutron star, and $\vec{r}(t)$ is the position vector of the detector relative to the SSB.

The signal can be written in the form

$$h(t) = A^\mu h_\mu(t), \tag{2}$$

where $A^\mu$ denotes the amplitudes associated with the four linearly independent components

$$h_1(t) = a(t) \cos \Phi(t), \tag{3}$$

$$h_2(t) = b(t) \cos \Phi(t), \tag{4}$$

$$h_3(t) = a(t) \sin \Phi(t), \tag{5}$$

$$h_4(t) = b(t) \sin \Phi(t). \tag{6}$$

$a(t)$ and $b(t)$ are the antenna-pattern functions defined by Equations (12) and (13) in Ref. [45], and $\Phi(t)$ is the signal phase given by [1]. In [2], $A^\mu$ depends on the star’s inclination, wave polarization, initial phase at $t = 0$ and strain amplitude $h_0$.

B. $\mathcal{F}$-statistic

The time-domain data collected by a detector takes the form

$$x(t) = A^\mu h_\mu(t) + n(t), \tag{7}$$

where $n(t)$ stands for stationary, additive noise. The $\mathcal{F}$-statistic maximizes the likelihood of detecting a signal in data $x(t)$ with respect to $A^\mu$ [42]. We define a scalar product $\langle x \mid y \rangle$ as a sum over single-detector inner products,

$$\langle x \mid y \rangle = \sum_x \langle x^X \mid y^X \rangle \tag{8}$$

$$= \sum_x 4 \Re \int_0^{\infty} df \frac{\hat{x}^X(f) \hat{y}^X(f)}{S_h^X(f)}, \tag{9}$$

$^2$ Here we assume a perpendicular rotor emitting at $2f_0$, only for simplicity. A non-perpendicular rotor also emits at $f_*$, and hence eight components are involved. A full description of the signal model can be found in Ref. [45]. The emission spectrum of a triaxial rotor contains additional lines [38, 46].
where $X$ indexes the detector, $S_h(f)$ is the single-sided power spectral density (PSD) of detector $X$, the tilde denotes a Fourier transform, and $\mathbb{R}$ returns the real part of a complex number $[47]$. The $F$-statistic is expressed in the form

$$F = \frac{1}{2} x_\mu M^{\mu\nu} x_\nu,$$  \hspace{1cm} (10)

where we write $x_\mu = (x|h_\mu)$, and $M^{\mu\nu}$ denotes the matrix inverse of $M_{\mu\nu} = (h_\mu|h_\nu)$. We leverage the existing, fully tested software infrastructure in the LSC Algorithm Library Applications (LALApps) \[^3\] to compute $F$ as a function of frequency and its time derivatives over an interval of length $T_{\text{drift}}$ [49]. The software operates on the raw data collected by LIGO in the form of short Fourier transforms (SFTs), usually with length $T_{\text{SFT}} = 30$ min for each SFT. It provides options to search up to the third time derivative of frequency, $\dot{f}_0$. The implementations described in Section III and V use the options to search over $f_0$ and $(f_0, \dot{f}_0)$, respectively.

C. Search parameter ranges

The ranges of $\dot{f}_0$ and $f_0$ to be considered in defining the parameters of the search can be reexpressed in terms of the range of braking index $n = f_0 \dot{f}_0/\dot{f}_0$ and the spin-down age of the source $\tau$, given by $[34] [36]$

$$-\frac{f_0}{(n_{\text{min}} - 1)\tau} \leq \dot{f}_0 \leq -\frac{f_0}{(n_{\text{max}} - 1)\tau},$$  \hspace{1cm} (14)

and

$$\frac{n_{\text{min}}\dot{f}_0^2}{f_0} \leq \dot{f}_0 \leq \frac{n_{\text{max}}\dot{f}_0^2}{f_0}.$$  \hspace{1cm} (15)

Radio timing observations yield $1.4 \leq n \leq 3$ for all pulsars, where $f_0$ can be measured reliably by absolute pulse numbering [9] [50] (cf. [51]). Gravitational radiation in the mass and current quadrupole channels corresponds to $n = 5$ and $n = 7$, respectively. In this quick study, we assume $2 \leq n \leq 7$. Strictly speaking, the ranges obtained from (14) and (15) are wider than needed. In a real search, one should ideally calculate the $\dot{f}_0$ and $f_0$ ranges for each fixed $2 \leq n \leq 7$ and choose the widest $f_0$ and $\dot{f}_0$ ranges. Equations (14) and (15) are valid, provided that $f_0(t)$ during the observation is much smaller than $f_0$ at birth, $f_{0\text{birth}}$. The regime $f_0(t) \sim f_{0\text{birth}}$ is considered in Section VI.

III. HMM TRACKING OF $f_0$

We begin this section by reviewing briefly the use of HMM tracking in gravitational wave searches (Section III A). We formulate the tracker as a one-dimensional HMM with a single hidden variable $f_0$ (Section III B), discuss the coherent drift time-scale $T_{\text{drift}}$ (Section III C), and illustrate that the method can track secular spin down and stochastic timing noise simultaneously (Section III D). A full description of HMMs can be found in Ref. [48]. The Viterbi algorithm used for solving the HMM is described in Appendix A.

A. HMM formulation

A Markov chain is a stochastic process transitioning between discrete states at discrete times $\{t_0, \cdots, t_{N_T}\}$. A HMM is an automaton, in which the state variable $q(t) \in \{q_1, \cdots, q_{N_q}\}$ is hidden (unobservable), and the measurement variable $o(t) \in \{o_1, \cdots, o_{N_o}\}$ is observable. The hidden state at time $t_{n+1}$ only depends on the state at time $t_n$ with transition probability

$$A_{q_i,q_j} = P[q(t_{n+1}) = q_j|q(t_n) = q_i].$$  \hspace{1cm} (16)

The hidden state $q_i$ is observed in state $o_j$ with emission probability

$$L_{o_j,q_i} = P[o(t_n) = o_j|q(t_n) = q_i].$$  \hspace{1cm} (17)

Given the prior defined by

$$\Pi_{q_i} = P[q(t_0) = q_i],$$  \hspace{1cm} (18)

the probability that the hidden state path $Q = \{q(t_0), \cdots, q(t_{N_T})\}$ gives rise to the observed sequence $O = \{o(t_0), \cdots, o(t_{N_T})\}$ via a Markov chain equals

$$P(Q|O) = L_{o(t_{N_T})q(t_{N_T})} A_{q(t_{N_T})q(t_{N_T-1})} \cdots L_{o(t_1)q(t_1)} A_{q(t_1)q(t_0)} \Pi_{q(t_0)}.$$  \hspace{1cm} (19)
The most probable path maximizing $P(Q|O)$, viz.

$$Q^*(O) = \arg \max P(Q|O),$$  

(20)
gives the best estimate of $q(t)$ over the total observation, where $\arg \max(\cdots)$ returns the argument that maximizes the function $(\cdots)$.

B. Transition and emission probabilities

We consider the one-dimensional hidden state variable $q(t) = f_0(t)$. The discrete hidden states are mapped one-to-one to the frequency bins in the output of a frequency-domain estimator $F(f_0)$ computed over an interval of length $T_{\text{drift}}$, with bin size $\Delta f_0$ selected using the metric described in Appendix A. For simplicity here, we take $\Delta f_0 = 1/(2T_{\text{drift}})$, with mismatch $m \leq 0.2$. The mismatch $m$ is defined as the fractional reduction of $F$-statistic power caused by discrete parameter sampling (see Appendix B). We choose $T_{\text{drift}}$ as described in Section III C to satisfy

$$|f_0(t)| T_{\text{drift}} = \delta f_0 = \frac{1}{2T_{\text{drift}}},$$  

(26)
to satisfy (27). Hence $T_{\text{drift}} = (2|f_0|)^{-1/2}$ depends solely on the spin-down rate of the source. To illustrate how $T_{\text{drift}}$ is determined by $\tau$, we calculate $f_0$ and hence $T_{\text{drift}}$ by assuming purely electromagnetic spin down ($f_0(t) \propto B_0^2 f_0(t)^{n}, n = 3$) as an example, where $B_0$ is the birth magnetic field strength; see Appendix C for a detailed derivation. Figure 1 displays contours of $T_{\text{drift}}$ as a function of the gravitational-wave signal frequency today, $f_{0\text{today}}$, and the ratio $f_{0\text{today}}/f_{0\text{birth}}$, where $f_{0\text{birth}}$ is the signal frequency at the birth of the star. The contours in the upper left panel (e.g., SNR 1987A; $\tau = 0.03$ kyr) show $T_{\text{drift}} \lesssim 5$ hr for most of the plot. By contrast, the contours in the upper right panel (e.g., Cas A; $\tau = 0.34$ kyr) satisfy $T_{\text{drift}} > 10$ hr over $\approx 1/3$ area of the plot. The $T_{\text{drift}}$ values estimated for older objects with $\tau = 0.6$ kyr and 1 kyr are plotted in the lower panels. All panels show that $T_{\text{drift}}$ decreases significantly for $f_{0\text{today}}/f_{0\text{birth}} \approx 0.8$ and $f_{0\text{today}} \gtrsim 100$ Hz.

Figure 1 is plotted for $n = 3$. In practice, the range of $f_0$ is given by equation (14) for $2 \leq n \leq 7$. We substitute $|f_0|_{\text{max}}$ into (26) to satisfy (27) for all $f_0$ and obtain

$$T_{\text{drift}} = (2|f_0|_{\text{max}})^{-1/2}.$$  

(27)

Hence for any given source with spin-down age $\tau$, we can always choose a coherent duration $T_{\text{drift}}$ and divide the data into $N_T = T_{\text{obs}}/T_{\text{drift}}$ coherent segments, without searching the $(f_0, f_0', \cdots)$ parameter space.

D. Timing noise

Stochastic spin wandering (often termed ‘‘timing noise’’) is a widespread phenomenon in isolated neutron stars pulsating at radio and X-ray wavelengths [52–54]. The auto-correlation time-scale ranges from days to years [54, 56]. The phenomenon can result from magneto-spheric changes [57], superfluid dynamics in the stellar interior [34, 55, 56], spin microjumps [61, 62], and fluctuations in the spin-down torque [63, 65].

