Resampling to accelerate cross-correlation searches for continuous gravitational waves from binary systems

Grant David Meadors, Badri Krishnan, Maria Alessandra Papa, John T. Whelan, and Yuanhao Zhang

Resampling to accelerate cross-correlation searches for continuous gravitational waves from binary systems

G.D. Meadors,[1, 2, *] B. Krishnan,[1] M.A. Papa,[1, 2, 3] John T. Whelan,[4] and Yuanhao Zhang[4, 1, 2]

[1]*Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-14476 Potsdam-Golm, Germany*
[2]*Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany*
[3]*University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201, USA*
[4]*Rochester Institute of Technology, Rochester, New York 14623, USA*

Continuous-wave (CW) gravitational waves (GWs) call for computationally-intensive methods. Low signal-to-noise ratio signals need templated searches with long coherent integration times and thus fine parameter-space resolution. Longer integration increases sensitivity. Low-Mass X-ray Binaries (LMXBs) such as Scorpius X-1 (Sco X-1) may emit accretion-driven CWs at strains reachable by current ground-based observatories. Binary orbital parameters induce phase modulation. This paper describes how resampling corrects binary and detector motion, yielding source-frame time series used for cross-correlation. Compared to the previous, detector-frame, templated cross-correlation method, used for Sco X-1 on data from the first Advanced LIGO observing run (O1), resampling is about $20\times$ faster in the costliest, most-sensitive frequency bands. Speed-up factors depend on integration time and search set-up. The speed could be reinvested into longer integration with a forecast sensitivity gain, 20 to 125 Hz median, of approximately 51%, or from 20 to 250 Hz, 11%, given the same per-band cost and set-up. This paper's timing model enables future set-up optimization. Resampling scales well with longer integration, and at $10\times$ unoptimized cost could reach respectively $2.83\times$ and $2.75\times$ median sensitivities, limited by spin-wandering. Then an O1 search could yield a marginalized-polarization upper limit reaching torque-balance at 100 Hz. Frequencies from 40 to 140 Hz might be probed in equal observing time with $2\times$ improved detectors.

I. INTRODUCTION

New gravitational-wave (GW) source types await sensitive analyses. Transient signals such as GW150914 [1] can reach strain amplitudes h_0 of approximately 10^{-21}. Yet-unseen continuous-wave (CW) signals, from sources such as non-axisymmetric neutron stars (NSs) [2], are constrained to be significantly weaker: for Scorpius X-1 (Sco X-1), the brightest Low Mass X-ray Binary (LMXB), the best 95%-confidence marginalized-polarization upper limit reaches 2.3×10^{-25} [3]. Accretion-driven torque-balance could drive GW emission from LMXBs: infalling matter's angular momentum is predicted to be balanced by that radiated gravitationally [4, 5]. Sco X-1 attracts attention [6] as the brightest persistent X-ray source [7]. Emission might be expected at a GW frequency f_0 equal to 2ν, for an NS spin frequency ν, assuming that the compact object in the system is an NS radiating *via* the $l = m = 2$ mass quadrupole moment. An NS could also emit *via* r-mode (Rossby) oscillations [8, 9], depending on the equation of state and dissipative mechanisms [10–12]. Its spin frequency is unknown, so an f_0 range must be searched.

In this paper, we discuss how to accelerate and increase the sensitivity of a broadband search for Sco X-1. CW analyses are computationally demanding. Long coherent integration times T_{coh}, for low signal-to-noise ratio (SNR) signals, induce a steep metric [2] on the parameter space, increasing the matched-filtering template density. While an optimal statistic [13] can maximize out *amplitude parameters*, the *Doppler parameters* need explicit templating. Sensitivity, which grows from longer integration, conflicts with computational cost, which grows faster. Semicoherent methods [14] tune this balance: an observing run of data is subdivided into coherent segments. Summing statistics from segments increases total sensitivity, while the metric depends mainly on the coherent-segment length. Sensitivity benefits from both total observing time T_{obs} and T_{coh}. Whole observation runs are typically used, with coherent segments as long as resources permit. Speed frees resources to be invested in coherent integration time. *Resampling* [13, 15] techniques can accelerate the cross-correlation methods (CrossCorr) [16–18] that have to date shown the most sensitive results for Sco X-1 in simulation [19] and Advanced LIGO data [3]. After over a decade of GW investigations into Sco X-1 [20–27], the nominal torque-balance level is near. Discovery may yield new astrophysics.

Detection becomes more likely as the GW strain amplitude h_0 sensitivity approaches torque-balance (TB). LMXB accretion torque could recycle NSs to higher ν [4]. If spin-up torque balances GW spin-down [5], the apparent speed limit on millisecond pulsars slightly over 700 Hz [28] may be explained. Sco X-1 and similar LMXBs could radiate GWs from NS asymmetries. By Bildsten Equation 4 [6], with characteristic strain h_c related by $h_c/h_0 = 2.9/4.0$, for an LMXB with flux $\mathcal{F}_{\mathrm{X-ray}}$,

* grant.meadors@ligo.org

$$h_c \approx 4 \times 10^{-27} \left[\frac{300 \text{ Hz}}{\nu_s} \times \frac{\mathcal{F}_{\text{X-ray}}}{10^{-8} \text{ erg cm}^{-2} \text{ s}^{-1}} \right]^{1/2}. \quad (1)$$

High X-ray flux (3.9×10^{-7} erg cm^{-2} s^{-1} [29]), assuming a nominal 1.4 solar mass, 10 km radius NS of unknown spin frequency, implies Sco X-1's h_0:

$$h_0 \approx 3.5 \times 10^{-26} [600 \text{ Hz}]^{1/2} f_0^{-1/2}. \quad (2)$$

Advanced LIGO Observing Run 1 (O1) data was searched [3] with the cross-correlation method [18], setting a 95%-confidence marginalized-polarization upper limit at 175 Hz of 2.3×10^{-25}, or 8.0×10^{-26} assuming optimal, circular polarization, respectively $3.5\times$ and $1.2\times$ above torque-balance. This analysis spanned 25 to 2000 Hz, with the detector noise curve and computational cost reducing the depth of the upper limits.

As Advanced LIGO [30], Advanced Virgo [31], and KAGRA [32] observatories improve, sensitivity varies linearly with noise amplitude spectral density(ASD), $S_H^{1/2}(f)$, for fixed T_{obs}. *Sensitivity depth $D^C(f_0)$* [33, 34] factors away this noise floor, to characterize analyses:

$$D^C(f_0) \equiv S_H^{1/2}(f_0)[h_0^C(f_0)]^{-1}. \quad (3)$$

Depth should be specified at a confidence C, such as $D^{95\%}(f_0)$, based on a strain upper limit h_0^C. *Coherent SNR* is proportional to $h_0 \sqrt{T_{\text{obs}}/S_H}$; deeper methods find lower SNR signals.

Methods vary [35] for finding CWs from NS in binary systems. Isolated CW techniques [13, 36–40] inform searches at unknown sky location, as well as for known ephemerides [41, 42], and also for the *directed* case of known sky location but uncertain ephemerides. Sco X-1 searches are directed (Table I). Five Doppler parameters arise from the binary orbit. New techniques address these parameters' computational cost [16, 18, 21, 23, 43–48], with more in development [34]. The cross-correlation method found all simulations in a 2015 Mock Data Challenge (MDC) [19] and sets O1 upper limits 3 to 4 times more stringent than others [26, 27]. The *resampled* cross-correlation method could surpass these limits.

Resampling was proposed [13] and detailed [15] for isolated-star $\mathcal{F}$-statistic calculations. Strain $h_0(t)$ is interpolated from the detector frame, where Earth and source motion introduce phase modulation, to the source frame. In the source frame, the statistic simplifies (with normalization factors determined by the detector antenna functions) to frequency bin power. Although interpolation is costly, subsequent computations can be faster than interpolating across time-varying frequency bins. We adapt the cross-correlation method for resampling. Speed-up and sensitivity performance projections are estimated from implemented code tested with simulated data. Deeper, resampled cross-correlation methods could bring CW analyses of Sco X-1 and similar LMXBs to the brink of detection.

Section II details the cross-correlation method, Section III explains resampling, and Section IV measures the cost and benefits. Figures 6 and 7 show predicted astrophysical reach. Section V concludes.

II. CROSS-CORRELATION METHOD

Detecting a sinusoid should be simple. Low SNR, amplitude- and phase-modulated sinusoids are hard. The 'CrossCorr' cross-correlation method [16, 18] intersects two paths to this problem: the stochastic radiometer [47, 54] and the multi-detector $\mathcal{F}$-statistic [13, 55, 56]. This cross-correlation method computes a statistic, ρ, which approaches the others in limiting cases. We summarize ρ to clarify, and to explain how resampling [13, 15], designed for the $\mathcal{F}$-statistic, is transferable. The principle remains – a semicoherent matched filter using a signal model for continuous, modulated GWs, then a frequentist statistic proportional to the power, $(h_0)^2$.

A. Signal model

Continuous waves from NS in binary systems are defined by a signal model in amplitude and Doppler parameters. Amplitude parameters $\bar{\mathcal{A}}^i$ [13], are factored out: reference phase Φ_0, polarization angle ψ, NS inclination angle ι (with respect to the line of sight), and strain amplitude h_0. Sky location is in right ascension α and declination δ.

The Doppler parameters λ for an isolated system include frequency f_0 and higher-order Taylor-expanded *spindown* (or *spinup*) terms $f^{(1)}$, $f^{(2)}$, *etc.* Assuming an NS source spinning at frequency ν, GW $f^{(k)} \equiv (\sigma d^k \nu(\tau)/d\tau^k | \tau = t_{\text{ref}})$, with emission time in the source frame τ, evaluated at arbitrary reference time t_{ref} (conventions follow [34]). For quadrupole emission, $\sigma = 2$. Assuming torque-balance, LMXB searches have set spindown terms to zero and instead consider *spin-wandering* [57], an unmodeled stochastic drift about f_0. For an isolated system without spindown, the measured frequency in the solar system barycenter (SSB) will be constant.

For a binary system, the λ parameters further include $(a_P, P, T_{\text{asc}}, T_{\text{p}}, e)$. Orbital projected semi-major axis, in time units, is $a_p \equiv (a \sin i)/c$ (a measured in light-seconds). Orbital period is P. Time of ascension is T_{asc}, when the compact object crosses the ascending node, heading away from an SSB observer. Because only the companion's inferior conjunction time T_0 [51] is known, the compact object $T_{\text{asc}} = T_0 - P/4$ [19] (stated in SSB GPS seconds). Time of periapsis passage is T_{p}. Orbital eccentricity is e. When $e = 0$, $T_{\text{p}} = T_{\text{asc}}$ by convention. For Sco X-1, α and δ are precise enough that one point covers uncertainty.

Sco X-1 parameter	Ref.	Value	Uncertainty	Units
Right ascension (α)	[49]	16:19:55.067	$\pm 0.06''$	—
Declination (δ)	[49]	$-15°38'25.02''$	$\pm 0.06''$	—
Distance (d)	[50]	2.8	± 0.3	kpc
X-ray flux at Earth ($\mathcal{F}_{\mathrm{X-ray}}$)	[29]	3.9×10^{-7}	—	erg cm^{-2} s^{-1}
Orbital eccentricity (e)	[19, 51]	< 0.068	(3σ)	—
Orbital period (P)	[51]	68023.70	± 0.04	s
Orbital projected semi-major axis (a_p)	[3, 52]	1.805	± 1.445	s
Compact object time of ascension (T_{asc})	[19, 51]	897753994	± 100	s
Companion mass (M_2)	[53]	0.42	—	M_{sol}

TABLE I. Sco X-1 prior measured parameters from electromagnetic observations (compare [3, 46]). Uncertainties are 1σ, except for a_p, explained below. The projected semi-major axis is in time units, $a_p = (a \sin i)/c$; it depends on velocity $K_1 = [10, 90]$ km s^{-1} with approximately-uniform uncertainty distribution (previously [19, 20, 51], increased uncertainty after private communication [52]). Time of ascension T_{asc} is defined here as the time when the compact object crosses the ascending node, heading away from the observer in the solar system barycenter (SSB). Because the companion's inferior conjugation T_0 is measured [19], we calculate $T_{\mathrm{asc}} = T_0 - P/4$ [51].

1. Strain and amplitude parameters

Strain amplitude $h(t)$ is measured; GW phase $\Phi(t; \lambda)$ is key to its signal model:

$$h(t) = [F_+(t; \alpha, \delta), F_\times(t; \alpha, \delta)] \begin{bmatrix} A_+ \cos \Phi(t; \lambda) \\ A_\times \sin \Phi(t; \lambda) \end{bmatrix}, \quad (4)$$

where t is detector GPS time measured; F_+ and $F_\times$ are called *beam-pattern functions*. The amplitude model factors loosely depend on time and sky location *via* the detector response in the *antenna functions* $a(t; \alpha, \delta)$ and $b(t; \alpha, \delta)$ [13, 16]. Since we discuss known sky location targets, (α, δ) will be implicit in $a(t)$, $b(t)$:

$$\begin{bmatrix} F_+(t) \\ F_\times(t) \end{bmatrix} = \begin{bmatrix} a(t) \cos 2\psi + b(t) \sin 2\psi \\ b(t) \cos 2\psi - a(t) \sin 2\psi \end{bmatrix}, \quad (5)$$

$$\begin{bmatrix} A_+ \\ A_\times \end{bmatrix} = h_0 \begin{bmatrix} \frac{1 + \cos^2 \iota}{2} \\ \cos \iota \end{bmatrix}. \quad (6)$$

Amplitude parameters can be projected into four new coordinates which affect the waveform linearly [13, 55, 56, 58, 59]. These *canonical* coordinates $\mathcal{A}^\mu$ satisfy, for basis functions $h_\mu(t; \lambda)$,

$$h(t; \lambda) = \sum_{\mu=1}^{4} \mathcal{A}^\mu h_\mu(t; \lambda). \quad (7)$$

Maximization, or marginalization, over $\mathcal{A}^\mu$ leads to approximately Neyman-Pearson optimal statistics (respectively $\mathcal{F}$, $\mathcal{B}$) [56]. The cross-correlation method obtains a similar statistic ρ with different motives [16] (see Section II B).

2. Doppler parameters

The $\Phi(t; \lambda)$ model is defined with source time τ as a function of t, *via* SSB time t_{SSB}. In the spindown-free source frame, $\Phi(\tau) = 2\pi f_0 \tau$, so $\Phi(t; \lambda) = 2\pi f_0 \tau(t_{\mathrm{SSB}}(t; \alpha, \delta); \lambda)$. One can find the barycentric time t_{SSB} from (α, δ), *via* the vector $\vec{r}(t)$ pointing from the SSB to the detector and the unit vector $\vec{n}$ pointing from the SSB to the source. The latter vector is defined as $\vec{n}(\alpha, \delta) = (\cos \alpha \cos \delta, \sin \alpha \cos \delta, \sin \delta)$ [34]. With GW unit wavevector $\vec{k} = -\vec{n}$ in the far-field approximation,

$$t_{\mathrm{SSB}}(t; \alpha, \delta) = t + \frac{\vec{r}(t) \cdot \vec{n}(\alpha, \delta)}{c}. \quad (8)$$

The relativistic t_{SSB} is corrected for Shapiro and Einstein delays, in addition to the Earth orbital and rotational Roemer delays encoded by $\vec{r}$ [34]. The binary orbital Roemer delay comes from the projected radial distance R along the line of sight. Following conventions [34, 60],

$$\tau(t_{\mathrm{SSB}}; \lambda) = t_{\mathrm{SSB}} - \frac{d}{c} - \frac{R(t_{\mathrm{SSB}}; \lambda)}{c}, \quad (9)$$

wherein larger R signifies greater distance from the binary barycenter (BB) along the line of sight, away from the observer. Source distance will affect h_0 and cause an overall time shift d/c equivalent to changing Φ_0, and inertial motion effects an overall constant Doppler shift to f_0. As d would also affect electromagnetic observations and is indistinguishable from other parameters, we now drop (d/c), in effect equating the SSB with the BB.