In the absence of a measured ephemeral, a HMM can track the evolution of $f_0(t)$ caused by both secular spin down and stochastic timing noise [58]. We approximate the timing noise by an unbiased random walk, in which $f_0(t)$ moves by at most one bin up or down during the timing-noise time-scale $T_{\text{drift}}$. The transition probability matrix for timing noise only is

$$A_{q_{i-1}q_i}' = A_{q_{q_i}}'(q_{q_i+1}q_i) = \frac{1}{3}.\quad (28)$$

For $T_{\text{drift}} \gg T_{\text{drift}}$, we can neglect timing noise, as discussed in previous sections. For $T_{\text{drift}} \ll T_{\text{drift}}$, we can...
FIG. 1. Contours of $T_{\text{drift}}$ (in hours) for targets with $\tau = 0.03 \text{ kyr}$ (top left), 0.34 kyr (top right), 0.6 kyr (bottom left) and 1 kyr (bottom right) as a function of the gravitational-wave signal frequency today $f_{0,\text{today}}$ and the ratio $f_{0,\text{today}} / f_{0,\text{birth}}$, where $f_{0,\text{birth}}$ is the signal frequency at the birth of the star, assuming pure electromagnetic spin down ($\dot{f}_0 \propto B_0^2 f_0^{n/3}$).

IV. SIMULATIONS AND SENSITIVITY

In this section, we firstly introduce the Viterbi score and the associated, score-based detection threshold in Section IV A. We demonstrate the performance of $f_0$ tracking for three different scenarios using synthetic data: (1) an older object with weak timing noise (Section IV B); (2) an older object with strong timing noise (Section IV C); and (3) a very young object (Section IV D). Simulations are conducted in an artificially restricted, 1-Hz sub-band to save time. The signals are injected at a fixed sky position, and $S_h(f)$ is set to Advanced LIGO’s design sensitivity [66]. We show the results in detail with a set of injections into Gaussian noise for each of the three scenarios, where the polarization and inclination angles and initial phase are arbitrarily chosen and fixed. Monte-Carlo simulations are conducted to generate the receiver operator characteristic (ROC) curves (Section IV E).

A. Viterbi score and threshold

The Viterbi score $S$ is defined, such that the log likelihood of the optimal Viterbi path equals the mean log likelihood of all paths plus $S$ standard deviations at $t_{N_T}$, viz.

$$S \equiv \frac{\ln \delta^*(t_{N_T}) - \mu_{\ln \delta}(t_{N_T})}{\sigma_{\ln \delta}(t_{N_T})}$$

(29)
with

$$\mu_{in}(t_{N_T}) = N_Q^{-1} \sum_{i=1}^{N_Q} \ln \delta_{q_i}(t_{N_T})$$

and

$$\sigma_{in}(t_{N_T})^2 = N_Q^{-1} \sum_{i=1}^{N_Q} [\ln \delta_{q_i}(t_{N_T}) - \mu_{in}(t_{N_T})]^2,$$

where $\delta_{q_i}(t_{N_T})$ denotes the maximum probability of the path ending in state $q_i$ ($1 \leq i \leq N_Q$) at step $N_T$ (see Appendix A), and $\delta_{q}(t_{N_T})$ is the likelihood of the optimal Viterbi path, i.e. $P[Q^*(O)|O]$. In a real search, we normally divide the full frequency band into multiple 1-Hz sub-bands to allow parallelized computing. In each 1-Hz sub-band, we consider the candidate for follow-up and further scrutiny, if $S$ exceeds a threshold $S_{th}$ set by the desired false alarm and false dismissal probabilities. The value of $S_{th}$ varies with $N_T$, $N_Q$, and the entries in $A_{q/q_i}$. Systematic Monte-Carlo simulations are always required in practice to calculate $S_{th}$ for each HMM implementation.

For the three scenarios in Sections IV B–IV D, $S_{th}$ is determined as follows. Searches are conducted on data sets containing pure Gaussian noise in 1-Hz sub-bands. For a given false alarm probability $P_a$ in a 1-Hz sub-band, the value of $S$ yielding a fraction $P_a$ of positive detections is $S_{th}$. The false alarm probability in a search over band $B$ is given by $P_{a\text{total}} = 1 - (1 - P_a)^B$. We set $P_a = 1\%$ and generate $10^3$ noise realizations for each scenario. Searches for the first two scenarios are based on the same $N_T$ and $N_Q$, and hence they both yield $S_{th} = 6.7$. The mean and standard deviation of $S$ in the $10^3$ realizations are $\mu_S = 5.5$ and $\sigma_S = 0.4$. In the last scenario, we have $S_{th} = 0.8$, with $\mu_S = 0.63$ and $\sigma_S = 0.06$. Because $\mu_{in}(t_{N_T})$ increases as $N_T$ grows, resulting lower $S$ normalized by $\sigma_{in}(t_{N_T})$ in (29), it is as expected that the $S_{th}$ is much lower in the last scenario ($N_T = 2000$) than the first two scenarios ($N_T = 40$).

B. $\tau \gtrsim 5 \text{ kyr}$, $T_{drift} \gg T_{drift}$

In the first group of tests, we consider a relatively older target with low timing noise, e.g., $\tau \gtrsim 5 \text{ kyr}$ and $T_{drift} \gg T_{drift}$. Four sets of synthetic data, containing injected signals with $h_0/10^{-26} = 10$, 5, 3, and 2, are generated for $T_{obs} = 83.3$ d at two detectors (the LIGO Hanford and Livingston observatories) using Makefakedata version 4 from LALApps. Detailed injection parameters are shown in Table II. The searches are conducted using the search parameters in Table II and $A_{q/q_i}$ in (22). The detection is deemed successful for $S > S_{th} = 6.7$. The results in Table III show that signals with $h_0 \geq 3 \times 10^{-26}$ are detected. We calculate the root-mean-square error (RMSE) $\varepsilon_{f_0}$ in $f_0$ between the optimal Viterbi path and the injected signal (in Hz and in units of $\Delta f_0$). All successful detections yield $\varepsilon_{f_0} < \Delta f_0$. The errors are introduced mostly because the HMM takes discrete values of $f_0$ (i.e., $\Delta f_0$ is the smallest step size), while the injected signal $f_0(t)$ can take any value within a bin.

Figure 2 presents the tracking results corresponding to Table III. Panels (a)–(c) show that the optimal Viterbi path is match the injected $f_0(t)$ closely, with $\varepsilon_{f_0} = 0.39\Delta f_0$, $0.46\Delta f_0$ and $0.59\Delta f_0$, respectively. Panel (d) shows that the signal is not tracked successfully. The detectability drops rapidly from $h_0 = 3 \times 10^{-26}$ to $h_0 = 2 \times 10^{-26}$, as expected near the detection limit (see detailed explanation in Section III B of Ref. [48]).

| Parameter          | Symbol | Value  |
|--------------------|--------|--------|
| Right ascension    | $\alpha$ | 23h 23m 26.0s |
| Declination        | $\delta$ | 58° 48’ 0.0” |
| Polarization angle | $\psi$ | 4.94278 rad |
| Inclination angle  | $\cos \iota$ | 0.718742 rad |
| Initial phase      | $\phi_0$ | 2.43037 rad |
| PSD                | $S_h(f)^{1/2}$ | $4 \times 10^{-24}$ Hz$^{-1/2}$ |
| Frequency          | $f_{0inj}$ | 151.23456789 Hz |
| First derivative of $f_{0inj}$ | $\dot{f}_{0inj}$ | $-1.0 \times 10^{-11}$ Hz$^{-1}$ |
| Second derivative of $f_{0inj}$ | $\ddot{f}_{0inj}$ | $2.0 \times 10^{-24}$ Hz$^{-2}$ |

TABLE I. Injection parameters used to create the synthetic data analyzed in Sections IV B and IV C

| Parameter | Value | Unit |
|-----------|-------|------|
| $f_0$    | 151–152 | Hz |
| $T_{drift}$ | 50 | hr |
| $\Delta f_0$ | $2.78 \times 10^{-6}$ | Hz |
| $T_{obs}$ | 83.3 | d |
| $N_T$ | 40 | – |

TABLE II. Search parameters for the synthetic signals with injection parameters quoted in Table II

| $h_0$ ($10^{-26}$) | Detect? | $S$ | $\varepsilon_{f_0}$ (Hz) | $\varepsilon_{f_0}/\Delta f_0$ |
|---------------------|---------|-----|-----------------------------|-------------------------------|
| 10.0                | ✓       | 90.9 | $1.07 \times 10^{-6}$       | 0.39                          |
| 5.0                 | ✓       | 32.0 | $1.29 \times 10^{-6}$       | 0.46                          |
| 3.0                 | ✓       | 10.6 | $1.64 \times 10^{-6}$       | 0.59                          |
| 2.0                 | ×       | 5.5  | 0.46                        | $1.6 \times 10^5$             |

TABLE III. Results of $f_0$ tracking for synthetic signals with the injection parameters in Table II $T_{obs} = 83.3$ d, $T_{drift} = 50$ hr, and characteristic wave strain $h_0$. The RMSE $\varepsilon_{f_0}$ between the optimal Viterbi path and the injected $f_0(t)$ is quoted in Hz and in units of $\Delta f_0 = 2.78 \times 10^{-6}$ Hz. The third column quotes the Viterbi score $S$. 
FIG. 2. Injected $f_0(t)$ (blue curve) and optimal Viterbi path (red curve) for the injected signals in Table III. Panels (a)–(d) display paths for $h_0/10^{-26} = 10, 5, 3, 2$, respectively. Good matches are obtained in (a)–(c), with $\varepsilon f_0 = 0.39\Delta f_0, 0.46\Delta f_0,$ and $0.59\Delta f_0$, respectively. In (d), the signal is not detected; the spin down of $f_0(t)$ is too slow to be seen in the plot ($\varepsilon f_0 \gg f_0T_{\text{obs}} = 7 \times 10^{-7}$ Hz). The horizontal axes are in units of HMM steps with $T_{\text{drift}} = 50$ hr for each step ($N_T = 40$).

C. $\tau \gtrsim 5$ kyr, $T'_{\text{drift}} \approx T_{\text{drift}}$

In the second group of tests, we show that the HMM can track secular spin down and timing noise simultaneously for synthetic signals injected in Gaussian noise. As an example, we assume the time-scale of the unbiased random walk is the same as the spin-down time-scale, i.e., $T'_{\text{drift}} = T_{\text{drift}} = 50$ hr. The modified transition probability matrix is the product of (22) and (28), given by

$$2A_{q_{i-2}q_i} = A_{q_{i-1}q_i} = A_{q_iq_i} = 2A_{q_{i+1}q_i} = \frac{1}{3},$$

(32)

with all other terms being zero.

Four data sets with signal strains $h_0/10^{-26} = 10, 5, 3$ and 2 are generated for $T_{\text{obs}} = 83.3$ d at two detectors. We use the same injection parameters in Table I at $t = 0$. In addition to the spin down, $f_0(t)$ wanders randomly by at most $\pm \Delta f_0$ over time-scale $T_{\text{drift}}$. The data sets are searched using the parameters in Table II and $A_{q_i/q_i}$ in (32). The results are shown in Table IV. The detection is deemed successful for $S > S_{th} = 6.7$. All successful detections yield $\varepsilon f_0 < \Delta f_0$.

Figure 3 presents the tracking results for signals with $h_0/10^{-26} = 10, 5, 3, 2$ in Table IV. The optimal Viterbi paths in panels (a)–(c) match the injected paths closely, indicating successful detections. The RMSE $\varepsilon f_0$ increases from $0.38\Delta f_0$ to $0.77\Delta f_0$, when $h_0$ decreases from $1 \times 10^{-25}$ to $3 \times 10^{-26}$. Panel (d) shows that the optimal Viterbi path does not match the injected $f_0(t)$ for $h_0 = 2 \times 10^{-26}$, i.e., the injected signal is not detected.
FIG. 3. Injected \( f_0(t) \) (blue curve) and optimal Viterbi path (red curve) for the injected signals in Table IV. Panels (a)–(d) display paths for \( h_0/10^{-26} = 10, 5, 3, 2 \), respectively. Good matches are obtained in (a)–(c), with \( \varepsilon_f = 0.38\Delta f_0, 0.59\Delta f_0, \) and \( 0.77\Delta f_0 \), respectively. In (d), the signal is not detected; the spin down of \( f_0(t) \) is too slow to be seen in the plot \( (\varepsilon_f \gg \dot{\varepsilon}_f \approx f_0 T_{\text{obs}} \sim 7 \times 10^{-5} \text{Hz}) \). The horizontal axes are in units of HMM steps with \( T_{\text{drift}} = 50 \text{ hr} \) for each step \( (N_T = 40) \).

 Given the signal-to-noise ratio, \( T_{\text{drift}} \) and \( N_T \), the sensitivity of the search remains the same for \( T'_{\text{drift}} \gg T_{\text{drift}} \) (Section IV B) and \( T'_{\text{drift}} = T_{\text{drift}} \) (Section IV C).

### Table IV. Results of \( f_0 \) tracking for synthetic spin-wandering signals with the injection parameters in Table I, \( T_{\text{obs}} = 83.3 \text{ d}, T_{\text{drift}} = 50 \text{ hr}, \) and characteristic wave strain \( h_0 \). The RMSE \( \varepsilon_f \) between the optimal Viterbi path and the injected \( f_0(t) \) is quoted in Hz and in units of \( \Delta f_0 = 2.78 \times 10^{-6} \text{Hz} \). The third column quotes the Viterbi score \( S \).

| \( h_0 (10^{-26}) \) | Detect? | \( S \) | \( \varepsilon_f \) (Hz) | \( \varepsilon_f/\Delta f_0 \) |
|-----------------|--------|----------|----------------|------------------|
| 10.0            | ✔      | 92.5     | 1.06 \times 10^{-6} | 0.38             |
| 5.0             | ✔      | 28.4     | 1.63 \times 10^{-6} | 0.59             |
| 3.0             | ✔      | 7.2      | 2.14 \times 10^{-6} | 0.77             |
| 2.0             | ×      | 5.1      | 0.29             | 1.1 \times 10^5  |

In the third group of tests, we consider a very young object with \( \tau \lesssim 0.03 \text{kyr} \), e.g., SNR 1987A. In stack-slide-based searches for such young sources, typically four or more frequency derivatives must be searched in order to accurately track the rapid phase evolution.

We inject four signals with \( h_0/10^{-26} = 15, 13, 11, 10 \) and high spin-down rates as quoted in Table V. Other injection parameters remain the same as those in Table...
TABLE V. Spin-down-related injection parameters used to create the synthetic data analyzed in Section V D

| Parameter | Symbol | Value       | Unit   |
|-----------|--------|-------------|--------|
| First derivative of \( f_{0\text{inj}} \) | \( \tilde{f}_{0\text{inj}} \) | \(-3.0 \times 10^{-8} \text{ Hz s}^{-1}\) |        |
| Second derivative of \( f_{0\text{inj}} \) | \( \tilde{\tilde{f}}_{0\text{inj}} \) | \(3.0 \times 10^{-17} \text{ Hz s}^{-2}\) |        |
| Third derivative of \( f_{0\text{inj}} \) | \( \tilde{	ilde{	ilde{f}}}_{0\text{inj}} \) | \(-3.0 \times 10^{-26} \text{ Hz s}^{-3}\) |        |

TABLE VI. Search parameters for the synthetic signals with injection parameters quoted in Tables I and V

| Parameter | Value | Unit |
|-----------|-------|------|
| \( f_0 \) | 151–152 | Hz   |
| \( T_{\text{drift}} \) | 1 | hr   |
| \( \Delta f_0 \) | 1.39 \times 10^{-4} | Hz |
| \( T_{\text{obs}} \) | 83.3 | d    |
| \( N_T \) | 2000 | –    |

We choose \( T_{\text{drift}} = 1 \text{ hr} \) to satisfy equation (21). In this case, we always have \( T_{\text{inj}} \geq T_{\text{drift}} \) and hence use \( A_{q,\psi} \) in (22). The search parameters and results are presented in Tables VI and VII, respectively. In this group, the detection is deemed successful for \( S > S_{\text{th}} \geq 0.8 \) given \( P_a = 1\% \). The successful detections yield \( \varepsilon_{f_0} < 2\Delta f_0 = 2.8 \times 10^{-4} \text{ Hz} \). We tolerate \( \varepsilon_{f_0} \) slightly larger than \( \Delta f_0 \) because \( T_{\text{drift}} \) is relatively short.

Figure 1 shows the tracking results corresponding to Table VII. Panels (a)–(c) show that the optimal Viterbi path matches the injected \( f_0(t) \) closely. The discrepancy between the optimal Viterbi path and the injected \( f_0(t) \) can hardly be seen, because \( \varepsilon_{f_0} \sim 10^{-4} \text{ Hz} \) is much smaller than the total change in \( f_0 \) over \( T_{\text{obs}} \) (≈ 0.2 Hz). Panel (d) shows that the signal is not detected for \( h_0 = 10 \times 10^{-26} \), with \( \varepsilon_{f_0} = 151.8\Delta f_0 \gg \Delta f_0 \).

TABLE VII. Results of \( f_0 \) tracking for synthetic signals with the injection parameters quoted in Tables I and V. \( T_{\text{obs}} = 83.3 \text{ d} \), \( T_{\text{drift}} = 1 \text{ hr} \), and characteristic wave strain \( h_0 \). The RMSE \( \varepsilon_{f_0} \) between the optimal Viterbi path and the injected \( f_0(t) \) is quoted in Hz and in units of \( \Delta f_0 = 1.39 \times 10^{-4} \text{ Hz} \). The third column quotes the Viterbi score \( S \).

| \( h_0 \) (10^{-26}) | Detect? | \( S \) | \( \varepsilon_{f_0} \) (Hz) | \( \varepsilon_{f_0}/\Delta f_0 \) |
|-----------------|---------|------|-----------------|-----------------|
| 15.0            | ✓       | 3.0  | 1.77 \times 10^{-4} | 1.3             |
| 13.0            | ✓       | 2.1  | 2.50 \times 10^{-4} | 1.8             |
| 11.0            | ✓       | 0.9  | 2.47 \times 10^{-4} | 1.8             |
| 10.0            | ×       | 0.5  | 0.02             | 151.8           |

E. ROC curve and sensitivity

The detection threshold \( S_{\text{th}} \) is set by \( P_a \). The probability that an injected signal yields \( S < S_{\text{th}} \) is the false dismissal probability, denoted by \( P_d \). We quantify the performance of the HMM in terms of its ROC curve, plotting the detection probability \( 1 - P_d \) against the false alarm probability \( P_a \) for various signal strengths. The signal-to-noise ratio for a biaxial rotor scales approximately in proportion to \( h_0 \), given by (33),

\[
h_0 = h_0 \ 2^{-1/2} \left\{ \left[ \left( 1 + \cos^2 \psi \right)/2 \right]^2 + \cos^2 \psi \right\}^{1/2},
\]

so we quote \( h_0 \) instead of \( h_0 \) as the signal strength. The simulations are conducted in an artificially restricted, 1-Hz sub-band, at a fixed sky location, with both polarization angle \( \psi \) and initial phase \( \Phi_0 \) randomly chosen with a uniform distribution within the range [0, 2\pi] rad.

The ROC curves are essentially indistinguishable for the two scenarios in Section IV B and IV C, because HMM tracking is insensitive to the exact choice of \( A_{q,\psi} \). Figure 5 shows the ROC curves for these two scenarios with four values of \( h_0 \), ranging from 1.8 \times 10^{-26} to 2.5 \times 10^{-26}. For \( P_a = 1\% \), we have 85% and 99% confidence to detect a signal with \( h_0 \) ranging from 2.2 \times 10^{-26} to 2.5 \times 10^{-26}, respectively, read off the top two curves in Figure 5. The 95% confidence sensitivity on effective strain is \( h_0 \) ≈ 2 \times 10^{-26} \( T_{\text{obs}} = 83.3 \text{ d} \).

Figure 6 shows the ROC curves for the searches in Section IV D with four \( h_0 \) values, ranging from 7.5 \times 10^{-26} to 9.0 \times 10^{-26}. The properties of the curves are similar to Figure 5. However, the overall sensitivity degrades by a factor of ≈ 3.5, with \( h_0 \) ≈ 8.5 \times 10^{-26} \( (T_{\text{obs}} = 83.3 \text{ d}) \).

V. HMM TRACKING OF \( f_0 \) AND \( \dot{f}_0 \)

In Sections IV and V we show that one-dimensional HMM tracking can be applied to search for any young objects, but the sensitivity degrades when the spin-down rate is too high, e.g., \( |\dot{f}_0| \gtrsim 10^{-8} \text{ Hz s}^{-1} \) and \( T_{\text{drift}} \lesssim \) a few hours. In this section, we describe a more costly alternative to \( f_0 \) tracking, which allows relatively longer \( T_{\text{drift}} \) when the spin-down rate is high. We formulate the tracker as a two-dimensional HMM with hidden state \( (f_0, \dot{f}_0) \) in Section V A and present simulation examples in Section V B.

A. Transition and emission probabilities

In this implementation, we define a two-dimensional hidden state variable \( q(t) = (f_0(t), \dot{f}_0(t)) \) and track \( f_0 \) and \( \dot{f}_0 \) jointly. The state variable can take \( N_Q = N_f_0 \times N_f_0 \) possible discrete values \( q_{ij} \in \{q_1, \cdots, q_{N_f_0}N_f_0 \} \), where
FIG. 4. Injected $f_0(t)$ (blue curve) and optimal Viterbi path (red curve) for the injected signals in Table VII. Panels (a)–(d) display paths for $h_0/10^{-26} = 15, 13, 11, 10$, respectively. Good matches are obtained in (a)–(c), with $\varepsilon f_0 < 2\Delta f_0$. In panel (d), the signal is not detected. The horizontal axes are in units of HMM steps with $T_{\text{drift}} = 1$ hr for each step ($N_T = 2000$).