Kepler's equations involve a constant argument of periapse ω (the angle from the ascending node to periapsis in the direction of motion, dependent on T_p and T_{asc}) and a time-varying eccentric anomaly E (implicit in τ [34]). These equations describe system dynamics:

$$\tau = T_{\mathrm{p}} + \frac{P}{2\pi}(E - e\sin E), \tag{10}$$

$$\frac{R}{c} = a_p\left[\sin\omega(\cos E - e) + \cos\omega\sin E\sqrt{1 - e^2}\right].\tag{11}$$

Sco X-1's orbit is near-circular ($e < 0.068$ at 3σ), so we will focus on $e = 0$, though resampling can handle elliptical orbits. Sco X-1's a_p is four orders of magnitude less than P, so we approximate $E(\tau) = E(t)$. Let $\Omega \equiv 2\pi/P$. In this circular case [18],

$$\frac{R(t;\lambda)}{c} = a_p\sin\left(\Omega[t - T_{\mathrm{asc}}]\right), \tag{12}$$

$$\phi(t;\lambda) = 2\pi f_0\left[t_{\mathrm{SSB}}(t;\alpha,\delta) - \frac{R(t;\lambda)}{c}\right], \tag{13}$$

$$\Phi(t;\lambda) = \Phi_0 + \phi(t;\lambda). \tag{14}$$

Phase modulation induces an effective frequency modulation depth, Δf_{obs}. This modulation adds to Doppler shift from detector velocity, $\vec{v} = d\vec{r}/dt$ (dominated by Earth's orbit v_{Earth}), when calculating the total physical frequency bandwidth $\Delta f_{\mathrm{drift}}$ through which the signal can drift:

$$\Delta f_{\mathrm{drift}} = 2 \times \left(\frac{\max\left(\vec{v}\cdot\vec{n}\right)}{c} + \Delta f_{\mathrm{obs}}\right), \tag{15}$$

$$\Delta f_{\mathrm{obs}} = a_p\Omega f_0. \tag{16}$$

With $|\vec{v}| \approx v_{\mathrm{Earth}}$, $\max(v/c) \approx 10^{-4}$ (lower off-ecliptic).

For an unmodulated signal, $d\Phi/dt = 2\pi f_0$, reducing to a Fourier transform [2]. For modulated signals, $\phi(t)$ must be tracked to maintain coherence. Given Equation 14, the cross-correlation method tracks a CW signal as the signal changes instantaneous frequency.

Mismatch in Doppler parameters can lead to false dismissal. The phase mismatch metric [2] ([18] for the cross-correlation method) sets the parameter-space density required for Doppler parameters. A mismatch in the phase-model Roemer delay of about a half-cycle of f_0 between the beginning and end of each integration time T_{coh} will lose the signal. (A 100 Hz signal accumulates $\mathcal{O}(10^2)$ cycles over a_p of Sco X-1 and $\mathcal{O}(10^5)$ cycles over 2 AU). The computational cost stems from the parameter-space density needed for the long T_{coh} that low-SNR signals require.

We define detection statistics for these signals. This paper will show that resampling is a more efficient way to compute the cross-correlation method's ρ statistic.

B. Cross-correlation statistic

The goal is to calculate the statistic, ρ, as efficiently as possible. See Figure 1 for a cost per template comparison of the previous 'demodulation' and resampled methods.

Let us define ρ as in Whelan *et al* [18]. (In Appendix A we compare ρ, like Dhurandhar *et al* [16], with the $\mathcal{F}$-statistic [13] and radiometer [47, 54]). Data start by being parcelled into short Fourier transforms (SFTs) [61], each of duration T_{sft}. The total data set spanning T_{obs} for Q detectors may contain up to $N_{\mathrm{sft}} \leq QT_{\mathrm{obs}}/T_{\mathrm{sft}}$ SFTs.

The *cross-correlation* in our method is made between pairs of SFTs: the first component of the pair is indexed by K, the second component by L. In the Whelan *et al* construction, K and L span all detectors, meaning they both can range from 0 up to N_{sft}. (Particulars are discussed in Section III B 1, where K and L are redefined). The sets of SFTs $\{K\}$ and $\{L\}$ are defined by an allowable lag-time, T_{max}, the difference between start times of given SFTs K and L. It is common to require $K \neq L$ (to avoid auto-correlation). SFT pairs KL in the set $\mathcal{P}$ are *cross-correlated.*

A time series $x_K(t) = h_K(t) + n_K(t)$, signal h and noise n, has one-sided power spectral density (PSD) S_K. Analyze the Fourier transform $\tilde{x}$ (using Equation 2.1 [18] conventions), with sampling time δt, SFT mid-time t_K:

$$\tilde{x}_{Km} = \sum_{j=0}^{N-1} x_K(t_K - T_{\mathrm{sft}}/2 + j\delta t)e^{-\mathrm{i}2\pi jm\delta t/T_{\mathrm{sft}}}\delta t \tag{17}$$

so normalized data z_{Km} in frequency bin m (k in [18] and our appendices) is,

$$z_{Km} = \tilde{x}_{Km}\sqrt{\frac{2}{T_{\mathrm{sft}}S_K}}. \tag{18}$$

SFT bin frequency is f_m, but the signal instantaneous frequency is f_K; these must not be confused. The discrepancy, to the nearest bin from the instantaneous frequency, is κ_{Km},

$$f_m = \frac{m}{T_{\mathrm{sft}}}, \tag{19}$$

$$\kappa_{Km} = m - f_K T_{\mathrm{sft}}. \tag{20}$$

Multiple bins in a set $\mathcal{K}_K$ are part of the *Dirichlet kernel*, discussed around Equation 6.5 of [62]. The signal contribution to each bin is found by the normalized *sinc* function, $\mathrm{sinc}\alpha = \frac{\sin\pi\alpha}{\pi\alpha}$. The total data vector $\mathbf{z}$ has elements z_K, which are the Fourier-transformed data. Each element is summed from all bins that could contain a signal at a frequency f (implicitly specified by the set $\mathcal{K}_K$ and the f_K model in κ_{Km}), then indexed by SFT K,

$$\Xi_K \equiv \sqrt{\sum_{m'\in\mathcal{K}_K}\mathrm{sinc}^2(\kappa_{Km'})}, \tag{21}$$

$$z_K = \frac{1}{\Xi_K}\sum_{m\in\mathcal{K}_K}(-1)^m\mathrm{sinc}(\kappa_{Km})z_{Km}. \tag{22}$$
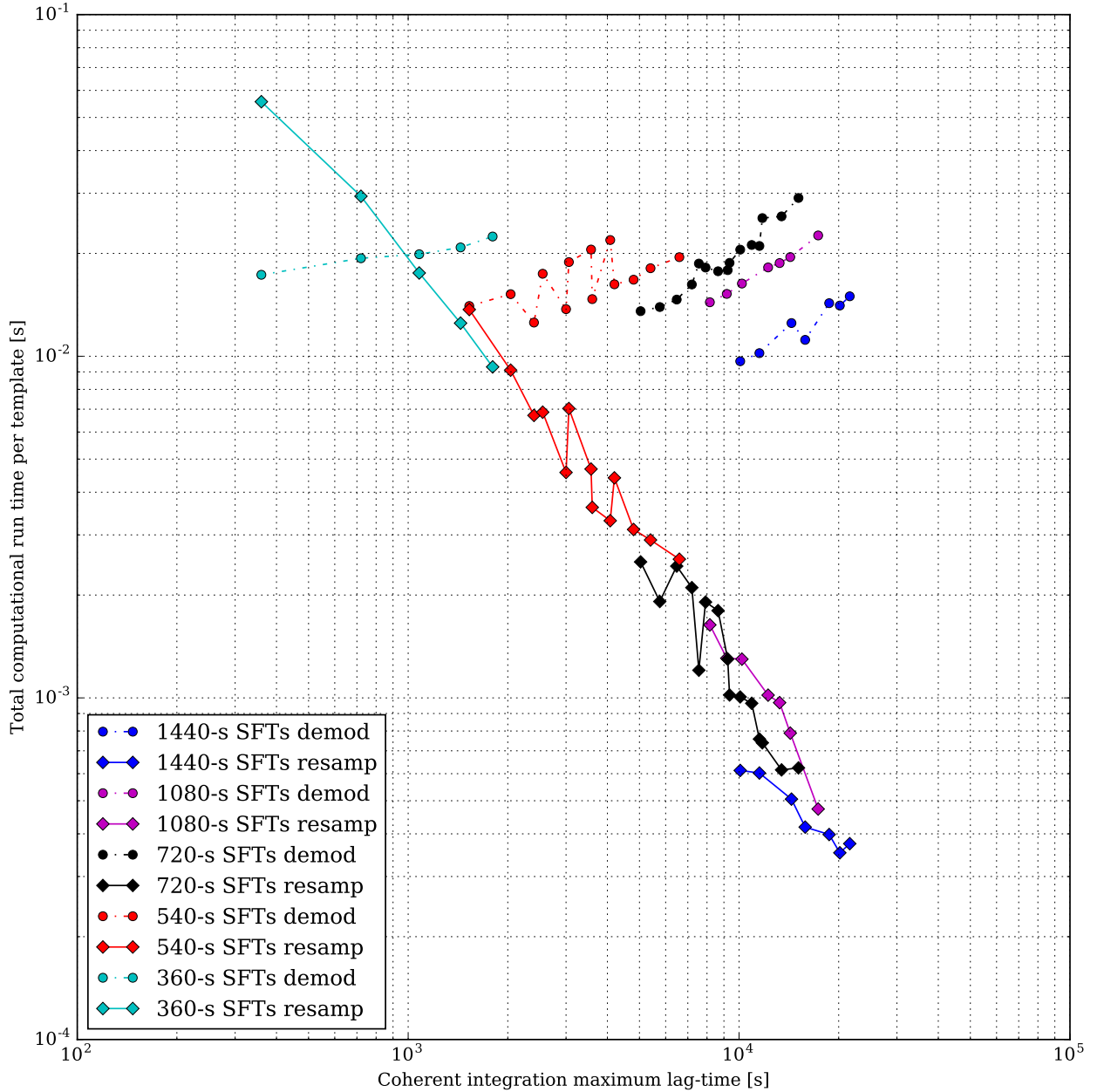
FIG. 1. Measured demodulation *(dotted lines, cirles)* and resampling *(solid lines, diamonds)* computational cost-per-template *vs* lag-time $T_{\max}$. Lag-time is the maximum difference in start times between the cross-correlation method's data segments. Results plotted for various Short Fourier Transform (SFT) [61] data durations T_{sft} *(colors)*. Average run time for a given lag-time is shown; roughly 190 'demod' and 180 'resamp' measurements done on the Atlas cluster. Run time shown is total (not per-segment). The overhead costs for resampling include the barycentering interpolation and FFT (with Doppler wings). Doppler wings large in proportion to f_{band} reduce resampling efficiency; these tests use 0.050-Hz f_{band}. At longer lag-times, these costs are proportionally less, so resampling is faster. Demodulation's cost increases with $T_{\max}$ for fixed T_{sft}, but because the setup uses longer T_{sft} with longer $T_{\max}$ (such as 360-s SFTs at 1000-s lag-time but 1440-s SFTs at 20000-s lag-time), the cost per-template is kept between about 10 and 30 ms. Per-template costs for demodulation jump when comparing one SFT duration to another. This effect is because longer SFTs require fewer pairing operations, but length is limited by source and detector acceleration timescales. At fixed T_{sft}, demodulation run time increases with lag-time. Resampling allows pairing to be done in the source frame regardless of acceleration, with a flexible timescale T_{short} instead of T_{sft}. Fewer pairs and use of the FFT make resampling better suited to long $T_{\max}$. From one SFT duration to another, resampling per-template costs are usually consistent. At long $T_{\max}$, resampling per-template costs are lower than demodulation.

The cross-correlation method constructs ρ with a *filter*, Hermitian weighting matrix $\mathbf{W}$. It uses the conjugate transpose †. With matrix entries KL that correlate elements K from the SFT vector $\mathbf{z}$,

$$\rho = \mathbf{z}^\dagger \mathbf{W} \mathbf{z}, \tag{23}$$

or in explicit notation, $\rho = \sum_K \left(\sum_L z_L^* W_{KL} \right) z_K$. Equation 23 depends, *via* $\mathbf{W}$, on the point in λ parameter space (including frequency f) of the signal model.

A near-optimal $\mathbf{W}$ is the geometrical factor $\hat{\Gamma}_{KL}^{\text{ave}}$ (chosen for ψ-independence, Whelan *et al* Equation 2.33 [18]). Let a hat symbol indicate noise-weighted normalization, *e.g.*, $\hat{a}^K \equiv \sqrt{2T_{\text{sft}}/S_K}\, a^K$. Taking a^K ($a(t)$ at the (mid-)time of SFT K) and likewise a^L, b^K, b^L, we can find $\hat{\Gamma}_{KL}^{\text{ave}}$. With overall normalization N (*ibid.* Equation 3.6),

$$\hat{\Gamma}_{KL}^{\text{ave}} = \frac{1}{10}\left(\hat{a}^K \hat{a}^L + \hat{b}^K \hat{b}^L \right), \tag{24}$$

$$N = \left(2 \sum_{KL \in \mathcal{P}} \Xi_K^2 \Xi_L^2 (\hat{\Gamma}_{KL}^{\text{ave}})^2 \right)^{-1/2}. \tag{25}$$

Another weight, $\hat{\Gamma}_{KL}^{\text{circ}} = \frac{1}{10}(\hat{a}^K \hat{b}^L - \hat{b}^K \hat{a}^L)$, is also ψ-independent. Combining $\hat{\Gamma}^{\text{ave}}$ and $\hat{\Gamma}^{\text{circ}}$ can fix ι.

To obtain ρ, Equation 2.36 of [18] (analogous to Equation 4.11 of [16]), we cross-correlate with the paired data in SFTs L indexed by bin n. We unite Fourier bins using the filter, complex conjugation $*$, and the signal model phase difference between SFTs, $\Delta \Phi_{KL} = \Phi_K - \Phi_L$:

$$\begin{aligned}
\rho = N \sum_{KL \in \mathcal{P}} \hat{\Gamma}_{KL}^{\text{ave}} \sum_{m \in \mathcal{K}_K} \sum_{n \in \mathcal{K}_L} (-1)^{m-n} \\
\times \operatorname{sinc}(\kappa_{Km}) \operatorname{sinc}(\kappa_{Ln}) \\
\times (e^{\mathrm{i}\Delta\Phi_{KL}} z_{Km}^* z_{Ln} + e^{-\mathrm{i}\Delta\Phi_{KL}} z_{Km} z_{Ln}^*).
\end{aligned} \tag{26}$$

Implicit in Φ_K is f_K, hence all Doppler parameters: ρ must be calculated for each λ template. Since billions [3] of templates are common, efficiency is paramount. Thanks to links between ρ and the $\mathcal{F}$-statistic that are explored in Appendix A, resampling speeds the search.

III. RESAMPLING

Many signals can be resampled into the source frame. (Compact binary coalescences were contemplated first [63]). This paper focuses on CWs [13] and adheres in notation to code documentation [64, 65]. Resampling abstractly moves phase demodulation from $\mathbf{W}$ onto $\mathbf{z}$.

Delay causes phase modulation: Equation 13 is Roemer-delayed by Earth and source binary motion. We want to sample $\phi(t; \lambda) = 2\pi f_0 \tau$, but in equally-spaced τ (source frame) instead of equally-spaced t (detector frame). Although they consider spindown rather than

binary parameters, $\tau \sim t_b$ in [15]; calculating Equation 8 and Equation 9 (with numerical solutions to Equations 10 and 11), $\tau(t; \lambda)$ can be found.

Because $x(t)$ is discrete, sampling $x(\tau)$ requires interpolation. The sinc function interpolates between time-domain samples, paralleling frequency-domain use [62]. As it is computationally-prudent to analyze small frequency bands f_{band} independently, data are heterodyned, by selecting the band of interest from a Fourier transform, then inverse Fourier transforming into a downsampled, complex time series, then interpolating. Since this procedure differs from [15], we describe it.

A time series $x(t)$ sampled at δt has a Nyquist frequency of $f_N = 1/(2\delta t)$. Each SFT K contains its own set of time indices j ranging from 0 to $N-1$, so j implicitly refers to K. With respect to an arbitrary reference time, $t = t_K - T_{\text{sft}}/2 + j\delta t$. Given a set of $M = T_{\text{obs}}/T_{\text{sft}}$ SFTs indexed by K, each with frequency bins k, spaced by $\delta f = 1/T_{\text{sft}}$, with $N = T_{\text{sft}}/(\delta t)$ samples, $x(t)$ can be reconstructed by the inverse FFT:

$$x_K(t_K - T_{\text{sft}}/2 + j\delta t) = \sum_{k=0}^{N-1} z_{Kk} e^{i2\pi jk\delta t/T_{\text{sft}}} \delta f. \tag{27}$$

Time series segments and frequency bands can be selected by indices.