$i$ and $j$ index $f_0$ and $\dot{f}_0$ bins, respectively, and $N_{f_0}$ and $N_{\dot{f}_0}$ are the total number of $f_0$ and $\dot{f}_0$ bins, respectively. The discrete hidden states are mapped one-to-one to the two-dimensional array of bins in the output of the estimator $F(f_0, \dot{f}_0)$ computed over $T_{\text{drift}}$. The $f_0$ and $\dot{f}_0$ bin sizes $\Delta f_0$ and $\Delta \dot{f}_0$ are selected using a phase metric described in Appendix B. Assuming that the spin-down evolution of a neutron star is smooth (i.e. no glitches) and that $f_0(t)$ is bounded, we can always choose an intermediate time-scale $T_{\text{drift}}$ for a particular astrophysical source, $T_{\text{SFT}} < T_{\text{drift}} < T_{\text{obs}}$, to satisfy

$$\left| \int_t^{t+T_{\text{drift}}} dt' \dot{f}_0(t') \right| < \Delta \dot{f}_0$$

for $0 < t < T_{\text{obs}}$. We calculate $f_0(t_{n+1})$ from the estimated $f_0(t_n)$ and $\dot{f}_0(t_n)$ according to

$$f_0(t_{n+1}) = f_0(t_n) + \dot{f}_0(t_n)T_{\text{drift}}.$$  

Alternatively, if we track $f_0(t)$ and $\dot{f}_0(t)$ independently, another constraint on $T_{\text{drift}}$ is imposed by $\Delta f_0$, given by (21). In other words, we cannot use longer $T_{\text{drift}}$ than that in the $\dot{f}_0$ tracking. Hence we do not track $f_0$ and $\dot{f}_0$ independently and choose $T_{\text{drift}}$ to satisfy (34) only.

---

5. The $F$-statistic is computed as a function of $f_0(t)$ and $\dot{f}_0(t)$ at a given reference time. We normally choose the start time $t_n$ of the interval as the reference time.
FIG. 5. Receiver operator characteristic (ROC) curves for the searches described in Section IVB and IV C (curves indistinguishable to the eye). The four curves (from top to bottom) correspond to the four representative wave strains h_0^eff/10^{-26} = 2.5, 2.2, 2.0, and 1.8. The horizontal and vertical axes indicate the false alarm probability P_a and detection probability 1 - P_d, respectively. Each curve is based on 500 realizations. Parameters: T_{\text{drift}} = 50 \text{ hr}, N_T = 40, S_h(\hat{f})^{1/2} = 4 \times 10^{-24} \text{ Hz}^{-1/2}.

FIG. 6. Receiver operator characteristic (ROC) curves for the searches described in Section IV D. The four curves (from top to bottom) correspond to the four representative wave strains h_0^eff/10^{-26} = 9.0, 8.5, 8.0, and 7.5. The horizontal and vertical axes indicate the false alarm probability P_a and detection probability 1 - P_d, respectively. Each curve is based on 200 realizations. Parameters: T_{\text{drift}} = 1 \text{ hr}, N_T = 2000, S_h(\hat{f})^{1/2} = 4 \times 10^{-24} \text{ Hz}^{-1/2}.

If we update f_0(t_{n+1}) according to (35), the transition probability matrix becomes

\[ A_{q_{i} - \Delta i, j + 1 | q_{ij}} = A_{q_{i} - \Delta i, j | q_{ij}} = \{ 2[\Delta t_{\text{max}}(j) - \Delta t_{\text{min}}(j) + 1] \}^{-1}, \tag{36} \]

with all other terms being zero. In (36), \Delta i takes integer values \Delta t_{\text{min}}(j) \leq \Delta i \leq \Delta t_{\text{max}}(j) with

\[ \Delta t_{\text{min}}(j) = \text{floor}(\hat{f}_{0,j} T_{\text{drift}} / \Delta f_0), \tag{37} \]
\[ \Delta t_{\text{max}}(j) = \text{ceil}(\hat{f}_{0,j} T_{\text{drift}} / \Delta f_0), \tag{38} \]

where floor(x) denotes the largest integer smaller than or equal to x, ceil(x) denotes the smallest integer larger than or equal to x, and \hat{f}_{0,j} is the value of \hat{f}_0 in the j-th \hat{f}_0 bin. The detailed derivation of (36) is given in Appendix D.

The emission probability is given by

\[ L_\omega(t | q_{ij}) = P[o(t) | f_0(t) \leq f_0(t) \leq f_0 + \Delta f_0, \hat{f}_{0,j} \leq \hat{f}_0(t) \leq \hat{f}_{0,j} + \Delta \hat{f}_0] \propto \exp[F(f_{0,i}, \hat{f}_{0,j})]. \tag{39} \]

We choose a uniform prior in both f_0 and \hat{f}_0, viz.

\[ \Pi_{q_{ij}} = N_Q^{-1}. \tag{41} \]

B. Abridged mock search

In this section we demonstrate the (f_0, \hat{f}_0) HMM tracker using synthetic data. To make a fair comparison with the f_0 tracker, we conduct an abridged version of a mock search for the rapidly spinning down signal simulated in Section IV D, with the same injection parameters as in Tables I and VI. We choose T_{\text{drift}} = 50 \text{ hr} (N_T = 40) to satisfy (34) and use the search parameters in Table VIII. The F-statistic is computed over a 1-Hz frequency band as a function of f_0 and \hat{f}_0 for each segment. For demonstration purposes, only five values of \hat{f}_0 are searched in a range containing the injected \hat{f}_0 to save time, i.e., the phase metric is not computed.

The results are presented in Table IX. Compared to the performance displayed in Table VII using f_0 tracking, the (f_0, \hat{f}_0) tracking can detect a signal about three times weaker. We calculate the RMSE \varepsilon_{f_0} in f_0 between the optimal Viterbi path and the injected signal (in Hz and in units of \Delta f_0). We do the same for the RMSE \varepsilon_{\hat{f}_0} in \hat{f}_0 (in Hz s^{-1} and in units of \Delta \hat{f}_0). In a real search, we consider candidates for follow-up and further scrutiny, if S exceeds a threshold S_h set by the desired false alarm and false dismissal probabilities, as shown in Section IV. The value of S_h depends strongly on N_Q and hence the two-dimensional (f_0, \hat{f}_0) parameter space. Systematic Monte-Carlo simulations are required in practice to calculate S_h for each HMM implementation, an exercise lying outside the scope of this paper. Instead, in this section, we adopt the following rule of
thumb: the injected signal is deemed to be detected if we obtain $\varepsilon_{f_0} < 0.5\Delta f_0$ and $\varepsilon_{f_0} < 10\Delta f_0$. The errors are introduced mostly because HMM takes discrete values of $f_0$ and $f_0$, while the injected signal $f_0(t)$ and $f_0(t)$ can be any value within a bin. Since we calculate $f_0(t_{n+1})$ from the estimated $f_0(t_n)$ and $f_0(t_n)$, $\varepsilon_{f_0}$ accumulates to a few $\Delta f_0$ after $N_T$ steps, introduced by $\varepsilon_{f_0}$.

Figure 7 displays the optimal Viterbi paths (red curves) and the true paths $f_0(t)$ and $f_0(t)$ (blue curves) for the two weakest injections (a) $h_0 = 4 \times 10^{-26}$ and (b) $h_0 = 3 \times 10^{-26}$. The left and right panels show $f_0$ and $f_0$, respectively. In Figure 7(a) the optimal Viterbi paths agree with $f_0(t)$ and $f_0(t)$ closely. In the right panel, it is shown that the estimated $f_0$ fluctuates within one bin around the injected $f_0(t)$. The fluctuations around $f_0$ cannot be seen clearly in the left panel, because $\varepsilon_{f_0} = 1.80 \times 10^{-5}$ Hz is much smaller than the total change of $f_0$ over $T_{\text{obs}} \approx 0.2$ Hz. In contrast, Figure 7(b) shows that the optimal Viterbi paths do not match the injected $f_0(t)$ and $f_0(t)$, i.e., the injected signal is not detected.

### VI. DISCUSSION

#### A. Cost-sensitivity trade-off

In this section, we start by comparing the HMM tracking method to existing stack-slide-based semi-coherent methods and then discuss the cost-sensitivity trade-off between $(f_0, \dot{f}_0)$ tracking and $f_0$ tracking. Analytic approximations for the computing cost and sensitivity are described briefly in Appendices B and E.

HMM tracking incoherently combines the $F$-statistic outputs from $N_T = T_{\text{obs}}/T_{\text{drift}}$ blocks of data. The computing cost is composed of two parts: (1) calculating the coherent $F$-statistic (i.e., $L_{\text{coh}}(\tau)$) for all $N_T$ segments; and (2) recursively maximizing $P(Q|O)$, i.e., solving the HMM. Assuming we use data from two interferometers and search up to the maximum frequency $f_{0,\text{max}}$, the computing costs of calculating $F(f_0)$ and $F(f_0, \dot{f}_0)$ over one block of coherent segment $T_{\text{drift}}$ are given by

$$T_{f_0} = 0.46 \left( \frac{f_{0,\text{max}}}{0.6 \text{kHz}} \right) \left( \frac{T_{\text{drift}}}{10 \text{d}} \right)^2 \left( \frac{1}{N_{\text{core}}} \right),$$

and

$$T_{f_0, \dot{f}_0} = 0.36 \left( \frac{f_{0,\text{max}}}{0.6 \text{kHz}} \right)^2 \left( \frac{0.3 \text{kyr}}{\tau} \right) \left( \frac{T_{\text{drift}}}{10 \text{d}} \right)^4 \left( \frac{10^3}{N_{\text{core}}} \right),$$

respectively, where $N_{\text{core}}$ is the number of cores running in parallel (see details in Appendix B). The Viterbi algorithm computes $Q^*(O)$ via $(N_T + 1)/N_Q$ in $N_Q$ operations. For example, in a 1-Hz sub-band with $N_Q = 2 \times 10^6$ and $N_T = 36$, it takes $\lesssim 30$ s to compute $Q^*(O)$ but $\lesssim 1$ hr to compute $N_T$ blocks of the $F$-statistic. Hence the total computing cost is dominated by the cost of computing $N_T$ blocks of the $F$-statistic, scaling as $T_{\text{f0}} \propto N_T f_{0,\text{max}}^2 T_{\text{drift}}$ for $f_0$ tracking, and $T_{f_0, \dot{f}_0} \propto N_T f_{0,\text{max}}^2 \tau^{-1} T_{\text{drift}}^4$ for $(f_0, \dot{f}_0)$ tracking.

When higher-order derivatives of the frequency are required to be searched for very young objects (e.g., $\tau \lesssim 0.03$ kyr), e.g., in a stack-slide search, the cost of computing $F$-statistic grows geometrically as $T_{\text{drift}}$ increases. Figure 8 shows the cost of calculating the $F$-statistic over one block of duration $T_{\text{drift}}$ for a target with $\tau = 0.03$ kyr up to $f_{0,\text{max}} = 600$ Hz. The three curves, from bottom to top, represent calculating $F(f_0)$, $F(f_0, \dot{f}_0)$, and $F(f_0, f_0, \ddot{f}_0)$, respectively. For example, it requires $\sim 10^3$ core-day to compute $F(f_0, f_0, \dot{f}_0)$ for a single block of duration $T_{\text{drift}} = 4$ d. When $f_0$ needs to be considered, the cost of calculating $F(f_0, f_0, \dot{f}_0, \ddot{f}_0)$ becomes prohibitive even for $T_{\text{drift}} < 1$ d. Under these circumstances, the HMM tracker comes into its own; it allows an efficient search for rapidly evolving signals without searching high-order frequency derivatives.

Given fixed $T_{\text{obs}}$, one can tune $T_{\text{drift}}$ to trade off sensitivity against computing cost for a particular target. Table X shows the theoretical scalings of sensitivity and cost as a function of $T_{\text{drift}}$ for the two HMM implementations described in Section III and IV. In reality, the scalings vary with many factors, including $N_T$, $P_a$, $P_b$, and the noise statistics, as discussed in detail in Refs. [68, 69]. In this paper, we include the theoretical scalings to allow quick order-of-magnitude comparisons, but we emphasize that they are not a substitute for Monte-Carlo simulations. The sensitivities of $f_0$ tracking and $(f_0, \dot{f}_0)$ tracking scale the same way with $T_{\text{drift}}$. An $(f_0, \dot{f}_0)$ search

| Parameter | Value | Unit |
|-----------|-------|------|
| $f_0$     | 151–152 | Hz   |
| $-\dot{f}_0$ | $2.98 \times 10^{-8}$–$3.02 \times 10^{-8}$ | Hz/s⁻¹ |
| $T_{\text{drift}}$ | 50 | hr |
| $\Delta f_0$ | $2.78 \times 10^{-6}$ | Hz |
| $\Delta \dot{f}_0$ | $1 \times 10^{-10}$ | Hz/s⁻¹ |
| $T_{\text{obs}}$ | 83.3 | d |
| $N_T$ | 40 | – |

**TABLE VIII.** Search parameters for the synthetic signals with injection parameters quoted in Tables II and V.
TABLE IX. Results of \((f_0, f_0')\) tracking for injected signals with the parameters in Tables I and V, \(T_{\text{obs}} = 83.3\) d, \(T_{\text{drift}} = 50\) hr, and wave strain \(h_0\). The RMSE \(\varepsilon_{f_0}\) between the frequency of the optimal Viterbi path and the injected \(f_0(t)\) is quoted in Hz and in units of \(\Delta f_0 = 2.78 \times 10^{-6}\) Hz. The RMSE \(\varepsilon_{f_0'}\) between the frequency derivative of the optimal Viterbi path and the injected \(f_0'(t)\) is quoted in Hz s\(^{-1}\) and in units of \(\Delta f_0' = 1 \times 10^{-10}\) Hz s\(^{-1}\). The third column quotes the Viterbi score \(S\).

| \(h_0\) \((10^{-26})\) | Detect? | \(S\) | \(\varepsilon_{f_0}\) \((\text{Hz})\) | \(\varepsilon_{f_0}/\Delta f_0\) | \(\varepsilon_{f_0'}\) \((\text{Hz s}^{-1})\) | \(\varepsilon_{f_0'}/\Delta f_0'\) |
|-----------------|---------|------|-----------------|-----------------|-----------------|-----------------|
| 8.0             | ✓       | 9.0  | \(2.89 \times 10^{-6}\) | 1.0             | \(2.97 \times 10^{-11}\) | 0.30            |
| 5.0             | ✓       | 2.9  | \(6.45 \times 10^{-6}\) | 2.3             | \(3.56 \times 10^{-11}\) | 0.36            |
| 4.0             | ✓       | 2.0  | \(1.80 \times 10^{-5}\) | 6.5             | \(2.85 \times 10^{-11}\) | 0.29            |
| 3.0             | ✓       | 1.2  | 0.45            | \(1.6 \times 10^{5}\) | \(1.659 \times 10^{-10}\) | 1.59            |

The table entries are as follows:

- **Detect?** column indicates whether the signal was detected. ✓ means detected, and × means not detected.
- **\(S\)** column represents the Viterbi score.
- **\(\varepsilon_{f_0}\)** column is the RMSE between the frequency of the optimal Viterbi path and the injected \(f_0(t)\).
- **\(\varepsilon_{f_0}/\Delta f_0\)** column is the ratio of the RMSE to the frequency drift rate.
- **\(\varepsilon_{f_0'}\)** column is the RMSE between the frequency derivative of the optimal Viterbi path and the injected \(f_0'(t)\).
- **\(\varepsilon_{f_0'}/\Delta f_0'\)** column is the ratio of the RMSE to the frequency derivative of the frequency drift rate.

**FIG. 7.** Injected \(f_0(t)\) and \(f_0'(t)\) (blue curves) and optimal Viterbi paths (red curves) for the last two injected signals in Table IX with (a) \(h_0 = 4 \times 10^{-26}\) and (b) \(h_0 = 3 \times 10^{-26}\). The top left and top right panels show good matches for both \(f_0\) and \(f_0'\), respectively. The red curve in the left panel fluctuates around the blue curve with \(\varepsilon_{f_0} = 1.80 \times 10^{-5}\) Hz, which is too small to be seen in the plot \((\varepsilon_{f_0} \ll \dot{f}_0 T_{\text{obs}} \approx 0.2\) Hz\). In the lower two panels, the injected signal is not detected. The horizontal axes are in units of HMM steps with \(T_{\text{drift}} = 50\) hr for each step \((N_T = 40)\).
allows longer \( T_{\text{drift}} \) and hence in practice is always more sensitive than an \( f_0 \) search. However, an \( f_0 \) search is always faster.

Figure 8 plots the ratio \( T_{f_0,f_0}/T_{f_0} \) as a function of \( \tau \). The curves, from bottom to top, represent achieving 10%–50% better sensitivity by switching from \( f_0 \) tracking to \((f_0,\dot{f}_0)\) tracking. We can always achieve better sensitivity using \((f_0,\dot{f}_0)\) tracking compared to \( f_0 \) tracking. However, \( T_{f_0,f_0}/T_{f_0} \) is approximately proportional to \( \tau \) and increases exponentially with the percentage of sensitivity improvement.

### B. Spin down of young objects with age \( \lesssim f_{0\text{birth}}/f_{0\text{birth}} \)

The true age of a young neutron star may be significantly less than its characteristic spin-down time-scale at birth, \( f_{0\text{birth}}/[(n-1)f_{0\text{birth}}] \), depending on its ellipticity and magnetization. To investigate this scenario, we approximate the braking law with a power law in the usual way, viz. \( f_0(t) = -\eta f_0(t)^n \), with \( \eta \propto B_0^2 \) and \( 2 \lesssim n < 3 \) if the torque is electromagnetically dominated, and \( \eta \propto \epsilon^2 \) and \( n = 5 \) if the torque is dominated by gravitational radiation reaction, where \( \epsilon \) is the ellipticity. Integrating the braking law with respect to \( t \), we find that the characteristic spin-down time-scale of the signal is given by \[ f_0/f_0 = |\xi|^{1-1}\tau \] (44) with

\[ \xi = \frac{1}{n-1} \left[ 1 - \left( \frac{f_{0\text{birth}}}{f_0} \right)^{1-n} \right]. \] (45)

The term \((f_{0\text{birth}}/f_0)^{1-n}\) is normally neglected under the assumption \( f_0 \ll f_{0\text{birth}} \), yielding \( \dot{f}_0 \approx -f_0(n-1)^{-1}\tau^{-1} \). However, this assumption does not necessarily apply to young objects (e.g. \( \tau = 0.03 \text{kyr} \) for SNR 1987A), for which we obtain \( |\xi| \lesssim 0.05 \) for \( f_{0\text{birth}} \lesssim 600 \text{Hz} \) and \( B_0 \lesssim 6 \times 10^{12} \text{G} \) with \( n = 3 \). A detailed discussion can be found in Section IIB of Ref. 33.

The indirect upper limit on \( h_0 \) derived from energy conservation is given by \[ h_0 \leq \frac{1}{D} \left( \frac{5GI |\xi|}{2c^3 \tau} \right)^{1/2}, \] (46)
where \( G \) is Newton’s gravitational constant, \( c \) is the speed of light, and \( D \) is the distance to the source. The indirect limit on \( h_0 \) is lowered because of the second term in (45). On the other hand, the slower spin-down rate \( \dot{f}_0 = -f_0|\xi|\tau^{-1} \) benefits HMM tracking by allowing longer \( T_{\text{drift}} \). If we consider a young object with \( \tau = 0.03 \) kyr as an example, the impact of having \( |\xi| = 0.05 \) translates into raising \( T_{\text{drift}} \) by a factor of \( \approx 3 \).

VII. CONCLUSION

In this paper, we describe two practical implementations of an efficient HMM tracker, combined with the maximum likelihood matched filter \( F \)-statistic, to economically search for continuous gravitational waves from young neutron stars in SNRs. The HMM incoherently combines the coherent \( F \)-statistic outputs from multiple \((N_T)\) data blocks of duration \( T_{\text{drift}} \). It tracks rapid, secular spin down without searching high-order derivatives of the signal frequency. The first implementation, tracking \( f_0 \) alone, can simultaneously surmount two challenges in young SNR searches: rapid spin down and stochastic timing noise. Three scenarios for different spin-down and timing-noise time-scales are discussed. Given \( T_{\text{obs}} = 83.3 \text{ d} \), we obtain \( h_0^{\text{eff}, 95\%} \approx 2.4 \times 10^{-26} \) for both weak and strong timing noise in the first two scenarios \((\tau \gtrsim 5 \text{ kyr})\) and \( h_0^{\text{eff}, 95\%} \approx 8.5 \times 10^{-26} \) in the last scenario \((\tau \lesssim 0.03 \text{ kyr})\), with \( P_a = 1\% \). We expect that \( h_0^{95\%} \) is more conservative than the quoted \( h_0^{\text{eff}, 95\%} \) for unknown \( \cos \iota \).