Equation 27 can be simplified by using the index $q_K \equiv (t_K - T_{\text{sft}}/2)/(\delta t) + j$ in its argument. The q_K is the index with respect to the start of T_{obs}. A new sampling interval $\delta t'$ and corresponding index q_K' for some time $t = q_K' \delta t'$ can define a downsampled time series. This time series (heterodyne frequency f_h) is $x'(q'\delta t')$ and is produced as in Appendix B.

A. Resampling theory

1. Interpolation

When data is unaliased and approximately stationary during each SFT, $x'(q'\delta t')$ is a complete representation. Sinc-interpolation allows us to interpolate $x'(\tau)$. The Shannon formula as implemented [64] states that for integer D Dirichlet elements, integer index j and $j^* \equiv \operatorname{round}(t/(\delta t))$, $j0 \equiv j^* - D$, and a window w_j (here, Hamming with length $2D + 1$),

$$\delta_j \equiv \frac{t - t_j}{\delta t}, \tag{28}$$

$$x(t) \approx \frac{\sin(\pi \delta_{j0})}{\pi} \sum_{j=j^*-D}^{j^*+D} (-1)^{(j-j0)} \frac{x_j w_j}{\delta_j}, \tag{29}$$

converging when $D \to \infty$. A typical $D = 8$, minimizing costs of sinc-interpolation (linear in D) plus subsequent FFTs (linear in Appendix B's Δf_{load}).

2. Resampling into the source frame

Let our source-frame time series be indexed by r with constant spacing $\delta t'$: $\tau = r\delta t'$. We use the function $t(\tau; \lambda)$, the functional inverse of the function $\tau(t; \lambda)$ from Equation 9. Over timescales T_{sft} when the signal stays in one frequency bin, $(d^2\tau/dt^2)T_{\text{sft}}^2 f_0 \ll 1$, Taylor approximation is valid around t_0:

$$\tau(t; \lambda) \approx \tau(t_0; \lambda) + \left[\frac{d\tau(t)}{dt}\big|_{t=t_0}\right]t, \qquad (30)$$

$$t(\tau; \lambda) \approx t(\tau_0; \lambda) + \left[\frac{d\tau(t)}{dt}\big|_{t=t_0}\right]^{-1}\tau, \qquad (31)$$

making computations practical.

Translating from detector time to source time introduces a timeshift $\Delta t^* = r\delta t' - t(r\delta t'; \lambda)$ to $x(t)$. The discrete source-frame time series is $x'(r\delta t') = \exp\left(-i2\pi f_h \Delta t^*\right)x'(t(r\delta t'; \lambda))$:

$$\delta_{q'} \equiv \frac{t(r\delta t'; \lambda) - q'\delta t'}{\delta t'}, \qquad (32)$$

$$r^* \equiv \text{round}\left(\frac{t(r\delta t'; \lambda)}{\delta t'}\right), \qquad (33)$$

$$x'(r\delta t') \approx \frac{\sin(\pi\delta_{q0'})}{\pi}e^{-2\pi f_h[r\delta t' - t(r\delta t'; \lambda)]}$$
$$\times \sum_{q'=r^*-\text{D}}^{r^*+\text{D}} (-1)^{(q'-q0')}\frac{x'_{q'}w_{q'}}{\delta_{q'}}. \qquad (34)$$

Then $x'_r \equiv x'(r\delta t')$ is the complex, heterodyned, downsampled, discrete time series that equally samples the source frame $x(\tau)$.

Roemer delays vanish in $x(\tau)$, if the Doppler parameters λ are accurate. Mismatch results in residual phase modulation. No finite lattice of λ can perfectly sample the space. The required resolution is determined by the phase mismatch metric g [2].

Derivatives $d/d\lambda$ for $\lambda \in (f, a_p, T_{\text{asc}}, P)$ have been calculated for the cross-correlation method's metric [18]. In the similar $\mathcal{F}$-statistic metric [34], e and T_p are discussed. The metric is computed in software over the phase mismatch $\Delta\Phi_{\alpha,i}$ for the cross-correlation method's pairs indexed by $\alpha = KL$ and Doppler parameters indexed by i,

$$g_{ij} \approx \frac{1}{2}\left\langle\left(\frac{\partial(\Phi_K - \Phi_L)}{\partial\lambda^i}\right)\left(\frac{\partial(\Phi_K - \Phi_L)}{\partial\lambda^j}\right)\right\rangle_\alpha, \quad (35)$$

extending to any Doppler parameters in the phase model. (Metric vielbeins represent the natural units of distance for a parameter-space vector).

Given the metric, a lattice is calculated with the spacing in each dimension set by the allowed mismatch, λ_μ.

Mismatch is a tunable choice about the statistic's acceptable *fractional loss*: $\mu_\lambda = (\max(\rho)-\rho)/\max(\rho)$. A simple cubic lattice grid for a diagonal metric has spacings $\delta\lambda^i$,

$$\delta\lambda_i = \sqrt{\frac{\mu_{\lambda_i}}{g_{ii}}}. \qquad (36)$$

However, the metric is only a local approximation [2]. The total derivative $d\tau$ contains many approximate degeneracies, for example when frequency mismatch df equals modulation depth mismatch $d\Delta f_{\text{obs}}$ arising from offset a_p or T_{asc} (see Appendix C). Mismatch studies are thus needed to verify the loss and chose spacings. Each lattice point in orbital parameter space must have its own resampled $x(\tau)$.

Resampling interpolation yields $x(\tau)$ so that a putative signal is concentrated at a single frequency f_0. Next, taking the Fourier transform [13, 15] generates ρ.

B. Resampled cross-correlation method implementation

Source-frame $x(\tau)$ speeds Section II B's ρ calculation. Supplied with T_{obs}, we divide data into semicoherent segments with a shortest timescale of T_{short}, replacing T_{sft}. This T_{short} is the duration we will take from each K side of a pair of the cross-correlation method. The L side of the pair will be composed of all other T_{short} intervals with start times up to a *maximum lag-time* T_{max} before or after. A total, cross-detector, coherent integration duration of T_{coh} includes a central T_{short} plus T_{max} on both sides:

$$|\tau_K - \tau_L| \leq T_{\text{max}}, \qquad (37)$$
$$T_{\text{coh}} = 2T_{\text{max}} + T_{\text{short}}. \qquad (38)$$

For same-detector correlations, only T_{max} on one side is typically used, to avoid auto-correlation and double-counting, but we preserve the above definition of T_{coh} to keep frequency resolution the same.

Times τ_K and τ_L evenly divide the resampled time series if calculated in the source frame, though this means that slightly unequal amounts of detector data go into T_{short}. As $|\tau_K - t_K| \leq |\vec{r} \cdot \vec{n}/c + a_p|$, the difference between an interval start time in detector and source frame is bounded by the Roemer delay. We neglect these effects because relative inequality from one interval to the next is proportional to $d\tau/dt \leq 2 \times 10^{-4}$. Based on prior experience [34], these delays do not affect the metric estimation. For the cross-correlation method's metric [18], the goal is to constrain the (pair-averaged) phase mismatch over T_{max} from offset $\delta\lambda_i$, which grows linearly proportionally to T_{max}, so it is negligible from the phase mismatch over $(1 + d\tau/dt)T_{\text{max}}$.

Nor are average noise weightings affected much by resampling, because the normalization N is a sum over

T_{obs}. However, weightings are based on average noise per SFT. To find the weights, we average noise for each T_{short} interval by interpolating with Equation 29. Terms T_{sft} in Section II B become replaced with T_{short}.

The current implementation zero-pad gaps instead of skipping them. These gaps contribute nothing to ρ, and, because the noise-weighted antenna functions $\hat{a}(t)$ and $\hat{b}(t)$ give gaps zero weight, they contribute nothing to N.

Compared to the non-resampling cross-correlation method [18], resampling yields two benefits. First, T_{short} supercedes T_{sft}, the latter being limited by modulation moving the signal out of bin. Increasing T_{short} reduces the number of (new) pairs, $N_{\mathrm{pairs}} \approx N_{\mathrm{det}}^2 T_{\mathrm{max}} T_{\mathrm{obs}} T_{\mathrm{short}}^{-2}$ (replacing T_{sft} from Equation 3.27 in [18]). Because sensitivity is, to zeroth order, proportional to $h_0^{\mathrm{sens}} \propto (N_{\mathrm{det}}^2 T_{\mathrm{obs}} T_{\mathrm{max}})$, independent of T_{sft}, but cost is linearly proportional to the number of templates times the number of pairs, it is optimal to minimize the number of pairs by maximizing T_{short}.

Second, the number of frequency templates required is automatically supplied by an FFT. An FFT over a time period T_{coh} is spaced at $1/T_{\mathrm{coh}} \propto T_{\mathrm{max}}^{-1}$. This scaling comes from the metric element g_{ff} for that lag-time, indepedent of T_{sft} and resampling. Rather than needing to repeat this fine frequency grid for every SFT, resampling allows all the data to be gathered into one FFT with time $T_{\mathrm{FFT}} \geq T_{\mathrm{coh}}$. (For finer sampling, the FFT can be zero-padded; for coarser, its output can be decimated).

1. Pair selection for resampled statistic

Resampled $x'(\tau)$ as given by Equation 34 must be divided into pairs to calculate the ρ statistic.

The set of pairs $\mathcal{P}$ must be constructed. Taking Q detectors, they are indexed by X for the first component of the cross-correlation method's pair and Y for the second component. These X, Y indices range from 0 to $Q - 1$. An option exists to exclude same-detector correlations, as in the stochastic radiometer. Here, we allow same-detector correlations, except same-detector same-time correlations, that is, the auto-correlation. We reuse indices K and L from previous sections but restrict the range of each to a single detector. Indexing T_{short} intervals is marked by K for detector X and L for detector Y. Indices K, L range from 0 to $M = T_{\mathrm{obs}}/T_{\mathrm{short}}$, regardless of any gaps. Approximating Equation 37 in the detector frame, such that

$$\{L|K\} : |KT_{\mathrm{short}} - LT_{\mathrm{short}}| \leq T_{\mathrm{max}}, \tag{39}$$
$$\implies \{K - T_{\mathrm{max}}/T_{\mathrm{short}}, \ldots, K + T_{\mathrm{max}}/T_{\mathrm{short}}\},$$

which is straightforward when T_{max} is an integer multiple R of T_{short}. (Performance is best in practice when $T_{\mathrm{short}} = T_{\mathrm{max}}$). This set $\{L|K\}$ contains $M_L = 2R + 1$ elements for cross-detector correlations and R for same-detector correlations, to avoid double-counting.

Detector-time pairing is predictable, and it is acceptable because K to $K + 1$ differences are of order $d\tau/dt \approx 2 \times 10^{-4}$. Yet the resampled time series do not start at precisely the same source frame time. Let $(\tau_X = \tau_K | K = 0)$, $(\tau_Y = \tau_L | L = 0)$. They can differ by $(\vec{r}_X - \vec{r}_Y) \cdot \vec{n}/c$, which for ground-based detectors is of order 10 ms at most. This $\Delta\tau_{XY}$ is still a full cycle at 100 Hz, and it must be accounted for, by timeshifting the resampled time series to the same starting epoch. The correct factor is the physical frequency f_0. Differences $\tau_K - \tau_L$ require a further timeshift at the heterodyned frequency, $f_0 - f_{\mathrm{het}}$, as they are internal to the resampled time series.

2. Fourier transform size and phase shift

The above definitions separate $\mathcal{P}$ pairs into intervals and detectors. To construct ρ from resampled data in these pairs using an FFT, we require the number of FFT samples, N_{FFT}. The metric resolution answers this question. Then we will substitute the pair definition into ρ to make an explicit quadruple sum.

The metric spacing $\delta\lambda_f$ will be achieved by an FFT of duration $1/(\delta\lambda_f)$. For typical mismatch μ_f, Equation 36 and Equation 4.31a of [18] yield $\delta\lambda_f < 1/T_{\mathrm{coh}}$. Specifically, Equation 4.33 [18] becomes $(3/4)T_{\mathrm{max}}^2$ on the right-hand side in the case $T_{\mathrm{max}} = T_{\mathrm{short}}$,

$$\delta\lambda_f = \sqrt{\frac{6\mu_f}{\pi}} \frac{1}{T_{\mathrm{coh}}}, \tag{40}$$

which provides $\delta\lambda_f T_{\mathrm{coh}} < 1$ up to $\mu_f \approx 0.52$. This is a high value of mismatch. Any FFT with that mismatch or finer frequency spacing is automatically long enough to include all the data in T_{coh}. (For coarser mismatch, decimation by a ratio $\nu_D \equiv \mathrm{ceil}(\delta\lambda_f \times T_{\mathrm{coh}})$ after the FFT can select the frequencies of interest). Conversely, if $\mu_f = 0.1$, $\delta\lambda_f$ implies FFT duration $\geq 2.3T_{\mathrm{coh}}$. Dirichlet frequency interpolation is replaced by zero-padding to the metric resolution.

The recovered fraction of spectral power is known from Equation 3.18 [18], $\langle\Xi^2\rangle$ (to which ρ is linearly proportional): for Dirichlet interpolation with m bins,

$$\langle\Xi^2\rangle = 2\int_0^{m/2} \mathrm{sinc}^2\kappa \, d\kappa. \tag{41}$$

In that paper, $m = 2$ was recommended to capture 0.903 of ρ. The function $\delta_{T_{\mathrm{sft}}}(f - f')$ is a continuous function determined by data; only κ_{Kk} are discrete. Zero-padding from T_{coh} to T_{FFT} (and taking only 1 bin of the FFT, so $m = 1$) gives,

$$\langle\Xi^2\rangle_{\mathrm{resamp}} = 2\int_0^{1/2} \mathrm{sinc}\left(\frac{T_{\mathrm{coh}}}{T_{\mathrm{FFT}}}\kappa\right) \mathrm{sinc}\left(\frac{T_{\mathrm{short}}}{T_{\mathrm{FFT}}}\kappa\right) d\kappa. \tag{42}$$

Hence ($T_{\mathrm{coh}} = 3T_{\mathrm{short}}$), $\langle \Xi^2 \rangle \approx 0.861$ when $T_{\mathrm{FFT}} = T_{\mathrm{coh}}$, the minimal possible by design. More typically, $\langle \Xi^2 \rangle \approx 0.963$ when $T_{\mathrm{FFT}} = 2T_{\mathrm{coh}}$, or ≈ 0.983 when $T_{\mathrm{FFT}} = 3T_{\mathrm{coh}}$. This is sufficient to forego the cost of Dirichlet interpolation in the frequency domain. Any desired improvement in $\langle \Xi^2 \rangle_{\mathrm{resamp}}$ can be obtained by requesting smaller μ_f.

Practical considerations mean that FFT speed is most predictable when N_{FFT} is an integer power of 2. Our resampled time series has a fixed $\delta t'$, so the only way to increase the number of samples is to zero-pad further in time. Starting with the required N_{FFT0},

$$N_{\mathrm{FFT0}} = \frac{\Delta f_{\mathrm{load}}}{\delta \lambda_f} \mathrm{ceil}(\delta \lambda_f \times T_{\mathrm{coh}}), \qquad (43)$$

$$N_{\mathrm{FFT}} = 2^{\mathrm{ceil}(\log_2 N_{\mathrm{FFT0}})}. \qquad (44)$$

In time, $T_{\mathrm{FFT}} = \delta t' N_{\mathrm{FFT}}$. The extension from N_{FFT0} to N_{FFT} causes over-sampling in the frequency domain. From this we decimate by rounding down to the nearest bin with a real-valued ratio ν_R,

$$\nu_R = (\delta \lambda_f)(\delta t') N_{\mathrm{FFT}}, \qquad (45)$$

To maximize recovered power, we use bin-centered frequency. Bin offset ($f_h \approx \bar{f}_h$ in Equation B14) is solved with a shift f_r^* to the nearest FFT bin:

$$\mathrm{remainder}(a, b) \equiv a - \frac{a}{|a|} \mathrm{floor}\left(\frac{|a|}{|b|}\right), \qquad (46)$$

$$f_r^* = \mathrm{remainder}\left(-f_{\mathrm{band}}/2, T_{\mathrm{FFT}}^{-1}\right). \qquad (47)$$

We will multiply a_r and b_r each by $\exp\left(-i2\pi f_r^* \tau\right)$. Preceeding time shifts using f_h remain valid. The smallest FFT frequency f_{FFT}, at k_0, causes the smallest output frequency $f_{\mathrm{min}} = f_h - f_{\mathrm{band}}/2$ to be found at bin k_0:

$$f_{\mathrm{FFT}} = f_h + f_r^* - \frac{1}{2}f_{\mathrm{band}}T_{\mathrm{FFT}}, \qquad (48)$$

$$k_0 = \mathrm{lround}\left(\frac{f_h - f_{\mathrm{band}}/2 - f_{\mathrm{FFT}}}{T_{\mathrm{FFT}}^{-1}}\right), \qquad (49)$$

where the lround function rounds to the integer less than its argument.