The second implementation, tracking \( f_0 \) and \( \dot{f}_0 \), allows longer \( T_{\text{drift}} \) and hence improves the sensitivity by a factor of a few. The first implementation is always faster and more robust against timing noise. One can achieve better sensitivity by switching from the first implementation to the second. However, it increases the computing cost by two to three orders of magnitude, depending on \( \tau \).

An optimized \( F \)-statistic-based semi-coherent Einstein@Home search for Cas A \((f_{\text{max}} = 1.5 \text{ kHz})\) in the Advanced LIGO O1 run costs approximately \( 2.7 \times 10^5 \text{ core-day} \), yielding \( 90\% \) confidence strain upper limit \( 1.4 \times 10^{-25} \) \([71]\). Assuming the same parameters, the method discussed in this paper is expected to provide comparable sensitivity but cost \( \sim 10^4 \text{ core-day} \). The advantage of HMM tracking grows in searches for younger targets, e.g., SNR 1987A.

The methods described in this paper can be applied to extending the searches for the SNRs listed in Ref. \([38]\), which are restricted to coherent segments of duration \( T_{\text{drift}} \leq 25.3 \text{ d} \), using the new data from a whole Advanced LIGO observing run. In addition, the recent work by Anderson et al. \([72]\) has identified 76 new Galactic SNR candidates, some of which may be promising candidates for gravitational-wave sources, if the SNR associates are confirmed. The \( f_0 \) tracker can be applied to search for targets that are poorly modelled, e.g., a long transient post-merger signal from the binary neutron star merger GW170817 \([73]\) with spin-down time-scale \( \sim 10^2–10^4 \text{ s} \). Some modifications are needed, e.g. \( L_{o,q} \) should be calculated from the power in SFT bins rather than the \( F \)-statistic, because the Earth’s rotation can be neglected.

To carry out a search using the methods presented in this paper, the following steps need to be completed in preparation. First, the search parameter ranges need to be determined systematically. The \( f_0 \) range is normally chosen to equal the band where the estimated strain sensitivity is below the indirect, \( \tau \)-based limit \([36]\). General equations \((14)\) and \((15)\) for calculating \( f_0 \) and \( \dot{f}_0 \) are given in Section II. Second, search parameter resolutions need to be calculated using the metric described in Appendix B given a desired mismatch. Third, a systematic Monte-Carlo simulation is required for each implementation to determine the detection threshold \( S_{\text{th}} \) given false alarm and false dismissal probabilities.

VIII. ACKNOWLEDGEMENTS

We would like to thank Ra Inta and the LIGO Scientific Collaboration Continuous Wave Working Group for detailed comments and informative discussions. L. Sun is supported by Australia Research Training Program Stipend Scholarship. The research was supported by Australian Research Council (ARC) Discovery Project DP110103347 and the ARC Centre of Excellence for Gravitational Wave Discovery CE170100004.

Appendix A: Viterbi algorithm

The principle of optimality \([74]\) demonstrates that in our special case, all subpaths \( Q^*(k) \) made up of the first \( k \) steps in \( Q^*(O) \) are optimal for \( 1 \leq k \leq N_T \). In that sense, the classic Viterbi algorithm \([43]\) provides a recursive, computationally efficient solution to computing \( Q^*(O) \) in a HMM, reducing the number of operations from \( N_Q^N_T \) to \((N_T + 1)N_Q \) by binary maximization \([42]\). A full description of the algorithm can be found in Section II D of Ref. \([48]\). At every forward step \( k \) \((1 \leq k \leq N_T) \) in the recursion, the algorithm eliminates all but \( N_Q \) possible state sequences, and stores the \( N_Q \) maximum probabilities

\[
\delta_{q_i}(t_k) = L_{o(t_k)|q_i} \max_{1 \leq j \leq N_Q} [A_{q_i,q_j}\delta_{q_j}(t_{k-1})], \quad (A1)
\]

and previous-step states leading to the retained most likely sequence

\[
\Phi_{q_i}(t_k) = \arg\max_{1 \leq j \leq N_Q} [A_{q_i,q_j}\delta_{q_j}(t_{k-1})]. \quad (A2)
\]

When backtracking, for \( 0 \leq k \leq N_T - 1 \), we reconstruct the optimal Viterbi path according to

\[
q^*(t_k) = \Phi_{q^*(t_{k+1})}(t_{k+1}). \quad (A3)
\]
Appendix B: Phase metric and computing cost

The costs of computing the $F$-statistic (i.e. $L_{\rho(q)}$) and recursively maximizing $P(Q|O)$ depend on the template spacing. We start by discussing the template spacing and cost for a general $F$-statistic search. In order to optimize the template spacing, a phase metric is defined. It expresses the signal-to-noise ratio as a function of template spacing along each parameter axis (e.g., $f_0$, $f_0$, $f_0$, ···). The mismatch $m$ is defined as the fractional reduction of $F$-statistic power caused by discrete parameter sampling, with \[ m = \sum_{i,j} g_{ij} \Delta \lambda^i \Delta \lambda^j, \] \[ g_{ij} = \frac{4\pi^2(i+1)(j+1)T_{\text{drift}}^{i+j+2}}{(i+2)(j+2)(i+j+3)}. \] \[ (B1) \]

The indices $i$ and $j$ take integer values from 0 to $k$, where $k$ indicates the highest-order frequency derivative considered (e.g. $k = 2$ for searching up to $f_0$), and $\Delta \lambda^i$ denotes the offset between the true value and the closest template of the $i$-th parameter. For example, the maximum value of $\Delta \lambda^0$ is half the frequency bin width $\Delta f_0$, because the signal frequency falls halfway between two templates in the worst case. We choose to adopt $m \leq 0.2$ in line with the Cas A search in S5 data [34] and the SNR searches in S6 data [36]. The highest frequency derivative needed is the largest integer $k$ satisfying $g_{kk}[f_{0\text{max}}(k) - f_{0\text{min}}(k)]^2 > m$ (no summation over $k$ implied), where $f_{0\text{max}}(k)$ and $f_{0\text{min}}(k)$ are the maximum and minimum $k$-th frequency derivative. In practice, we can choose the bin size of the $i$-th frequency derivative $\Delta f_0^{(i)}$ using the diagonal terms of (B2) to satisfy

$$\sum_{i=0}^{k} g_{ii}[\Delta f_0^{(i)}]^2 < 4m.$$ \[ (B3) \]

Monte-Carlo simulations are needed to accurately calculate the required bin sizes for a given $m$. Taking into consideration the off-diagonal terms of (B2) yields bin sizes close to the empirical Monte-Carlo results. A tiling algorithm is described in detail in Ref. [38]. Combining [14], [15] and (B2), the number of templates $N$ needed for $k = 2$ is [50, 78]

$$N = 0.35m^{-3/2}(\det g)^{1/2} f_{0\text{max}}^3 \tau^{-3},$$ \[ (B4) \]

with $f_{0\text{min}} \ll f_{0\text{max}}$ typically.

The computing time of a coherent $F$-statistic search over one block of duration $T_{\text{drift}}$ is given by

$$T = \kappa N \beta N_{\text{ifo}} T_{\text{drift}} T_{\text{SFT}}^{-1}.$$ \[ (B5) \]

where $\kappa$ is the time to compute the $F$-statistic per template per SFT. \[ N_{\text{ifo}} \] is the number of interferometers, and $\beta$ is the percentage of time that the interferometers collect data (i.e., duty cycle). For most of the young targets discussed in Ref. [36], $f_0$ is normally small. Only a few $f_0$ values need to be searched. For example, we obtain $f_0 < 1 \times 10^{-18} \text{Hz}^{-2}$ from [14] and [15] for $\tau \gtrsim 1 \text{kyr}$ and $f_0 < 600 \text{Hz}$ and $\Delta f_0 \sim 10^{-18} \text{Hz}^{-2}$ from [12] and [33] with $T_{\text{drift}} = 10 \text{d}$. In this example, only one value of $f_0$ is searched and the cost scaling in Equation (B5) reduces to $f_{0\text{max}}^{-1} T_{\text{drift}}^{-1}$. If we assume that only one value of $f_0$ is searched. For $\kappa = 4 \times 10^{-8} \text{s}, m = 0.2, \beta = 1, N_{\text{ifo}} = 2, T_{\text{SFT}} = 1800 \text{s}$ and $N_{\text{core}} = 10^3$ cores running in parallel, we obtain

$$T_{f_0,f_0} = 0.36 \frac{d}{0.6 \text{kHz}} \left(\frac{f_{0\text{max}}}{0.3 \text{kyr}}\right)^2 \left(\frac{T_{\text{drift}}}{10 \text{d}}\right)^4 \left(\frac{10^3}{N_{\text{core}}}\right).$$ \[ (B6) \]

However, for very young objects (e.g., $\tau \lesssim 0.03 \text{kyr}$) with larger $f_0$, the cost scales as $T_{f_0,f_0} \propto f_{0\text{max}}^{-1}$.

Figure 10 shows the cost of computing the $F$-statistic over a coherent segment $T_{\text{drift}}$ (in units of core-day). For concreteness, we fix $f_{0\text{max}} = 600 \text{Hz}$. In a real search, $f_{0\text{max}}$ is a function of $\tau$, because we determine $f_{0\text{max}}$ to be the maximum frequency where the estimated strain sensitivity of the search beats the indirect spin-down limit [see Equation (10)]. If we compute $F(f_0, f_0)$ (or search a single $f_0$ value), the costs for objects with $\tau = 0.3 \text{kyr}$ and 1 kyr are indicated by the two solid curves. A coherent $F$-statistic search or a stack-slide-based semi-coherent $F$-statistic search requires searching higher-order derivatives for objects with $\tau \lesssim 0.3 \text{kyr}$. The two dashed curves (top and bottom) represent the cost of computing $F(f_0, f_0, f_0)$ for objects with $\tau = 0.03 \text{kyr}$ and 0.1 kyr, respectively.

The serial clock time for computing can be reduced by parallelization. For $10^3$ nodes running in parallel, a coherent $F$-statistic search over $T_{\text{drift}} = 10 \text{d}$ takes about 9 hr for an object with $\tau = 0.3 \text{kyr}$ (e.g., Cas A), and about 10 d for an object with $\tau = 0.1 \text{kyr}$ (e.g., G1.9+0.3). In reality, the cost indicated by the top dashed curve for an object with $\tau = 0.03 \text{kyr}$ (e.g., SNR 1987A) is still underestimated, because $f_0$ and higher-order derivatives must be searched using a stack-slide-based semi-coherent method.