3. Antenna function weighting

Equation 34 expresses a discrete time series $x'_r = x'(r\delta t')$ of x' in $\tau = r\delta t'$. Time series accounting for amplitude modulation by antenna functions a and b are returned. The noise-weighted $\sqrt{2/(T_{\mathrm{sft}} S_h)}x'_r$ are multiplied by the noise-normalized $\hat{a}, \hat{b}$ antenna function time series. This hat symbol equals multiplication by $\sqrt{2T_{\mathrm{sft}}/S_h}$.

In the following paragraphs, let us outline some practical considerations, because the implementation may otherwise be ambiguous. When computed, elements a_r, b_r should be normed to order unity for numerical stability [66]. A noise-normalization $\mathcal{S}_a = \sqrt{2T_{\mathrm{sft}}/S_h}\langle a \rangle$ should be used. Multiplication by $\mathcal{S}_a$ can subsequently restore $\hat{a}, \hat{b}$ for the resampled time series. An error-prone point is that we must use a factor of $\mathcal{S}_a \sqrt{T_{\mathrm{short}}/T_{\mathrm{sft}}}$ in the implementation of $\hat{\Gamma}_{KL}^{\mathrm{ave}}$ (because we have written new indices K, L in terms of T_{short}). As the statistic contains factors of a^2, b^2, we track the ratio $T_{\mathrm{short}}/T_{\mathrm{sft}}$. This choice preserves the correct normalization factor and ensures numerical stability at each stage. (A clean-slate code implementation could be more straightforward). The physically-meaningful values a, b remain unchanged throughout.

The product of the normalizations equals $2/S_h$ (for S_h approximated by the nearest SFT). The kernel timestep is $\delta t'$ (in implementation, after the FFT). Multiplication by the requisite frequency shift f_r^* obtains a_r, b_r:

$$a_r \equiv \frac{2\delta t'}{S_h}a(r\delta t')x'(r\delta t')e^{-i2\pi f_r^* \tau}, \qquad (50)$$

$$b_r \equiv \frac{2\delta t'}{S_h}b(r\delta t')x'(r\delta t')e^{-i2\pi f_r^* \tau}. \qquad (51)$$

Here $a(t)$ and $b(t)$ are real-valued amplitude modulations with period of one sidereal day. They are not heterodyned. (Their period is also greater than the maximum Roemer delay, giving $a(r\delta t') \approx a(t(r\delta t'))$, $b(r\delta t') \approx b(t(r\delta t'))$). Antenna functions are effectively constant over T_{sft}. Multiplying $a(t)$ and $b(t)$ by $x(t)$ prepares the optimal filter for the $\mathcal{F}$-statistic [13] as well as for our inner product.

4. Phase shifts after Fourier transform

Subsequent shifts are labeled Φ_{out} and Φ_{in}. Φ_{out_K} is the shift at the physical frequency of bin k, $f = f_h - f_{\mathrm{band}}/2 + k(\delta \lambda_f)$, due to start time (epoch) for that detector's (X for K, Y for L) resampled time series. Φ_{in_K} is the shift at the heterodyned frequency of bin k, $[k_0 + \mathrm{floor}(\nu_R k)]T_{\mathrm{FFT}}^{-1}$, from different start times KT_{sft} within the resampled time series.

$$\Phi_{\mathrm{out}_K}(k) = 2\pi[f_h - f_{\mathrm{band}}/2 + k(\delta \lambda_f)](\tau_X), \qquad (52)$$

$$\Phi_{\mathrm{in}_K}(k) = 2\pi[k_0 + \mathrm{floor}(\nu_R k)]T_{\mathrm{FFT}}^{-1}KT_{\mathrm{short}}. \qquad (53)$$

Considerations include the antenna-weighted, phase-model corrected frequency-domain data $\hat{a}^K \zeta_K$. The ζ_K term equals the product $\Xi_K z_K \exp\left(-i\Phi_K\right)$. This term is explored in Appendix A, Equation A5; in contrast with Equation 17, k refers to a frequency bin, instead of m. In the Appendix, the index $m \equiv j - T_{\mathrm{sft}}/(2\delta t)$ is introduced for a time-domain sample.

Let us now reconstruct $\hat{a}^K\zeta_K$ with resampling: δt becomes $\delta t'$, T_{sft} becomes T_{short}. The index m increases with r. Precisely, $t_m = t_K + m\delta t$ is the overall time, analogous to $r\delta t'$. So m becomes $r - t_K/(\delta t')$.

We look at the time-domain limits of the data $\hat{a}^K\zeta_K$ as defined in Equation A5. The lower limit, $m = -T_{\text{sft}}/(2\delta t')$, becomes $r = (t_K - T_{\text{short}}/2)/(\delta t')$. The upper limit becomes $r = (t_K + T_{\text{short}}/2)/(\delta t')$. We call them (non-integer) $r_{B,K}$ and $r_{U,K}$. The discrete sum must round them. No samples are missed when $r_{U,K} = r_{B,K+1}$. As long as the ideal sample number, $N'_{\text{ideal}} = T_{\text{short}}/(\delta t')$, is $N'_{\text{ideal}} \gg 1$, rounding is tolerable. We will soon replace $r_{U,K}$ with the zero-padded $r_{B,K} + N_{\text{FFT}}$.

The term $a^K x_K(t_m)$ contains $t_m = r\delta t'$. Allowing $r_k \equiv \text{round}(t_K/(\delta t'))$, then $r = m + t_K/(\delta t')$ is simply $r = m + r_K$. So $a^K x_K(t_M)$ translates to a_{m+r_K}. This is the a_r weighted in Equation 50.

Substitute the above into $\hat{a}^K\zeta_K$:

$$\hat{a}^K\zeta_K = \sum_{r=r_{B,K}}^{r=r_{U,K}} a_r e^{-i2\pi f_K r \delta t'}, \tag{54}$$

observing that $\Phi_K = f_K t_K$ (source-frame frequency is constant). Equation 54 foretells a Fourier transform from r into k. Heterodyning has $f_K = f_0 - f_h$, discretely indexed as $k = f_K T_{\text{FFT}}$. Raise $r_{U,K}$ to $r_{B,K} + N_{\text{FFT}}$. Zero-padding (mathematically, using the Heaviside step function H) keeps the sum constant:

$$\hat{a}^K\zeta_K = \sum_{r=r_{B,K}}^{r=r_{B,K}+N_{\text{FFT}}} \frac{H(r_{U,K} - r)a_r}{\exp\left(i2\pi k r/N_{\text{FFT}}\right)}, \tag{55}$$

In practice, an FFT starts at $s = r - r_{B,K}$. Re-indexing,

$$\hat{a}^K\zeta_K = \sum_{s=k}^{N_{\text{FFT}}} \frac{H(r_{U,K} - r_{B,K} - s)a_{s+r_{B,K}}}{\exp\left(i2\pi k[s + r_{B,K}]/N_{\text{FFT}}\right)}, \tag{56}$$

wherein $r_{B,K}$ factors in the kernel:

$$\frac{2\pi k r_{B,K}}{N_{\text{FFT}}} = 2\pi\frac{k_0 + (k - k_0)}{T_{\text{FFT}}}K T_{\text{short}}, \tag{57}$$

expressing k in terms of distance from a minimum k_0. If we pick bins $\bar{k}$ above k_0 at a continuous decimation rate ν_R,

$$\frac{2\pi \bar{k} r_{B,K}}{N_{\text{FFT}}} = \Phi_{\text{in}_K}(\bar{k}). \tag{58}$$

Finally, as in Equation 52, Φ_{out_K} corrects an overall time shift in the resampling epoch, τ_X. When the heterodyning starts at epoch τ_X after reference time τ_0,

$$\frac{2\pi k[s + r_{B,K}]}{N_{\text{FFT}}} = 2\pi\Phi(\tau - \tau_0), \tag{59}$$

$$\Phi(\tau - \tau_0) = f_0\tau H([\tau_0 + \tau_X] - \tau) \tag{60}$$
$$+ (f_0 - f_H)\tau H(\tau - [\tau_0 + \tau_X]),$$

expanding the first Heaviside function into a Boxcar B,

$$H(\tau_0 + \tau_X - \tau) = H(\tau_0 - \tau) + B(\tau_0, \tau_0 + \tau_X), \tag{61}$$

so during τ_0 to τ_X, $f_0\tau_X$ cycles are accumulated, justifying Φ_{out_K}. (The second Heaviside function is null, because $r\delta t'$ starts at τ_X).

5. Frequencies returned from Fourier transform

With a discrete Fourier transform (DFT) from time samples s into frequencies k being the operation $\mathcal{F}_k^s$,

$$\mathcal{F}_k^s y_s = \sum_{s=k}^{N_{\text{FFT}}} e^{-(i2\pi ks/N_{\text{FFT}})}y_s, \tag{62}$$

$$(a^K\zeta_K)_k = e^{-i2\pi[\Phi_{\text{in}_K}+\Phi_{\text{out}_K}](k)}, \tag{63}$$
$$\times \mathcal{F}_k^s\left(H(r_{U,K} - r_{B,K} - s)a_{s+r_{B,K}}\right),$$

DFTs return a frequency *vector* indexed by k, rather than a scalar as in the previous demodulation search [18]. We select the set of frequencies $\bar{k}$. Mathematically, we represent this as a selection function $\delta_{\bar{k}}^k$ that reduces to the Dirichlet delta function when $\nu_R = 1$, so

$$(a^K\zeta_K)_{\bar{k}} = \delta_{\bar{k}}^k(a^K\zeta_K)_k. \tag{64}$$

In the case of $\{L|K\}$, where M_L multiple, often consecutive, T_{short} intervals are present at a single detector, we can do one Fourier transform, because $T_{\text{FFT}} \geq T_{\text{coh}}$. Call the sum $S_{\bar{k}}^{(L|K)}$:

$$S_{\bar{k}}^{(L|K)} = \delta_{\bar{k}}^k e^{-i2\pi[\Phi_{\text{in}_{L_0}}+\Phi_{\text{out}_{L_0}}](k)} \tag{65}$$
$$\times \mathcal{F}_k^s\left(H(r_{U,(L_0+M_L)} - r_{B,L_0} - s)a_{s+r_{B,L_0}}\right),$$

so the whole sum can be done in a single FFT. This is because Φ_{out_L} depends only on L for its detector time epoch (τ_Y), and Φ_{out_L} is proportional to LT_{short}, which is absorbed into the Fourier transform kernel.

If $S_{(L|K)}$ skips some term, *e.g.*, auto-correlation where $L = K$ in the same detector, this is handled, both in theory (by subtracting a Boxcar function) and practice (by skipping that time and putting the next T_{short} at the following place in the zero-padded time series). Segment L depends implicitly on its detector Y.

6. Statistic in resampled data and physical meaning

Taking a look at ρ from Equation 26, we see it can be phrased in terms of $\hat{a}^K \zeta_K$ in the Appendix A, Equation A2. We will break it into explicit pairs over Q detectors (first cross-correlation pair component indexed by X, second by Y), each of which has $M \equiv T_{\rm obs}/T_{\rm short}$ (zero-padded gapless) data segments, as in Section III B 1. A data segment index K for the first component of a pair is matched by M_L terms of the second component, starting from L_0. Eliding b terms,

$$\rho = \frac{N}{5}\Re \sum_{X=0}^{Q}\sum_{K=0}^{M}\hat{a}^K\zeta_K^* \sum_{Y=0}^{Q}\sum_{L=L_0}^{M_L}\hat{a}^L\zeta_L + \ldots, \quad (66)$$

so we insert the Fourier transforms to get the vector $\rho_{\bar{k}}$,

$$\rho_{\bar{k}} = \frac{N}{5}\Re \sum_{X=0}^{Q}\sum_{K=0}^{M}(\hat{a}^K\zeta_K)_{\bar{k}}^* \sum_{Y=0}^{Q} S_{\bar{k}}^{(L|K)} + \ldots, \quad (67)$$

A commonly-used projection of the in-phase data onto $a(t)$, the sub-interval integral F_{a_I} (as in Appendix A, Equation A8) motivates us to name the key quantities:

$$\bar{F}_{a_{K,\bar{k}}} = (\hat{a}^K\zeta_K)_{\bar{k}}, \quad (68)$$

$$\bar{F}_{a_{L,\bar{k}}} = S_{\bar{k}}^{(L|K)}. \quad (69)$$

Note: unlike terms F_a and F_b in Appendix A, the above quantities include noise normalization. Overall normalizations $\hat{A}_\mathcal{P}^2$, $\hat{B}_\mathcal{P}^2$, $\hat{C}_\mathcal{P}^2$ are the sums over pairs of $(\hat{a}^K\hat{a}^L)^2$, $(\hat{b}^K\hat{b}^L)^2$, and $(\hat{a}^K\hat{b}^L\hat{a}^L\hat{b}^K)$, respectively. Then the resampled ρ statistic parallels Equation A15:

$$\rho_{\bar{k}} = \frac{\sqrt{2}}{\sqrt{\hat{A}_\mathcal{P}^2 + 2\hat{C}_\mathcal{P}^2 + \hat{B}_\mathcal{P}^2}} \times \ldots \quad (70)$$

$$\Re \sum_{X=0}^{Q}\sum_{K=0}^{M}\sum_{Y=0}^{Q}\left[\bar{F}_{a_{K,\bar{k}}}\bar{F}_{a_{L,\bar{k}}} + \bar{F}_{b_{K,\bar{k}}}\bar{F}_{b_{L,\bar{k}}}\right].$$

Equation 70 holds in any reference frame. Dependence on detector and source motion has been absorbed by resampling, so the remaining formula is manifestly invariant. This formula for ρ is a semicoherent matched filter assuming a sinusoidal waveform. In the (non-physical) case of zero Roemer delay, frequency is constant and no resampling is needed, so Equation 70 exactly equals Equation A15. Resampling is elegantly interpreted as a shift to a frame with zero Roemer delay, where the frequency is effectively constant (up to the accuracy of the resampling parameters and numerical precision). It is unsurprising but reassuring that the result is independent of the original frame.

7. Summary of resampling implementation

Resampling has been ported from the $\mathcal{F}$-statistic computation into the cross-correlation method. The implementation differs in that $\mathcal{F}$ needs no concept of $T_{\rm short}$: its coherence time is the FFT time, and because each segment is resampled individually without being subdivided into pairs, $\Phi_{\rm in} = 0$. The $\mathcal{F}$-statistic includes autocorrelation, and there is no extra overlap. (Some inefficiency in recalculating the same overlapping pairs in the cross-correlation method could be reduced by caching partial terms F_a, F_b from Appendix A).

For the $\mathcal{F}$-statistic, resampling has already accelerated long $T_{\rm coh}$ searches. Resampling should also speed-up the cross-correlation method. Considering Equation 23, we have offloaded phase-correction from the $\mathbf{W}$ matrix onto the $\mathbf{z}$ vector, turning a quadratic operation into a linear one. That the remaining matrix can be evaluated by an FFT is a further improvement. In the next section, we measure computational speed and sensitivity.

IV. COMPUTATIONAL COST AND SENSITIVITY

The computational speed and cost of resampling for the cross-correlation method is to be measured. A first comparison (Figure 1) takes overall run times of the demodulation and resampling techniques for a given number of templates. The relative speed-up, in Figure 2, governs how much can be re-invested in search depth. Deeper understanding helps predict the computational cost in time required for conceivable use cases: the *timing model*.

A. Demodulation timing model

First, define the timing model for the demodulation search. Let each dimension have spacing $\delta\lambda$ determined by the metric, requiring N_λ templates be searched in each dimension to cover a range $\Delta\lambda = N_\lambda \delta\lambda$. Using a simple cubic lattice,

$$N_{\rm template} = \prod_\lambda \frac{\Delta\lambda}{\delta\lambda}. \quad (71)$$

Take a test case for a single point in orbital parameter space. With $n_{\rm bin} = 2$ Dirichlet interpolation bins, $N_{\rm template} = 55488$, $T_{\rm max} = 22800$ s, $T_{\rm obs} = 3.0 \times 10^6$ s, $T_{\rm sft} = 1440$ s ($N_{\rm det} = 2$) this case is measured to take a total time of $T_{\rm demod} = 159.80$ s. (Single-threaded without SIMD instructions on an Intel Core i7-4980HQ at 2.8 GHz). Normalizing these parameters into a single timing constant, $\tau_{\rm demod}$ for two detectors, and with scalings taken from [18], we have a timing function,
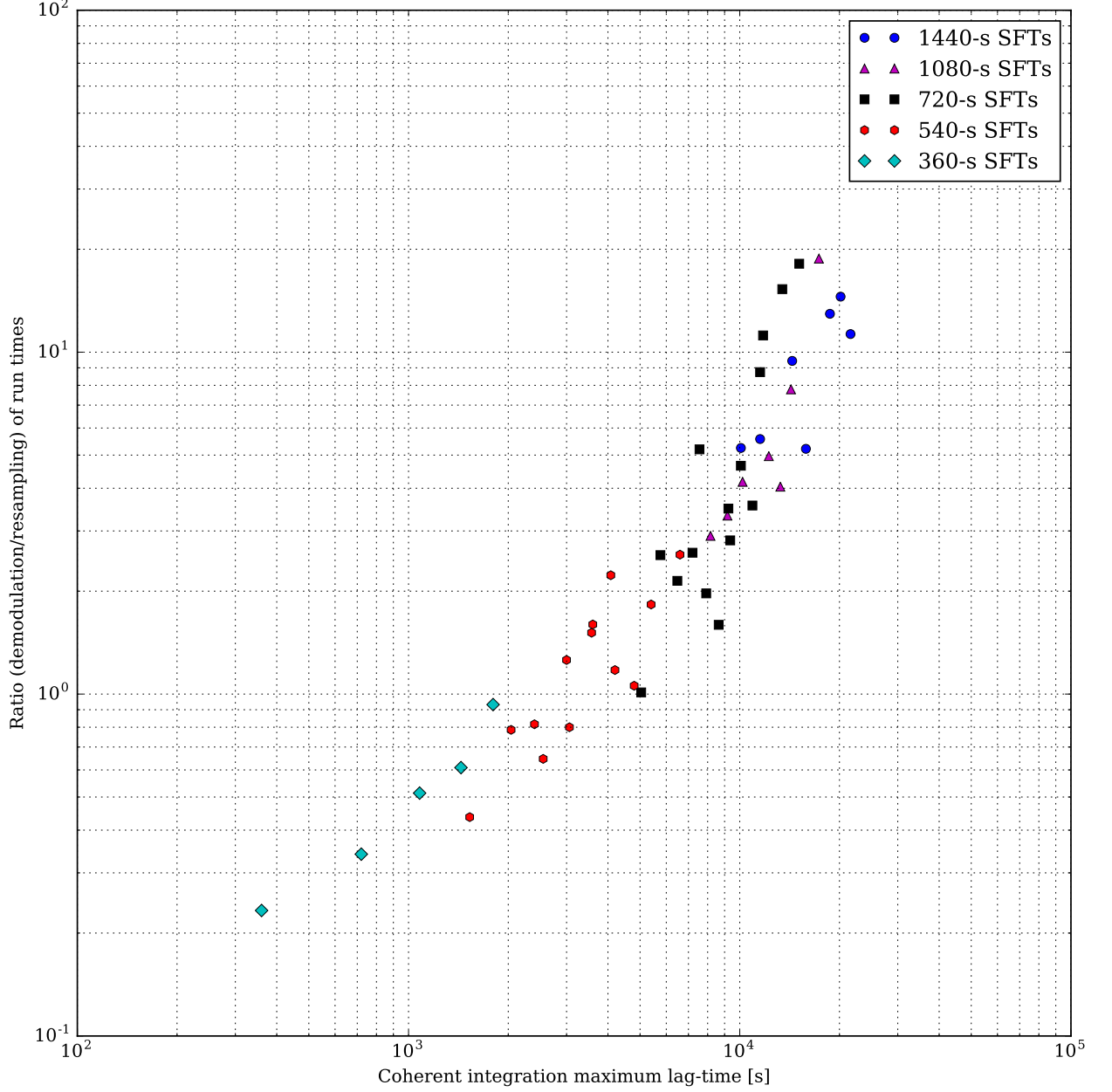
FIG. 2. Ratio of demodulation run time to resampling run time *vs* maximum lag-time $T_{\max}$ for representative T_{sft}. At longer $T_{\max}$, the relative advantage of resampling grows. It is roughly 20 for the longest typical O1 set-ups, resulting from both the ability to pair independently of T_{sft} (using T_{short}) and from the FFT yielding ρ as a function of f_0 for a given set of binary orbital parameters. Even longer $T_{\max}$ are attainable because of the asymptotic metric of orbital parameter space. As in Figure 1, Doppler wings large in proportion to f_{band} reduce resampling efficiency; these tests use 0.050-Hz f_{band}.

$$N_{\text{pairs}} \approx \frac{N_{\text{det}}(N_{\text{det}}+1)}{2} T_{\max} T_{\text{obs}} T_{\text{sft}}^{-2}, \qquad (72)$$

$$T_{\text{demod}} = \tau_{\text{demod}} n_{\text{bin}} N_{\text{template}} N_{\text{pairs}}. \qquad (73)$$

Using this measurement, τ_{demod} is about 1.5×10^{-8} s.

Note that this single measurement is based on gapless data. In the presence of gaps, the demodulation search can easily skip to the next SFT (at present, resampling

cannot skip gaps).

Template count N_{template} depends on every parameter's $\delta\lambda$. Because $\delta\lambda$ depends on T_{max} for all four Doppler parameters, the computational cost increases with longer lag-time. Each $\delta\lambda$ is proportional to the inverse square root of the corresponding metric element $g_{\lambda\lambda}$ as in Equation 36. Whelan *et al* [18] note that the metric element g_{ff} increases with T_{max}^2, while the orbital parameter elements also increase as T_{max}^2 for $T_{\text{max}} \ll P_{\text{orb}}$ before asymptoting as T_{max} approaches the P_{orb}. Uncertainty in P_{orb} is low enough that a single template is enough to cover it for short T_{max}, but not generally at high T_{max}. So the computational cost scaling for demodulation has $1 + 2 + 1 = 4$ powers of T_{max}: it is $T_{\text{demod}} \propto T_{\text{max}}^4$ for short lag-time. After the orbital period resolves and also asymptotes for long lag-time, the scaling is $\propto T_{\text{max}}^2$, with a larger coefficient. Contrast this case with resampling.

B. Resampling timing model

Better *scaling* is sought from the resampling timing function. Longer lag-times are theoretically easier to achieve with resampling. It is the measurements of the coefficients that determine whether the overall computational cost is affordable.

The resampling timing function is complicated: it involves three timing constants. Table II lists these constants. First is the timing constant $\tau_{0,\text{CCbin}}$ for per-template (per-bin) operations, such as multiplying, adding, copying, and phase-shifting results to and from the FFT. Second is the timing constant $\tau_{0,\text{bary}}$: the cost for barycentering for each point in orbital parameter space. Third and last is the timing $\tau_{0,\text{FFT}}$: the cost of the FFT operation (using the *FFTW* library) for each template. This division into three parts is motivated by a pre-existing timing model for the $\mathcal{F}$-statistic [65].

The τ constants are measured using Atlas, the cluster at AEI Hannover, Germany. A typical cluster node uses an Intel Xeon E3-1220v3 at 3.1 GHz; a smaller set of E3-1231v3 (3.4 GHz) and E5-1650v2 (3.5 GHz) CPUs are also in use. Approximately 120 configurations, varying frequency band (f_{band}), observing time (T_{obs}), lag-time (T_{max}), number of observatories (N_{det}), starting frequency ($f_h - f_{\text{band}}/2$), projected semi-major axis (a_p), allowed frequency mismatch (μ_f) in the statistic, and number of Dirichlet kernel terms (D), are tested and fit to the three search parameters. This fit minimizes the discrepancy between predicted and measured time, as shown in Figure 3.

Time T_{resamp} is predicted as follows. It is most efficient to take $T_{\text{max}} = T_{\text{short}}$. We divide the analysis into bands of Δf_{load} (Equation B5). We next separate $N_{\text{template}} = N_{\text{orb}} N_f$, into orbital N_{orb} and frequency N_f template counts. The FFT size is N_{FFT} (by Equation 44). A 'triangular' function accounts for detector pairings,
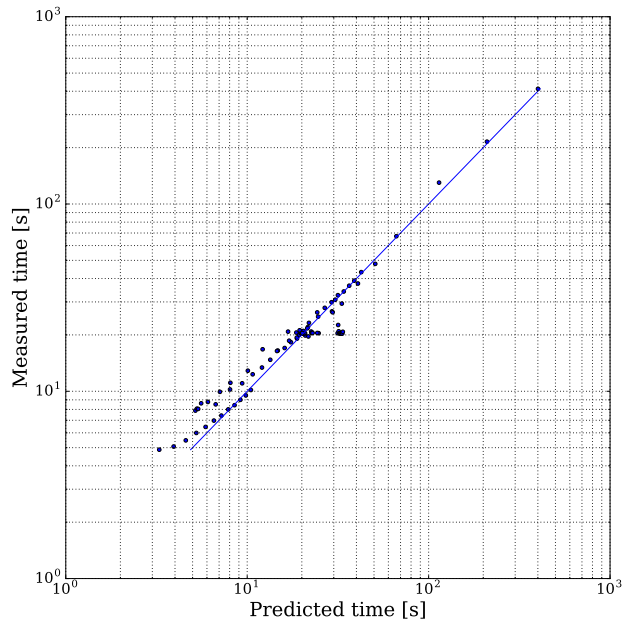
FIG. 3. Resampling timing model. Measured *vs* predicted run time for (about 120) use cases, varying frequency band f_{band}, observing time T_{obs}, lag-time T_{max}, number of observatories N_{det}, starting frequency f_{min}, projected semi-major axis a_p, allowed frequency mismatch μ_f in the statistic, and number of Dirichlet kernel terms D for resampling. This set captures the *low* N_{FFT} case, where the threshold for *high* N_{FFT} is 2^{18} bins: for more bins, the cost per FFT bin ($\tau_{(0,\text{FFT})}$) is approximately 3 times greater (see Table II). The diagonal line marks an exact prediction of run time of the resampling code. A fit is made using the three resampling timing constants, $\tau_{(0,\text{CCbin})}$, $\tau_{(0,\text{bary})}$, and $\tau_{(0,\text{FFT})}$, of the timing model in Equation 76. Measurement done on Atlas cluster and may vary due to machines under realistic use conditions.

$$\text{triang}(N) = 1 + \frac{N+1}{2}. \tag{74}$$

Taking a prefactor of 5 for the FFT logarithmic term is based on [65], from which the basic scheme of our model is motivated. Absorbing a typical number, $D = 8$, into $\tau_{0,\text{bary}}$, is efficient. The total time is then T_{resamp}:

$$
\begin{aligned}
T_{\text{resamp}} = N_{\text{orb}} N_{\text{det}} (T_{\text{obs}}/T_{\text{max}}) [\ldots \qquad (75) \\
\tau_{0,\text{CCbin}} N_f \text{triang}(N_{\text{det}}) + \ldots \\
\tau_{0,\text{bary}} (2\Delta f_{\text{load}} \times T_{\text{max}} \times (D/8)) + \ldots \\
\tau_{0,\text{FFT}} N_{\text{FFT}} \times 5\log_2(N_{\text{FFT}}) \times \text{triang}(N_{\text{det}})]
\end{aligned}
$$

Observe that N_{FFT} is proportional, albeit through power-of-two steps, to N_f, and N_f is proportional to T_{max} as before. At low lag-time, $N_{\text{orb}} = 2$, so the resampling time scales $T_{\text{resamp}} \propto T_{\text{max}}^2 \log T_{\text{max}}$. At high lag-time, after the number of orbital templates has asymptoted and period dimension resolved, it is, with a larger

Coefficient	Low N_{FFT} value [s]	Low N_{FFT} uncertainty [s]	High N_{FFT} value [s]	High N_{FFT} uncertainty [s]
$\tau_{0,\mathrm{CCbin}}$	1.01×10^{-7}	$\pm 1.10 \times 10^{-8}$	1.34×10^{-7}	$\pm 2.25 \times 10^{-8}$
$\tau_{0,\mathrm{bary}}$	1.62×10^{-7}	$\pm 7.48 \times 10^{-10}$	1.62×10^{-7}	$\pm 3.20 \times 10^{-9}$
$\tau_{0,\mathrm{FFT}}$	5.27×10^{-10}	$\pm 5.19 \times 10^{-11}$	1.40×10^{-9}	$\pm 6.00 \times 10^{-11}$

TABLE II. Table of timing coefficients. Higher values indicate greater computational cost. Values obtained from fit to overall external total run time (see Figure 3). Timing coefficients are divided into low and high N_{FFT} values, with the threshold being $N_{\mathrm{FFT}} = 2^{18}$, above which $\tau_{0,FFT}$ is about $3\times$ higher. This difference is expected from the *FFTW* library performance profile and may arise from cache sizes. Uncertainty reported is $\pm 1\sigma$ to the fit. About 120 measurements each done for low and high N_{FFT} on the Atlas cluster, using a mix of Intel Xeon E3-1220v3 and E3-1231v3 processors. Results accord with single-processor *Callgrind* performance profiling, but future precise internal per-function measurements may be valuable.

coefficient, $T_{\mathrm{resamp}} \propto \log T_{\mathrm{max}}$. The improvement stems from two parts of the new code: the 'SFT gain' by reducing the number of pairs saves a factor of T_{max}, and the 'FFT gain' by converting the **W** weights matrix into an FFT operator effects $T_{\mathrm{max}}^2 \rightarrow T_{\mathrm{max}} \log T_{\mathrm{max}}$.

Caveats: the *FFTW* functions for FFTs alert us to a $3\times$ increase in $\tau_{0,\mathrm{FFT}}$ for N_{FFT} above about 2^{18}. This behavior is observed and is why Table II is divided into low and high N_{FFT} sections. Our prediction for T_{resamp} applies a factor of 3 multiplier when N_{FFT} is predicted to be in this slow regime. A key *caveat* is that the precise N_{FFT} is difficult to calculate *a priori*. (The *post hoc* N_{FFT} is used to make τ estimates more accurate). This difficulty comes from the metric calculation depending on the true phase derivatives instead of a simpler diagonal approximation (as explained in [18]). Slight misprediction in metric-derived spacing can be amplified by power-of-2 rounding in N_{FFT}. Future improvement in T_{resamp} estimation can be expected from reusing the exact code used for metric calculation in the timing predictor.

C. Sensitivity of optimized set-up

Sensitivity depth D^C for the semi-coherent cross-correlation method search scales $T_{\mathrm{max}}^{1/4}$ [18], up to an uncertain time where spin-wandering makes longer integration incoherent. The demodulation technique gives an effective scaling of $D^C \propto (T_{\mathrm{demod}})^{1/16}$ for low lag-time T_{max}, compared to P_{orb}, or $\propto (T_{\mathrm{demod}})^{1/8}$ for high lag-time. Resampling, dropping the logarithmic term, offers $D^C \propto T_{\mathrm{resamp}}^{1/8}$ for low lag-time or $D^C \approx$ (constant) for high.

Once the computational cost reaches the orbital parameter metric plateau and asymptotes, additional sensitivity is nearly cost-free with resampling. Surprisingly, in the frequency dimension, the number of templates continues to increase $\propto T_{\mathrm{max}}$, but because $T_{\mathrm{short}} = T_{\mathrm{max}}$, the number of semicoherent segments decreases linearly as T_{max} increases, so there are longer but fewer FFTs to do. Small cost increases do continue, in the logarithmic FFT term. Two caveats: the number of period templates still depends on T_{obs}, and power-law scalings assume a large number of semicoherent segments. The conceivable case of $T_{\mathrm{max}} = 10$ days, $T_{\mathrm{obs}} = 3$ months may be close to the limit where this approximation holds, and excluding the auto-correlation means that the ratios of $T_{\mathrm{obs}}/T_{\mathrm{max}} < 5$ (approximately) may exclude some data. (The latter is partly-solvable by decoupling T_{short} from T_{max}). Nevertheless, the ease of high T_{max} with resampling helps both future searches and follow-ups.