We do not search second- or higher-order derivatives of the frequency in the HMM tracking. We start by discussing the $(f_0, f_0)$ tracking. In the two-dimensional tracking, we can always substitute the maximum $f_0$ in

\[ \tau \gtrsim \frac{1}{4} \text{kyr} \]
the range into \((B4)\) to choose \(T_{\text{drift}}\) and ignore \(\tilde{f}_0\) (or search a single value). However, for very young objects (e.g., \(\tau \lesssim 0.03 \text{ kyr}\)) with larger \(\tilde{f}_0\), there can be sensitivity loss due to the short \(T_{\text{drift}}\) derived from \((B4)\). The relation between theoretical sensitivity and \(T_{\text{drift}}\) is discussed in Appendix E. The estimate of \(T\) in \((B6)\) stands for the time required to calculate the \(F\)-statistic over one block of coherent segment \(T_{\text{drift}}\). For HMM tracking, we need to add up the time required to calculate all the required values of \(F(f_0, \tilde{f}_0)\). In addition there is a second component to the computing cost, namely solving the HMM. HMM tracking incoherently combines the \(F\)-statistic outputs from \(N_T = T_{\text{obs}}/T_{\text{drift}}\) blocks of data. The Viterbi algorithm computes \(Q'(O)\) via \((N_T + 1)N_Q \ln N_Q\) operations [48].

The total computing cost is dominated by the cost of computing \(N_T\) blocks of the \(F\)-statistic, scaling as \(T \propto N_T \eta f_0 t_{\text{drift}}^{-1} T_{\text{drift}}\). If we take \(T_{\text{drift}} = 5 \text{ d}\) as an example, Figure 11 shows the computing cost of a semi-coherent HMM search (in units of core-day) as a function of \(T_{\text{obs}}\). When tracking \(f_0\) alone, \(N_T \eta f_0 t_{\text{drift}}^{-1} T_{\text{drift}}\) when tracking \(f_0\) alone, saving a factor \(\sim f_0 t_{\text{drift}}^{-1} T_{\text{drift}}\) compared to the \((f_0, \tilde{f}_0)\) tracking.

**Appendix C: \(T_{\text{drift}}\) given \(\tau\)**

We assume purely electromagnetic spin down \((\dot{f}_0 \propto B_0^2 f_0^3, n = 3)\) for simplicity, which gives

\[
\dot{f}_0(t) = -\eta f_0(t)^3, \tag{C1}
\]

where the coefficient \(\eta \propto B_0^2\) is a positive constant. At time \(t = 0\) when the star was born, we have \(f_0(t) = f_{0\text{birth}}\). The differential equation \((C1)\) has the solution

\[
1 - \left(\frac{f_0}{f_{0\text{birth}}}\right)^2 = -\frac{2\tau \dot{f}_0}{f_0}. \tag{C2}
\]
Substituting $\dot{f}_0 = -(2T_{\text{drift}}^2)^{-1}$ from [26] into (C2), we obtain

$$T_{\text{drift}} = \tau^{1/2} f_0^{-1/2} \left[ 1 - \left( \frac{f_0}{f_0^{\text{birth}}} \right)^2 \right]^{-1/2}. \quad (C3)$$

**Appendix D: Transition probability matrix for $f_0$, $\dot{f}_0$ tracking**

We first derive the transition probabilities corresponding to the substate $\dot{f}_0$. Using Equations (14) and (15), the range of $\dot{f}_0$ is given by

$$\frac{0.2f_{0\text{min}}}{\tau^2} \leq \dot{f}_0 \leq \frac{2f_{0\text{max}}}{\tau^2}, \quad (D1)$$

where $f_{0\text{min}}$ and $f_{0\text{max}}$ are the minimum and maximum $f_0$ being searched. The maximum $\dot{f}_0$ is more than two orders of magnitude larger than the minimum $\dot{f}_0$. We assume that $\dot{f}_0$ is uniformly distributed in the range $0 \leq \dot{f}_0 \leq f_{0\text{max}}$. At each step, given Equation (34), $\dot{f}_0$ jumps at most one bin up or stays in the same bin with equal probability 1/2.

We then estimate the number of bins $f_0$ moves during each step. Equation (35) is more precisely given by

$$f_0(t_n) + \dot{f}_0(t_n) T_{\text{drift}} \leq f_0(t_{n+1}) \leq f_0(t_n) + [\dot{f}_0(t_n) + \Delta \dot{f}_0] T_{\text{drift}}. \quad (D2)$$

Let us write $q(t_n) = [f_0(t_n), \dot{f}_0(t_n)] = q_{ij}$, where $i$ and $j$ index $f_0$ and $\dot{f}_0$ bins, respectively. Then the number of bins that $f_0$ moves from step $t_n$ to step $t_{n+1}$, denoted by $\Delta i$, takes the minimum and maximum values

$$\Delta i_{\text{min}}(j) = \text{floor}(\dot{f}_0 j + 1 | T_{\text{drift}}/\Delta \dot{f}_0), \quad (D3)$$

$$\Delta i_{\text{max}}(j) = \text{ceil}(\dot{f}_0 j + 1 | T_{\text{drift}}/\Delta \dot{f}_0), \quad (D4)$$

where floor($x$) denotes the largest integer smaller than or equal to $x$, ceil($x$) denotes the smallest integer larger than or equal to $x$, and $\dot{f}_0 j$ is the value of $\dot{f}_0$ in the $j$-th $\dot{f}_0$ bin. In other words, $f_0(t_{n+1})$ can be located in any bin within the range $[i - \Delta i_{\text{max}}(j), i - \Delta i_{\text{min}}(j)]$ with uniform probability. The two dimensional transition probability matrix is given by

$$A_{qi−\Delta i,j+1q_{ij}} = A_{qi−\Delta i,jq_{ij}} \quad (D5)$$

$$= \left\{2[\Delta i_{\text{max}}(j) - \Delta i_{\text{min}}(j)] + 1\right\}^{-1}, \quad (D6)$$

where $\Delta i$ takes integer values $\Delta i_{\text{min}}(j) \leq i \leq \Delta i_{\text{max}}(j)$, and all other terms are zero.

---

Appendix E: Analytic sensitivity scalings

In this section, we present an approximate analytic formula for the search sensitivity, based on a few general assumptions. Deviations are discussed in detail in Ref. [68, 69]. Accurate sensitivity scalings require Monte-Carlo simulations for each implementation of the search, as shown in Section IV.

The sensitivity of a search can be defined in terms of the characteristic gravitational-wave strain corresponding to 95% detection efficiency. For a coherent $\mathcal{F}$-statistic search over one block of $T_{\text{drift}}$, searching up to the highest frequency diversity required for a given mismatch, it takes the form [30, 50]

$$h_{0}^{95\%}(f) = \Theta S_h(f)^{1/2}(\beta T_{\text{drift}})^{-1/2}, \quad (E1)$$

where $\Theta$ is a statistical threshold, depending on the shape of the parameter space manifold. One finds $30 \leq \Theta \leq 40$ for a directed search of the type discussed in this paper [50]. The term $\beta T_{\text{drift}}$ gives the length of the interferometer data in the timespan $T_{\text{drift}}$.

As every block of $\mathcal{F}$-statistic output over $T_{\text{drift}}$ is chi-squared distributed with four degrees of freedom and the chi-squared distribution is additive, we can calculate the PDF of $z = \ln \mathcal{P}(Q|O)$ along the true signal path from (11) and (13) by multiplying both the degrees of freedom and the noncentrality parameter by $N_T = T_{\text{obs}}/T_{\text{drift}}$. If $Q^*(O)$ coincides exactly with the true path, we obtain

$$p(z) = \chi^2 \left( z; \frac{4T_{\text{obs}}}{T_{\text{drift}}}, \frac{Kh_0^2T_{\text{obs}}}{S_h(f)} \right), \quad (E2)$$

If $Q^*(O)$ does not intersect the true path anywhere, we have

$$p(z) = \chi^2 \left( z; \frac{4T_{\text{obs}}}{T_{\text{drift}}}, 0 \right), \quad (E3)$$

Combining (E2) and (E3), the signal-to-noise ratio after $N_T$ steps of the HMM equals $\rho_0^2$, given by

$$\rho_0^2 = \frac{\mu_{\text{signal}} - \mu_{\text{noise}}}{\sigma_{\text{noise}}} \propto \frac{h_0^2}{S_h(f)} (T_{\text{obs}}T_{\text{drift}})^{1/2}, \quad (E4)$$

where $\mu_{\text{signal}} = N_T \rho_0^2$ and $\mu_{\text{noise}} = 0$ are the noncentralities of the distributions in (E2) and (E3), respectively, and $\sigma_{\text{noise}} = (8N_T)^{1/2}$ is the standard deviation of the distribution in (E3). Hence we obtain

$$h_{0}^{95\%}(f) = \Theta S_h(f)^{1/2}(T_{\text{obs}}T_{\text{drift}})^{-1/4}, \quad (E6)$$

---

9 Since $f_0$ is negative, we always have $f_0(t_{n+1}) \leq f_0(t_n)$.

10 Here we assume that the $\mathcal{F}$-statistic is independently and identically distributed. The estimate requires modification when applied to real interferometer data, where the noise is non-stationary and/or non-Gaussian. A more robust Bayesian framework is introduced in Ref. [79] to analyze the $\mathcal{F}$-statistic in the presence of instrumental artifacts.
assuming $\beta = 1$. When $T_{\text{obs}} = T_{\text{drift}}$, equation (E6) reduces to (E1).

By way of illustration, we compute $h_{05\%}^{\text{obs}}$ for HMM tracking with $T_{\text{obs}} = N_T T_{\text{drift}} = 360\text{ d}$ and compare the result with a coherent $F$-statistic search limited to $T_{\text{obs}} = T_{\text{drift}}$ like in Ref. [34]. We take $\Theta = 35$ and let $S_0(f)$ be the Advanced LIGO design noise PSD in (E1). Figure 12 shows the results. The blue solid curve, red solid curve, and blue dashed curve indicate an $F$-statistic search of duration $N_T T_{\text{drift}}$, a HMM search of duration $N_T T_{\text{drift}}$, and a fully coherent $F$-statistic search with $T_{\text{obs}} = 360\text{ d}$, respectively. Figure 12(a) plots $h_{05\%}^{\text{obs}}$ as a function of signal frequency. When $T_{\text{drift}}$ is fixed [e.g., $T_{\text{drift}} = 10\text{ d}$ in Figure 12(a)], the HMM tracking of duration $N_T T_{\text{drift}}$ improves upon the sensitivity of the $F$-statistic search of duration $T_{\text{drift}}$ by a factor of $\sim N_T^{1/4}$. Figure 12(b) plots the minimum $h_{05\%}^{\text{obs}}$ in Figure 12(a) achieved in the band where the detectors are most sensitive (at 245 Hz) as a function of $T_{\text{drift}}$. The sensitivity achievable by the HMM scales as $\sim T_{\text{drift}}^{-1/4}$ for fixed $T_{\text{obs}} = 1\text{ yr}$. Figures 12(a) and 12(b) demonstrate together the scaling indicated by (E6). A fully coherent search using all the data (duration $N_T T_{\text{drift}}$) indicated by the blue dashed curve is more sensitive than the HMM of course. However it is computationally expensive. The theoretical scalings here also apply approximately to other $F$-statistic-based semi-coherent searches.