Gains in search sensitivity depend on the measured timing constants. We iteratively estimate the maximum T_{max} possible with the resampling code for the same computing resources made available, in a given band, as to the demodulation O1 search [3]. For future searches, the distribution across bands can be re-allocated to maximize detection probability.

For now, we consult Figures 4 and 5. These figures show, using the $D^C \propto T_{\mathrm{max}}^{1/4}$ assumption, the forecast sensitivity gain from resampling's speed-up relative to demodulation. The exact same test set-ups are run for both resampling and demodulation and the run time is measured. Then Equation 76 is used to predict the run time of resampling with longer T_{max}, iteratively increasing T_{max} by 1% until the original demodulation cost is predicted to be reached. The quarter root of the final T_{max} is taken as the forecast gain. (If resampling already takes at least as long as demodulation for a given set-up, this gain defaults to unity). Gains depend on the test bands' set-ups (Table III). Figure 4's right side contrasts empirical sensitivity gains with predictions. The actual relative gain, given by the square root of the ratio of the ρ statistic with given T_{max}, tends to be less than the power-law prediction. Sensitivity forecasts in Figures 6 and 7 should thus be read as cautiously optimistic.

Long lag-times benefit the most from resampling. Figure 2 illustrates that resampling is only faster than demodulation for bands of $T_{\mathrm{max}} \gtrsim 2000$ to 5000 s, which Table III shows to be in frequency bands less than roughly 200 to 400 Hz in the O1 setup. These T_{max} allocations [3] were designed to maximize detection probability by investing integration time in high-probability regions of orbital parameter space and frequencies near the torque balance level. Where T_{max} is already large, resampling offers more acceleration, thus more computing to be reinvested, and Figures 4 and 5 show bigger gains. In principle, the cost allocation is a global problem: we want to

min f_0 [Hz]	max f_0 [Hz]	max $T_{\max}$ [s]	min $T_{\max}$ [s]	$f_{\text{band}}[Hz]$	T_{sft} [s]
25	50	25920	10080	0.050	1440
50	100	19380	8160	0.050	1080
100	150	15120	6720	0.050	720
150	200	11520	5040	0.050	720
200	300	6600	2400	0.050	540
300	400	4080	1530	0.050	540
400	600	1800	360	0.050	360
600	800	720	360	0.050	360
800	1200	300	300	0.050	300
1200	2000	240	240	0.050	240

TABLE III. O1 search set-up [3]. Set-up depends on the GW frequency, f_0, of a search band, as well as its position in orbital parameter space. More likely regions in parameter space are allocated longer $T_{\max}$ to increase detection probability. All search bands in O1 are 50 mHz wide in f_{band}. The SFT duration T_{sft} varies with frequency to limit spectral leakage from orbital acceleration. To find the set-up for a given point in parameter space, find the line bounding f_0 between min f_0 and max f_0 columns and consult [3] to determine its placement in orbital parameters. This set-up offers significant potential for re-optimization [67].

maximize the detection probability of the entire search, not one band. This problem has been addressed not only in [3] but also [67]. In the future, these methods can be turned to the complicated task of re-optimizing the resampling cost allocation to maximize detection probability. For this paper, forecasts are based on the O1 allocation. Also note that we assume that the sensitivity gains $\propto T_{\max}$ will uniformly scale the detection efficiency curves that set upper limits. Taking this product of averages is only approximate: the true sensitivity is an average constructed from the products of gains in each band. As the ρ statistic ratio from long $T_{\max}$ is less than predicted, a systematic study is needed about the sensitivity gain from computational reinvestment. In the future, we expect our assumptions to be tested by a second Mock Data Challenge (following [19]).

At present, results are suggestive. Figure 6 shows the projected upper limits that are forecast based on O1 results [3], divided by the sensitivity gain estimated for each band. Figure 7 shows these upper limits divided by the noise ASD of the detector, to show sensitivity depth D^C, which is easier to compare with other methods. Both figures refer to results marginalized over $\cos \iota$, as the inclination angle of Sco X-1 is unknown. Long $T_{\max}$ bands at low frequencies can potentially double to triple in sensitivity. Given equal cost allowance and the assumption of $T_{\max}$ limited to 3 days by spin-wandering, the gain is limited: from 20 to 125 Hz, the median gain is 51%, and from 20 to 250 Hz, it is only 11%, with minimal benefit at higher frequencies. The sensitivity depth varies between the mid-30s and mid-60s Hz$^{-1/2}$, depending on position in orbital parameter space. Given tenfold resources and the assumption of $T_{\max}$ limited to 10 days, the gains are respectively 2.83× and 2.75× over O1. This sensitivity depth is approximately 100 Hz$^{-1/2}$. Given O1 noise, the latter scenario would just touch the torque-balance level

at 100 Hz. Given twofold detector improvement, the upper limits would scale linearly, and resampling could potentially reach below torque balance from approximately 40 to 140 Hz. Longer observing runs should improve sensitivity with the usual $T_{\text{obs}}^{1/4}$ scaling [18].

Future computational enhancements in the cross-correlation method, such as GPU acceleration for the barycentering and FFT operations, may make the tenfold gain in cost allowance realistic, as may access to larger computing resources. For example, one *Einstein@home* Month (EM) of computing power assumes 12 thousand cores [67], or roughly 8.64 million CPU hours. Depending on CPU performance compared to the Atlas cluster, multi-EM allocations could extend the cross-correlation method's depth. It may be possible to use Bessel functions, as in [68], or a loosely-coherent approach [39], to accelerate moving through the orbital parameter space: the phase modulation can in principle be 'resampled' in the frequency domain as well as our time-domain approach, and some fusion of the two may be faster. Even now, resampling can accelerate longer lag-time follow-ups (progressive 4× increases in $T_{\max}$ for search candidates [3]) and improve the low-frequency search.

The 'CrossCorr' cross-correlation method is not the only method that may reach such performance. The 'Sideband' method [44] is under active development, and a binary-oriented, resampled $\mathcal{F}$-statistic code [34] has offered even greater sensitivity depth. The latter predicts that torque-balance could be reach up to $160 - 200$ Hz for conservative assumptions about eccentricity or $500 - 600$ Hz if eccentricity is assumed to be well-constrained. (By assuming the eccentricity to be circular, our result of 140 Hz is comparable to the latter case). Predictions are highly sensitive to the timing function and cost allowances of the final code, as well as to assumptions about spin-wandering. Here we have

presented our estimates based on working search code and extrapolations from the finished O1 search using the cross-correlation method.

V. CONCLUSIONS

Resampling accelerates the deepest current search for Sco X-1 and similar LMXBs, the cross-correlation method [3]. By calculating the cross-correlation method's ρ statistic using barycentric interpolation to the source frame, followed by an FFT, speed-up is possible for long coherent integration lag-times. Because of the plateauing of the binary orbital parameter space, this acceleration can drive the cross-correlation method's forecast sensitivity to the torque-balance level in conceivable scenarios. In the most optimistic case with O1-like data, it may graze this level at 100 Hz; with a detector twice as sensitive (closer to Advanced LIGO design sensitivity), this range may extend from 40 to 140 Hz. Re-optimization of the computational cost distribution across parameter space [67] can focus resources where detection is most probable. Reaching torque-balance might then be possible without large increases in computing power. Future improvement may allow it to compete up to higher frequencies, as might other proposed methods [34]. The cross-correlation method with resampling works already. This success is possible thanks to the deep similarity between the $\mathcal{F}$-statistic and ρ-statistic and the shared codebase of the *LIGO Applications Library*, which allowed the importation of large portions of the resampling algorithm, once the mathematics were understood. Future improvements to any of this family of methods might be transplanted to benefit all.

Many unknowns remain in Sco X-1. The depicted torque-balance level assumes a 10-km radius and 1.4-solar mass for a NS that itself has not been confirmed in the system; the level varies with the object's moment of inertia. Expectation has held that Sco X-1's luminosity makes it a promising target. Other systems may prove promising alternative targets, particularly if they have a known spin frequency. Known frequency, or much more precise orbital parameters, could reduce the cost of the cross-correlation method and similar semicoherent searches by many orders of magnitude. Then a sensitivity limited only by spin-wandering might be easily reached, regardless of location on the spectrum. Until then, computational optimizations will play a pivotal role in broadband searches. We see potential in applying this proven method to Advanced LIGO searches – gravitational waves from Sco X-1 have never been closer to detection.

Appendix A: Relationships to other optimal statistics

Terms called F_a and F_b [16] relate ρ to the $\mathcal{F}$-statistic, already amenable to resampling [15]. These F_a and F_b are the components of the statistic that are respectively projections of data along the a and b time series. To investigate these components, we will look at the phase-model corrected frequency-domain data, ζ_K. (Precisely, $\zeta_K = \Xi_K z_K \exp\left(-i\Phi_K\right)$ for z_K, Ξ_K from Equation 22). We can arrange the data z_{Kk}, indexed by frequency bin k, to include phase shift $\exp\left(-\mathrm{i}\Phi_K\right)$,

$$\zeta_K \equiv \sum_{k \in \mathcal{K}_K} (\mathrm{i})^{2k}\mathrm{sinc}(\kappa_{Kk})z_{Kk}e^{-\mathrm{i}\Phi_K}, \qquad \text{(A1)}$$

and likewise ζ_L, substituting (real-valued) Equation 24 into ρ ($\Re$ denoting the real part) and grouping terms:

$$\rho = \frac{N}{5}\Re \sum_{KL \in \mathcal{P}} \left[(\hat{a}^K \zeta_K)^* \hat{a}^L \zeta_L + (\hat{b}^K \zeta_K)^* \hat{b}^L \zeta_L\right], \text{(A2)}$$

which merits inspection of $\hat{a}^K \zeta_K$. Insert $\hat{a}^K$ and Equation 17, noting $(-1)^{k-l} = (-1)^{l-k}$, $(\mathrm{i})^{2k} = \exp(\mathrm{i}\pi k)$:

$$\hat{a}^K \zeta_K = \frac{2}{S_K}a^K \sum_{k \in \mathcal{K}_K} \mathrm{sinc}(\kappa_{Kk}) \qquad \text{(A3)}$$
$$\times \sum_{j=0}^{N-1} x_K(t_K - T_{\mathrm{sft}}/2 + j\delta t)$$
$$\times e^{-\mathrm{i}2\pi k(j\delta t - T_{\mathrm{sft}}/2)/T_{\mathrm{sft}}}\delta t e^{-\mathrm{i}\Phi_K},$$
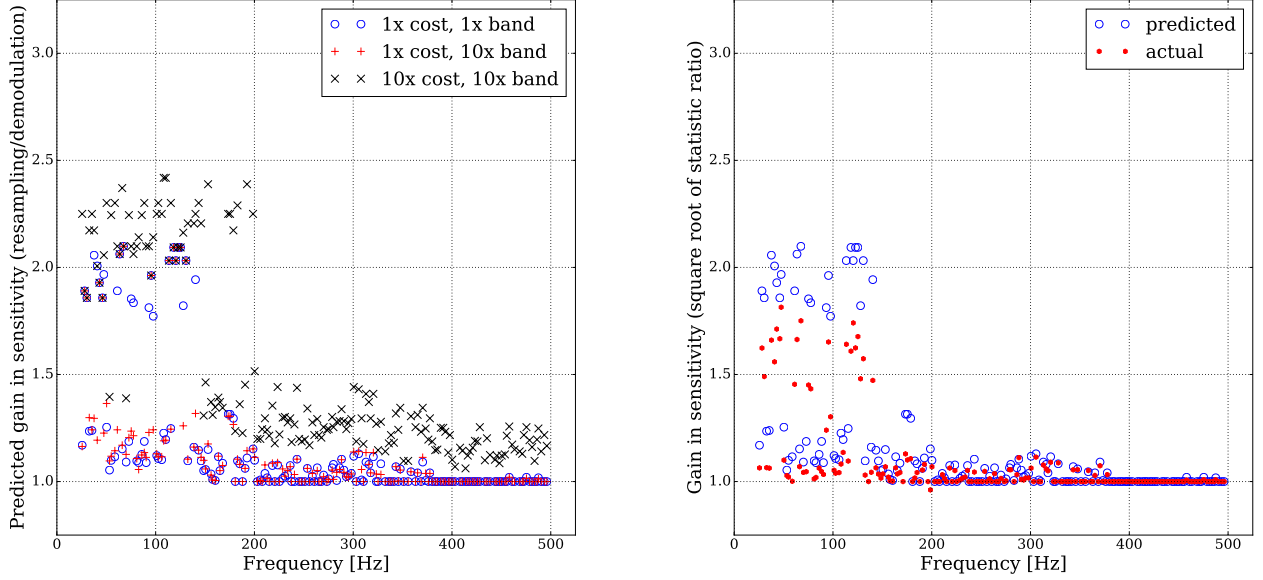
FIG. 4. *(Left)* Predicted gain in sensitivity for resampling over demodulation, *vs* frequency, based on h_0 upper limits being proportional to $T_{\max}^{-1/4}$. $T_{\max}$ is capped at 3 days in this Figure; compare Figure 5. An observation time of $T_{\mathrm{obs}} = 1.5 \times 10^7$ s is assumed. The timing model (Equation 76) estimates cost for incrementally-longer $T_{\max}$ until constraints reached. Our cost allowance is predicated on the measured ratio of resampling to demodulation times on the Atlas cluster. Whenever resampling is slower, the result defaults to 1. Symbols on the figure indicate the following constraints: ($\circ$) equal computational cost, ($+$) f_{band} 10× wider, ($\times$) f_{band} 10× wider and given 10× computing time. No other re-optimization of set-up is done. Fluctation in the results occurs, because benefit scales non-linearly with increased $T_{\max}$; less-probable regions of orbital parameter space are allocated lower $T_{\max}$ and see less benefit at fixed cost. A distinct high-gain population is seen where benefits are limited by spin wandering, computing cost having asymptoted in the binary metric. The ($+$) set is worse for lower frequencies because some bands move into high N_{fft}, but mediun frequencies benefit from reduced Doppler wings. The ($\times$) set shows improvement up to about 500 Hz, with more relative gain because Table III set-ups allocate shorter lag-time to those frequencies, so affording greater room for improvement. For equal cost ($\circ$), median gain from 20 Hz to 125 Hz is 51% and from 20 Hz to 250 Hz is 11%. *(Right)* Empirical results of simulation for 3-day limit, equal-cost, equal-band. *(Red hexagons)* show the square-root of the ratio of ρ (resampling) divided by ρ (demodulation). Improvement exists but is less than predicted, possibly because $h_0 \propto \rho^2 \propto T_{\max}^4$ scaling laws are imprecise.

Whereas $a(t)$, $b(t)$ amplitude modulations have a period on the order of a sidereal day, as in Section III B 3, the a^K, S_K terms vary much more slowly than f_K, So we take $m \equiv j - T_{\mathrm{sft}}/(2\delta t)$, $t_j \equiv t_K - T_{\mathrm{sft}}/2 + j\delta t$, $t_m \equiv t_K + m\delta t$, moving the antenna functions inside the sum over m,

$$\hat{a}^K \zeta_K = \sum_{m=-T_{\mathrm{sft}}/(2\delta t)}^{m=T_{\mathrm{sft}}/(2\delta t)-1} \delta t \frac{2}{S_K} a^K x_K(t_m) \tag{A4}$$
$$\times \sum_{k \in \mathcal{K}_K} \mathrm{sinc}(\kappa_{Kk}) e^{-\mathrm{i}(2\pi km\delta t/T_{\mathrm{sft}} + \Phi_K)}.$$

Instead of including all frequency bins k for z_K of Equation 22, the SFT signal-resolution can be zero-padded (see Equation 9.3 of [62]). Zero-padding k brings the nearest bin closer to f_K, approaching $\kappa_{Kk} \approx 0$,

$$\hat{a}^K \zeta_K \approx \sum_{m=-T_{\mathrm{sft}}/(2\delta t)}^{m=T_{\mathrm{sft}}/(2\delta t)-1} \delta t \frac{2}{S_K} a^K x_K(t_m) \tag{A5}$$
$$\times e^{-\mathrm{i}(2\pi(f_K T_{\mathrm{sft}})m\delta t/T_{\mathrm{sft}} + \Phi_K)}.$$

1. Statistic in conventional quantities

Proceeding to continuous time, Equation 5.10 of [16] has sub-interval integral F_{a_I},

$$F_{a_I} = \int_{T_I - \Delta T/2}^{T_I + \Delta T/2} a(t) x(t) e^{-\mathrm{i}\varphi(t)} dt, \tag{A6}$$

$$F_a = \sum_I F_{a_I}. \tag{A7}$$

Treating sums as integrals, $\delta t \rightarrow dt$, $T_{\mathrm{sft}} \rightarrow \Delta T$, $a_K \rightarrow a(t)$, $x_K(t_m) \rightarrow x(t)$, and
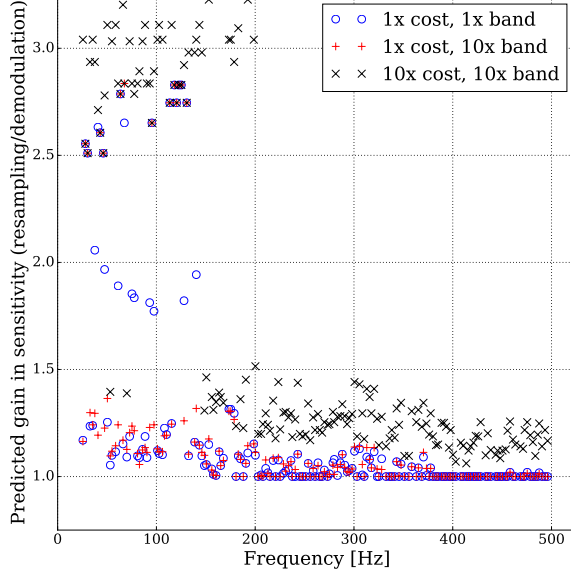
FIG. 5. Predicted gain in sensitivity for resampling over demodulation, *vs* frequency. $T_{\max}$ capped at 10 days; compare Figure 4. Bands limited by computational cost rather than spin-wandering see no change. In this optimistic scenario, the median gain from the (×) 10× cost, 10× f_{band} case, from 20 Hz to 125 Hz, is 2.83×, and from 20 Hz to 250 Hz it is 2.75×. This is the best improvement that we consider. As in Figure 4, actual gains may be less.

$\exp\left(-\mathrm{i}[2\pi(f_K T_{\text{sft}})m\delta t/T_{\text{sft}} + \Phi_K]\right) \rightarrow \exp\left(-\mathrm{i}\varphi(t)\right)$. Observe that with $2\pi f_K = (d\Phi_K/dt|t = t^K)$, $t = m\delta t$, φ is a Taylor approximation of Φ:

$$\hat{a}^K \zeta_K = \frac{2}{S_K} F_{a_K}. \qquad (A8)$$

Taking $T_0 \rightarrow T_{\text{sft}}$ in Equation 42 of [13], we write an inner product,

$$(x||y) \equiv \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} x(t)y(t)dt, \qquad (A9)$$

so $F_{a_K} = [T_{\text{sft}}/2](a \cdot x || \exp(-\mathrm{i}\varphi))$ can be viewed as a projection of the amplitude-modulated data onto the phase-model basis. A reader may wonder whether this is not a Fourier transform. Not quite: $\varphi(t)$ is phase-modulated and does not increase linearly in evenly-sampled detector time $m\delta t$. Before addressing this problem with resampling (Section III), we connect ρ to related statistics.

Using F_{a_K} in Equation A2,

$$\rho = \frac{N}{5} 4\Re \sum_{KL \in \mathcal{P}} \left[\frac{F_{a_K}^*}{S_K} \frac{F_{a_L}}{S_L} + \frac{F_{b_K}^*}{S_K} \frac{F_{b_L}}{S_L} \right]. \qquad (A10)$$

In the bin-centered limit (Equation 3.18 in [18]), $\langle \Xi^2 \rangle \approx 1$. Establishing N without Ξ but with,

$$\hat{A}_{\mathcal{P}}^2 \equiv \sum_{KL \in \mathcal{P}} (\hat{a}^K \hat{a}^L)^2, \qquad (A11)$$

$$\hat{B}_{\mathcal{P}}^2 \equiv \sum_{KL \in \mathcal{P}} (\hat{b}^K \hat{b}^L)^2, \qquad (A12)$$

$$\hat{C}_{\mathcal{P}}^2 \equiv \sum_{KL \in \mathcal{P}} (\hat{a}^K \hat{b}^L \hat{a}^L \hat{b}^K), \qquad (A13)$$

we obtain,

$$N = \frac{10}{\sqrt{2}} \left[\hat{A}_{\mathcal{P}}^2 + 2\hat{C}_{\mathcal{P}}^2 + \hat{B}_{\mathcal{P}}^2 \right]^{-1/2}, \qquad (A14)$$

$$\rho = \frac{4\sqrt{2}\Re \sum_{KL \in \mathcal{P}} \left[\frac{F_{a_K}^*}{S_K} \frac{F_{a_L}}{S_L} + \frac{F_{b_K}^*}{S_L} \frac{F_{b_L}}{S_L} \right]}{\sqrt{\hat{A}_{\mathcal{P}}^2 + 2\hat{C}_{\mathcal{P}}^2 + \hat{B}_{\mathcal{P}}^2}}. \qquad (A15)$$

Compare ρ to the $\mathcal{F}$-statistic in a specific case. Take Q detectors indexed by X, Y, each with $M \equiv T_{\text{obs}}/T_{\text{sft}}$ SFTs. Assume a frequency-dependent, stationary noise PSD $S_h(f)$. Allow all pairs $\mathcal{P}$, so the sum expands into a double sum of a double sum seen in Equation A7,

$$\rho = 4\sqrt{2} \left(S_h^2(f)\sqrt{\hat{A}_{\mathcal{P}}^2 + 2\hat{C}_{\mathcal{P}}^2 + \hat{B}_{\mathcal{P}}^2} \right)^{-1} \qquad (A16)$$

$$\times \Re \left[\sum_X^Q \sum_{K(X)}^M F_{a_{K(X)}}^* \sum_Y^Q \sum_{L(X)}^M F_{a_{L(Y)}} \cdots \right.$$
$$\left. + \sum_X^Q \sum_{K(X)}^M F_{b_{K(X)}}^* \sum_Y^Q \sum_{L(Y)}^M F_{b_{L(Y)}} \right],$$

As the index I in Equation A7 is detector independent,

$$\sum_X^Q \sum_{K(X)}^M F_{a_{K(X)}}^* = F_a^*, \qquad (A17)$$

so too for the L index and b terms, allowing (self-) *auto-correlations*. Normally, the cross-correlation method does not allow auto-correlations [18], but it can [16], such that,

$$\rho = 4\sqrt{2} \frac{|F_a|^2 + |F_b|^2}{S_h^2(f)\sqrt{\hat{A}_{\mathcal{P}}^2 + 2\hat{C}_{\mathcal{P}}^2 + \hat{B}_{\mathcal{P}}^2}}, \qquad (A18)$$

Simplifying the denominator,

$$\hat{A}_{\mathcal{P}}^2 = \left(\sum_I^{QM} (\hat{a}^I)^2 \right)^2, \qquad (A19)$$

$$\hat{A}_{\mathcal{P}} = \frac{1}{T_{\text{sft}}} \sum_I^{QM} \hat{a}^I \hat{a}^I T_{\text{sft}}, \qquad (A20)$$
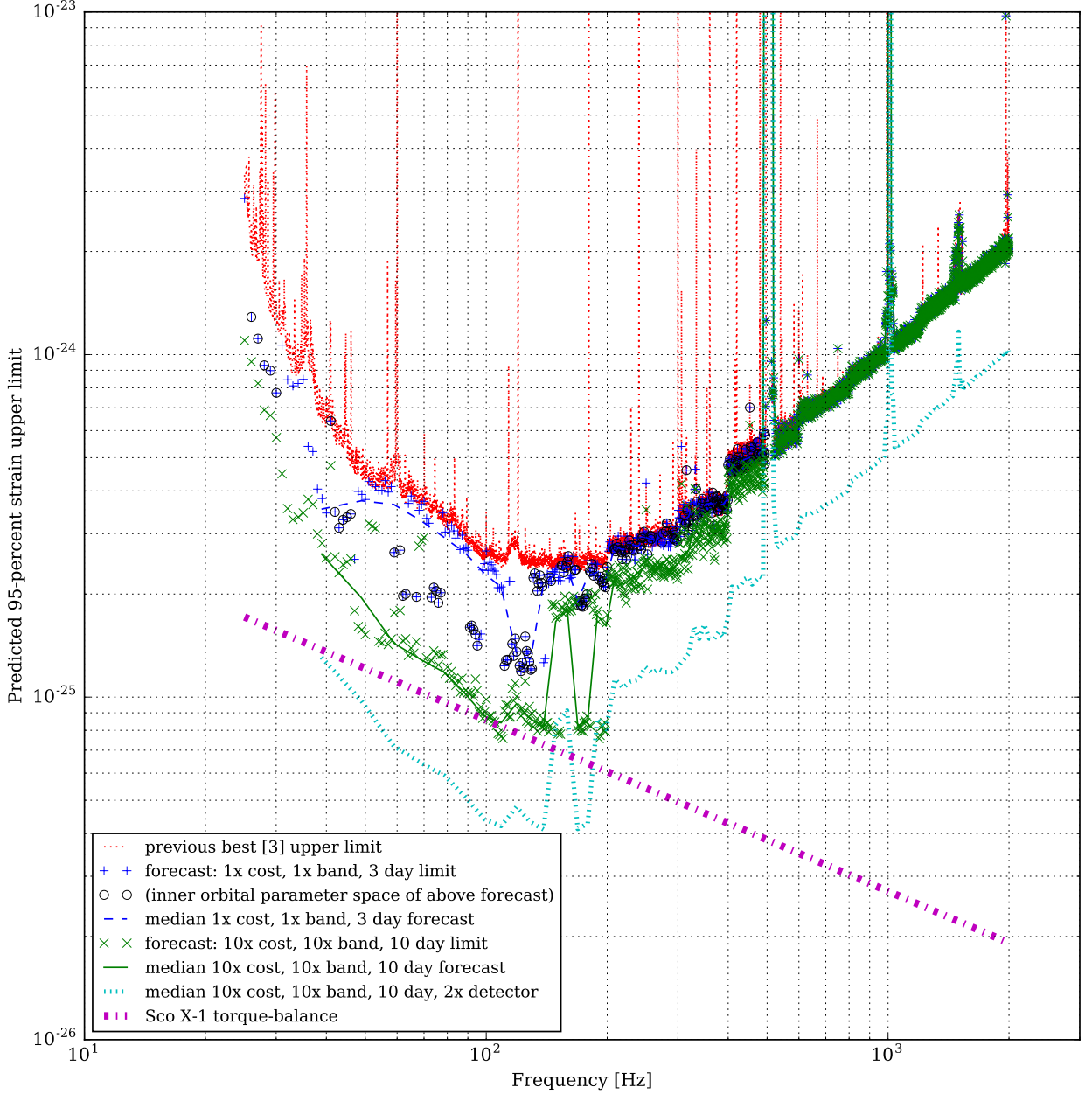
FIG. 6. Theoretically-forecast (hypothetical) upper limits extrapolated from the O1 cross-correlation method's 95% marginalized Sco X-1 search, without $\cos \iota$ information [3] Extrapolation based on gains in Figures 4 and 5. As in the former, actual gains may be less. O1 limit shown in *(red dots)*. Extrapolation based on equal cost, 3-day spin-wandering limit in *(blue dashed line)*, and based on 10× cost, $10 \times f_{\text{band}}$, 10-day spin-wandering in *(green solid line)*. Respective *(blue +)* and *((green ×)* indicate every 1-Hz interval (original upper limit used 50-mHz intervals). Lines trace (non-running) median of 10-Hz bins. Fluctuations seen in lines because some bands limited by spin-wandering, others not. The (+) bands are circled (○) if they are in a long T_{max} part of orbital parameter space, defined as $a_p \leq 2.1663$ s, $1131415225 < T_{\text{asc}} \leq 1131415583$ based on O1 setup [3]. Longer T_{max} times benefit more from resampling, as noted in Figure 4. Caution: correct upper limits would require estimation of detection efficiency, not done here. Present extrapolation suggests torque-balance might be attained in the best *(green ×)* case at 100 Hz, or *(small cyan dashes)*, with a 2× improved detector noise floor (T_{obs} same as O1), from 40 to 140 Hz. Compare to Figure 7.
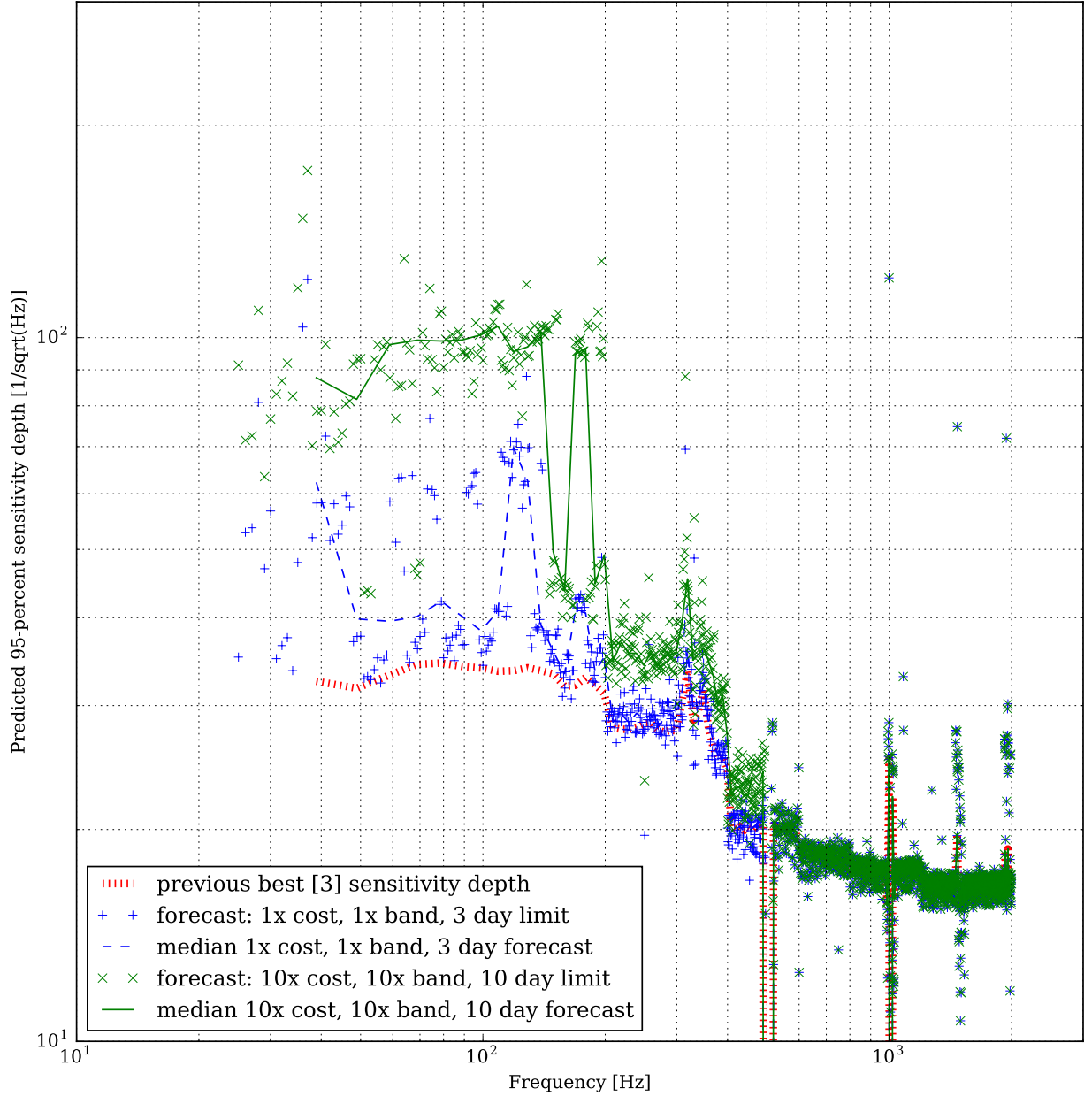
FIG. 7. Projected sensitivity depth *vs* frequency. Compare to Figure 6, from which this graph is derived by dividing by the harmonic mean of the Hanford (H1) and Livingston (L1) detector ASD noise floors. Additional uncertainty is introduced by contaminating lines in the detectors. The equal-cost case (*blue* +) shows bands up to about 200 Hz fluctuating between the mid-30 and mid-60 $Hz^{-1/2}$ range; as noted in Figure 4, gain is less at fixed cost for less-probable parts of orbital parameter space, because they are allocated shorter T_{max} in O1. For the best, $10\times$ cost, $10 \times f_{band}$, 10-day spin-wandering case (*green* $\times$), most parts of orbital parameter space reach approximately 100 $Hz^{-1/2}$. Caution: results are hypothetical and theoretically extrapolated from the timing model, as noted in Figure 6; also, eccentricity is assumed negligible. Results may improve depending on computational cost re-optimization (confer [67]).