[1] B. Abbott et al., Physical Review D 76, 082001 (2007).
[2] The LIGO Scientific Collaboration and The Virgo Collaboration, eprint arXiv:1203.2674 (2012).
[3] K. Riles, Progress in Particle and Nuclear Physics 68, 1 (2013).
[4] G. Usmanovski, C. Cutler, and L. Bildsten, MNRAS 319, 902 (2000).
[5] N. K. Johnson-McDaniel and B. J. Owen, Physical Review D 88, 044004 (2013).
[6] C. Cutler, Physical Review D 66, 084025 (2002).
[7] A. Mastrano, A. Melatos, A. Reisenegger, and T. Akgünn, Monthly Notices of the Royal Astronomical Society 417, 2988 (2011).
[8] P. D. Lasky and A. Melatos, Physical Review D 88, 103005 (2013). [astro-ph.HE].
[9] B. J. Owen, L. Lindblom, C. Cutler, B. F. Schutz, A. Vecchio, and N. Andersson, Physical Review D 58, 084020 (1998).
[10] J. S. Heyl, The Astrophysical Journal 574, L57 (2002).
[11] P. Arras, E. E. Flanagan, S. M. Morsink, A. K. Schenk, S. A. Teukolsky, and I. Wasserman, The Astrophysical Journal 591, 1129 (2003).
[12] R. Bondarescu, S. A. Teukolsky, and I. Wasserman, Phys. Rev. D 79, 104003 (2009). [astro-ph.CO].
[13] C. Peralta, A. Melatos, M. Giacobello, and A. Ooi, The Astrophysical Journal 644, L53 (2006).
[14] C. A. van Eysden and A. Melatos, Classical and Quantum Gravity 25, 225020 (2008).
[15] M. F. Bennett, C. A. Van Eysden, and A. Melatos, Monthly Notices of the Royal Astronomical Society 409, 1705 (2010).
[16] A. Melatos, J. A. Douglass, and T. P. Simula, The Astrophysical Journal 807, 132 (2015).
[17] P. Haensel, V. A. Urpin, and D. G. Iakovlev, Astronomy and Astrophysics, vol. 229, no. 1, March 1990, p. 133-137. 229, 133 (1990).

---

11 One fully coherent search directed at the isolated compact object Calvera has been carried out, using Initial LIGO S5 data ($T_{\text{obs}} = 2\text{ yr}$). It is based on a resampling technique and saves computing cost by a factor $T_{\text{obs}}(\log T_{\text{obs}})^{-1}$, yielding a minimum upper limit $h_{05\%}^{\text{obs}} = 1.14 \times 10^{-25}$ at 152 Hz. [90].
FIG. 12. Analytical sensitivity $h_{0}^{95\%}$ of two coherent $F$-statistic searches (blue curves) and a semi-coherent HMM search (red curve) for an isolated young neutron star at the design sensitivity of Advanced LIGO. The blue solid curve indicates a coherent $F$-statistic search of duration $T_{\text{obs}}$, the maximum practical duration set by spin down in Ref. [36]. The blue dashed curve indicates a coherent $F$-statistic search with $T_{\text{obs}} = 360$ d. The red curve indicates a HMM search with $T_{\text{obs}} = 360$ d and $N_{T} = T_{\text{obs}}/T_{\text{drift}}$. (a) $h_{0}^{95\%}$ as a function of signal frequency $f_{0}$. The blue dashed curve achieves $h_{0}^{95\%}$ at 245 Hz in (a) as a function of $T_{\text{drift}}$. The HMM sensitivity scales as $\sim T_{\text{drift}}^{-1/4}$ for fixed $T_{\text{obs}}$. When $T_{\text{drift}}$ is fixed, the HMM (red curve) improves upon the $F$-statistic search (blue solid curve) by a factor of $\sim N_{T}^{1/4}$. The spikes in the first panel are instrumental drift for fixed $h_{0}$. When $h_{0}$ is fixed, the HMM (red curve) improves upon the $F$-statistic search (blue solid curve) by a factor of $\sim N_{T}^{1/4}$. The spikes in the first panel are instrumental drift at 245 Hz in (a)]

(2005).

[40] S. Dhurandhar, B. Krishnan, M. Mukhopadhyay, and J. T. Whelan, Phys. Rev. D 77, 082001 (2008).
[41] P. R. Brady and T. Creighton, Physical Review D 61, 082001 (2000).

[42] B. G. Quinn and E. J. Hannan, The Estimation and Tracking of Frequency (Cambridge University Press, 2001) p. 266.

[43] A. Viterbi, IEEE Transactions on Information Theory 13, 260 (1967).

[44] B. P. Abbott et al., (2017), arXiv:1704.03719

[45] P. Jaranowski, A. Królak, and B. F. Schutz, Physical Review D 58, 063001 (1998)

[46] C. Van Den Broeck, Classical and Quantum Gravity 22, 1825 (2005).

[47] R. Prix, Physical Review D 75, 023004 (2007)

[48] S. Suvorova, L. Sun, A. Melatos, W. Moran, and R. J. Evans, Physical Review D 93, 123009 (2016).

[49] R. Prix, LIGO Report T0900149 (June 2011).

[50] K. Wette, B. J. Owen, B. Allen, M. Ashley, J. Betzwieser, N. Christensen, T. D. Creighton, V. Dergachev, I. Gholami, E. Goetz, R. Gustafson, D. Hammer, D. I. Jones, B. Krishnan, M. Landry, B. Machenschalk, D. E. McClelland, G. Mendell, C. J. Messenger, M. A. Papa, P. Patel, M. Pitkin, H. J. Pletsch, R. Prix, K. Riles, L. S. de la Jordana, S. M. Scott, A. M. Sintes, M. Trias, J. T. Whelan, and G. Woan, Classical and Quantum Gravity, Volume 25, Issue 23, id. 235011 (2008).

[51] N. Andersson, D. Antonopoulou, C. M. Espinoza, B. Haskell, and W. C. G. Ho, (2017), arXiv:1711.05550

[52] G. Hobbs, A. G. Lyne, and M. Kramer, Monthly Notices of the Royal Astronomical Society 402, 1027 (2010).

[53] R. M. Shannon and J. M. Cordes, Astrophys. J. 725, 1607 (2010).

[54] G. Ashton, D. I. Jones, and R. Prix, Physical Review D 91, 062009 (2015).

[55] J. M. Cordes and D. J. Helfand, Astrophys. J. 239, 640 (1980).

[56] S. Price, B. Link, S. N. Shore, and D. J. Nice, Monthly Notices of the Royal Astronomical Society 426, 2507 (2012).

[57] A. Lyne, G. Hobbs, M. Kramer, I. Stairs, and B. Stappers, Science 329, 408 (2010).

[58] M. A. Alpar, R. Nandkumar, and D. Pines, Astrophys. J. 311, 197 (1986).

[59] P. Jones, Monthly Notices of the Royal Astronomical Society 246 (1990).

[60] A. Melatos and B. Link, Monthly Notices of the Royal Astronomical Society 437, 21 (2014).

[61] J. M. Cordes and G. S. Downs, The Astrophysical Journal Supplement Series 59, 343 (1985).

[62] G. H. Jansen and B. W. Stappers, Astronomy and Astrophysics 457, 611 (2006).

[63] K. S. Cheng, The Astrophysical Journal 321, 799 (1987).

[64] K. S. Cheng, The Astrophysical Journal 321, 805 (1987).

[65] J. O. Urama, B. Link, and J. M. Weisberg, Monthly Notices of the Royal Astronomical Society: Letters 370, L76 (2006).

[66] D. H. Shoemaker et al., LIGO Report No. T0900288 (2009).

[67] C. Messenger, H. J. Bulten, S. G. Crowder, V. Dergachev, D. K. Galloway, E. Goetz, R. J. G. Jonker, P. D. Lasky, G. D. Meadors, A. Melatos, S. Premachandra, K. Riles, L. Sammut, E. H. Thrane, J. T. Whelan, and Y. Zhang, Physical Review D 92, 023006 (2015).

[68] K. Wette, Physical Review D 85, 042003 (2012).
[69] R. Prix and M. Shaltev, Physical Review D 85, 084010 (2012).

[70] C. T. Y. Chung, A. Melatos, B. Krishnan, and J. T. Whelan, MNRAS 414, 2650 (2011).

[71] J. Ming, M. Alessandra Papa, B. Krishnan, R. Prix, C. Beer, S. J. Zhu, H.-B. Eggenstein, O. Bock, and B. Machenschalk, (2017), arXiv:1708.02173.

[72] L. D. Anderson, Y. Wang, S. Bihir, H. Beuther, F. Bigiel, E. Churchwell, S. C. O. Glover, A. A. Goodman, T. Henning, M. Heyer, R. S. Klessen, H. Linz, S. N. Longmore, K. M. Menten, J. Ott, N. Roy, M. Rugel, J. D. Soler, J. M. Stil, and J. S. Urquhart, Astronomy and Astrophysics (2017), 10.1051/0004-6361/201731019.

[73] B. P. Abbott et al., The Astrophysical Journal 851, L16 (2017).

[74] R. Bellman, Princeton University Press Princeton New Jersey, Vol. 70 (1957) p. 342.

[75] P. R. Brady, T. Creighton, C. Cutler, and B. F. Schutz, Physical Review D 57, 2101 (1998).

[76] B. J. Owen, Physical Review D 53, 6749 (1996).

[77] D. M. Whitbeck, PhD Thesis, The Pennsylvania State University (2006).

[78] R. Prix, Classical and Quantum Gravity 24, S481 (2007).

[79] D. Keitel, R. Prix, M. A. Papa, P. Leaci, and M. Siddiqi, Physical Review D 89, 064023 (2014).

[80] P. K. Patel, Search for gravitational waves from a nearby neutron star using barycentric resampling, Ph.D. thesis, California Institute of Technology (2011).