which Riemann integrates for $a(t)$, $b(t)$ that vary slowly compared to T_{sft} (faster than T_{obs}, so an overall shift is negligible and $T_0 \to T_{\mathrm{obs}}$ in Equation A9),

$$\hat{A}_{\mathcal{P}} \approx \frac{QT_{\mathrm{obs}}}{S_h(f)}(a||a) \tag{A21}$$

Forming norms $A \equiv (a||a)$, $B \equiv (b||b)$, $C \equiv (a||b)$ [13]:

$$\rho = 4\sqrt{2}\frac{|F_a|^2 + |F_b|^2}{S_h(f)QT_{\mathrm{obs}}\sqrt{A^2 + 2C^2 + B^2}}. \tag{A22}$$

2. Comparison to the $\mathcal{F}$-statistic

The $\mathcal{F}$-statistic is a maximum-likelihood (ML) estimator. Values of $\mathcal{A}^\mu$ are chosen where the likelihood ratio is a maximum, Λ_{ML}. Composing frequency-integrated projections x_μ onto the basis h^μ in Equation 7, with $\mathcal{M}^{\mu\nu}$ the ML projections of h^μ onto h^ν [13, 15, 55, 56, 59]:

$$\Lambda_{\mathrm{ML}} = e^{\mathcal{F}}, \tag{A23}$$

$$\mathcal{F} \equiv \frac{1}{2}x_\mu \mathcal{M}^{\mu\nu}x_\nu, \tag{A24}$$

$$= \frac{4}{S_h(f)T_0}\frac{B|F_a|^2 + A|F_b|^2 - 2C\Re(F_aF_b^*)}{A \cdot B - C^2}.$$

Both ρ and $\mathcal{F}$ are dimensionless. As after Equation 5.15 in [16], ρ and $\mathcal{F}$ are proportional when $A \approx B$, $C \ll A, B$:

$$\rho \approx 4\sqrt{2}\frac{|F_a|^2 + |F_b|^2}{S_h(f)QT_{\mathrm{obs}}\sqrt{2A^2}}, \tag{A25}$$

equating $\mathcal{F}$ with $T_0 = QT_{\mathrm{obs}}$,

$$\mathcal{F} \approx \frac{4}{S_h(f)T_0}\frac{|F_a|^2 + |F_b|^2}{A}. \tag{A26}$$

Even for multiple detectors, (all-pairs) ρ can converge to the (fully-coherent) $\mathcal{F}$-statistic. Illustrating the crossover is now possible.

Dhurandhar *et al* [16] introduce the cross-correlation method starting from two data streams, like the stochastic radiometer [47], instead of the multi-detector $\mathcal{F}$-statistic [55]. The weight matrix $\mathbf{W}$ of Whelan *et al* [18] can merge these viewpoints. Any SFT, from any detector, is a dimension in $\mathbf{z}$ ('flattening' SFTs over the Greek indices also represented as boldface in [55] to represent different detectors). Cutler & Schutz Equation 3.8 [55] has $2\mathcal{F} = \sum_{a,d}(\Gamma^{-1})^{ad}(\mathbf{x}|\mathbf{h_a})(\mathbf{x}|\mathbf{h_d})$: a, d are the waveform components μ, ν in our Equation A23. Their inner products of $\mathbf{x}$ with the waveforms $\mathbf{h_a}, \mathbf{h_d}$ are scalar-valued vectors indexed by a and d, equivalent to summing F_a or F_b from multiple detectors. Only then is $\mathcal{F}$ computed. The sum of fully-coherent single-detector $\mathcal{F}$ does not equal the *fully*-coherent multiple-detector $\mathcal{F}$, which takes into account the cross-detector terms and converges with the ideal cross-correlation method.

Divergence can occur with *semi*-coherent methods [69, 70]. Semicoherent calculations with $T_{\mathrm{coh}} < T_{\mathrm{obs}}$ are more efficient, having higher sensitivity at fixed computational cost, than fully-coherent methods [14, 69, 70]. The sum of $\mathcal{F}$-statistics over $T_{\mathrm{obs}}/T_{\mathrm{coh}}$ segments of $\mathcal{F}$-statistics is computed, albeit with reduced sensitivity compared to the much more expensive fully-coherent search. Joint- and single-detector $\mathcal{F}$ can both be computed for each T_{coh}. (Comparison between joint and single is the basis of the $\mathcal{F}$-statistic consistency veto [71]). The main difference between the cross-correlation method and the semicoherent $\mathcal{F}$-statistic is that the former, distinguishing K and L, helps to exclude auto-correlations.

Examine the optimal amplitude parameters in $\mathcal{M}^{\mu\nu}$ and weights $\mathbf{W}$. Despite Equation 23's resemblance to Equation A25, $\mathbf{W}$ and $\mathcal{M}^{\mu\nu}$ are matrices over different spaces. $\mathbf{W}$ (implicit indices) is of SFTs, whereas $\mathcal{M}^{\mu\nu}$ (explicit indices) is of four amplitude parameters. The amplitude-parameter space metric is $\mathcal{M}^{\mu\nu}$, so $2\mathcal{F} = x^\mu x_\nu$ [56]. In principle, ρ_{ideal} might not use $\hat{\Gamma}^{\mathrm{ave}}_{KL}$ (chosen to avoid specifying $\cos\iota$ and ψ [18]), but instead Γ based on maximization or marginalization [56] of $\mathcal{A}^\mu$. A start would be projections, $\mathbf{z}_\mu$, of $\mathbf{z}$ onto the h_μ basis. Each $\mathbf{z}$ (a data vector, implicitly indexed, *e.g.*, by SFTs) can be projected to extract the components along the 4 amplitude-parameter space dimensions, producing the $N_{\mathrm{sft}} \times 4$ matrix, $\mathbf{z}_\mu$. Schematically,

$$\rho_{\mathrm{ideal}} = \frac{1}{2}\mathcal{M}^{\mu\nu}\mathbf{z}_\mu^\dagger \mathbf{W}\mathbf{z}_\nu. \tag{A27}$$

Hence $\mathbf{z}_\mu$ absorb Φ and can be thought of as Fourier transforms of source-frame data. The matrix $\mathcal{M}$ absorbs Γ from $\mathbf{W}$, leaving $\mathbf{W}$ a binary-valued index of which SFTs to pair. Such a statistic would echo the likelihood ratio mentioned in Section V of [16]. Recall, ρ and $\mathcal{F}$ converge when $A \approx B$ and $C \ll A, B$, *i.e.*, when $\mathcal{M}$ is proportional to the identity matrix. So, when amplitude space is flat and auto-correlations are included, $\rho \approx \mathcal{F}$.

One impetus for cross-detector correlation is that only signal should be coherent. This ground underlies stochastic searches: GW strains between detectors are related, but the noise is statistically independent (see Section III of [54]). Appendix D will revisit the robustness and merits of pairing choices for $\mathbf{W}$.

3. Comparison to the stochastic search

The stochastic search [47, 54, 72] is built on same-time cross-correlation of multiple detectors [73], whereby sensitivity depends on an overlap reduction function [74]. In stochastic literature, S often indicates detector strains and P indicates noise PSDs. To keep consistency in this paper, S will be replaced by X for strain and P by S for noise. For the stochastic radiometer Y-statistic [47, 74],

$$Y = \int_{-\infty}^{+\infty} df \int_{-\infty}^{+\infty} df \, \delta_T(f - f') X_1^*(f) Q(f') X_2(f'),$$
(A28)

where δ_T is a finite-time Dirac delta function approximation, X_1, X_2 are Fourier-transformed detector strains, and Q is an optimal filter. For sky direction $\hat{\Omega}' = \hat{\Omega}$,

$$Y_{\hat{\Omega}'} = (\lambda T) \int_{-\infty}^{+\infty} df \frac{\gamma_{\hat{\Omega}'}^* H}{S_1 S_2} X_1^* X_2.$$
(A29)

Expanding, with normalization factor λ, measurement duration T, and S_1 and S_2 the noise PSD, as well as $H(f)$ the strain power of the stochastic background, with overlap reduction function $\gamma_{\hat{\Omega}'}$ and polarizations $A \in \{+, \times\}$ and detector separation vector $\Delta\vec{x}$:

$$\gamma_{\hat{\Omega}'} = \frac{1}{2} \sum_A e^{i2\pi f \hat{\Omega} \cdot \frac{\Delta\vec{x}}{c}} F_1^A(\hat{\Omega}) F_2^A(\hat{\Omega}).$$
(A30)

This Y is effectively a case of the simultaneous cross-correlation method's ρ restricted to different detectors [18]. Notationally, $\hat{\Omega} = \vec{n}$. Radiometer Δx is c times detector arrival time difference $\Delta d_{KL} \equiv (\vec{r}_K(t) - \vec{r}_L(t)) \cdot \vec{n}/c$, stemming from Equation 8. Then, the radiometer phase difference $2\pi f \hat{\Omega} \cdot (\Delta\vec{x}/c)$ equals $\Delta\Phi_{KL}$ in Equation 26 and is $2\pi f_0 \Delta d_{KL}$. Because $10\Gamma_{KL}^{\text{ave}} = F_+^K F_+^L + F_\times^K F_\times^L$ [18],

$$\gamma_{\hat{\Omega}'} = 5 \frac{\sqrt{S_K S_L}}{2T_{\text{sft}}} e^{i\Delta\Phi_{KL}} \hat{\Gamma}_{KL}^{\text{ave}}.$$
(A31)

To be exact [75], where $\tilde{Q}(\hat{\Omega}, t, f; H) = Q(f')$, $\Delta t = T_{\text{sft}}$ is time segment length and $\gamma^*(\Omega, t, f) = \gamma_{\hat{\Omega}'}$,

$$\tilde{Q}(\hat{\Omega}, t, f; H) = \lambda(\Omega, t) \frac{H(f)\gamma^*(\Omega, t, f)}{S_1(t; |f|) S_2(t; |f|)}, \quad \text{(A32)}$$

and $\lambda(\Omega, t) = \lambda T$, absorbing λ in Equation A29.

Absent a Φ model, radiometer must equally sum frequency bin contributions over $\Delta f \geq \Delta f_{\text{obs}}$ (Equation 16). This width means $\Xi \approx 1$ and $X_1 = \sum_K \sum_k \tilde{x}_{Kk}$, $X_2 = \sum_L \sum_l \tilde{x}_{Ll}$ (referring to Equation 18; this is imprecise when radiometer uses overlapping, windowed bins [72] and the cross-correlation method uses non-overlapping rectangular bins). Moreover, $S_1 = S_K$, $S_2 = S_L$, $T = \Delta t$.

Compare to looking for an isolated point with no other sources and refer to the discussion following Equation 3.36 of [75]. If the stochastic background is taken as constant in frequency, $H^2(f) = 1$, λ simplifies (integrating over frequency and substituting Equation 3.34 of [75] as directed for network power P_{NW}^2),

$$\lambda(t) \approx [\Delta t P_{NW}^2(t)]^{-1},$$
(A33)

$$[\lambda(\Omega, t)\Delta t]^{-1} \approx \frac{5}{T_{\text{sft}}} \frac{1}{\sqrt{S_K S_L}} \hat{\Gamma}_{KL}^{\text{ave}},$$
(A34)

$$\lambda(\Omega, t) = \frac{\sqrt{S_K S_L}}{5} \frac{1}{\hat{\Gamma}_{KL}^{\text{ave}}},$$
(A35)

$$N = \frac{5}{\sqrt{2}} \bar{\lambda}(\Omega),$$
(A36)

where N is the cross-correlation method's normalization and $\bar{\lambda}(\Omega)$ is harmonic root mean square radiometer normalization.

In that case, after all substitutions and considering the cross-correlation method's ρ evaluated over all bins and only between the same SFT pairs as radiometer,

$$\rho \approx 4\sqrt{2}\Re\left(Y_{\Omega'}\right),$$
(A37)

by taking sums over cross-correlation method indices K and L to produce radiometer S_1 and S_2. Exact equality results for a single pair, such as the fully-coherent, cross-detector-only ρ. This conclusion bolsters Whelan *et al* [18] (notably Section III.D), stating that the cross-correlation method is similar to the radiometer with a phase model to allow different-time correlations.

The cross-correlation method, the radiometer, and the $\mathcal{F}$-statistic, which all are described as near-*optimal* under different conditions, do converge in certain limits. Understanding the cross-correlation method's intersections aids theory and practice.

In theory, viewing $\mathcal{F}$ as approximating the Bayesian $\mathcal{B}$-statistic [56] informs ρ as an approximate function of the likelihood ratio [16]. This perspective might facilitate Bayesian model selection for vetoes using alternative line hypotheses to compare against the signal hypothesis [76]. It may also link search set-up optimization for detection probability to rigorous statements about posterior probability [67]. Radiometric techniques might generate a deconvolved sky map of future detections [72, 75].

This paper should resolve confusion about the cross-correlation method. It does not use cross-detector data as its template. The cross-correlation method is a matched-filter-based semicoherent search, with the template corresponding to the signal model with chosen amplitude parameters, searched over the Doppler parameters. It differs in which filtered data are conjugated in the real-valued statistic.

In practice, at present, ties between the statistics help bring resampling from the $\mathcal{F}$-statistic into the cross-correlation method. Resampling solves the problem that $\Phi(t)$, particularly the time-varying $\phi(t)$, is not increasing at uniform frequency f_0, because of the time-varying Doppler shifts. If Doppler modulation were constant, then a Fourier transform could supply F_a (or the cross-correlation method's components F_{a_K} and F_{a_L}) and F_b, providing an entire frequency band. The data must be moved into the *source frame*, in which velocity with respect to the source is constant.

Appendix B: Downsampling and heterodyning

Section III is done with downsampled data, heterodyned downwards in frequency by f_h. Consider a bandpass-limited sample (subscript p) of Short Fourier Transform data for SFT K, equivalent to a rectangular frequency-domain window with starting bin k_a and ending bin k_b. Gaps in the set of SFTs are zero-padded to yield $M = T_{\rm obs}/T_{\rm sft}$. Equation 27 says that the time index with respect to SFT start time is j, and with respect to the observation run is q_K.

The q_K are non-overlapping integers from 0 to $MN - 1$, whereas j (implicitly depending on K) range from 0 to $N - 1$. Start times and SFT durations are integer multiples of the sampling time, $t_K - T_{\rm sft}/2 \equiv KT_{\rm sft}$, and $j\delta t = q_K \delta t - t_K + T_{\rm sft}/2$.

The ideal *bandpassed* data $x_{K,p}$ from an inverse Fourier transform of the whole $T_{\rm obs}$ would be,

$$
x_{K,p}(q_K \delta t) \equiv \sum_{k=k_a}^{k=k_b} e^{\mathrm{i}2\pi q_K \delta t \frac{k}{T_{\rm sft}}} \frac{z_{Kk}}{T_{\rm sft}}, \qquad \text{(B1)}
$$
$$
= \sum_{k=k_a}^{k=k_b} e^{\mathrm{i}2\pi j\delta t \frac{k}{T_{\rm sft}}} \frac{z_{Kk}}{T_{\rm sft}},
$$

because $t_K - T_{\rm sft}/2$ is an integer. When $k_a = 0$, $k_b = N - 1$, $x_{K,p}$ is equivalent to x_K in Equation 27. Yet we want not simply bandpassed data, but downsampled, heterodyned data. Since $M > 1$, we handle the sum over K. The difficulty is keeping phase coherence between inverse Fourier transforms.

Heterodyne frequency f_h is in the center of the band, near central bin $k_h \equiv (k_a + k_b)/2$. For discrete bins, the nearest frequency $\bar{f}_h \equiv k_h T_{\rm sft}^{-1}$. The frequency $\bar{f}_h = f_h - f_r$ is of the nearest integer bin to the ideal heterodyne f_h, where f_r is the remainder.

Let $l \equiv k - k_h$, so $k = l + k_h$:

$$
x_{K,p}(q_K \delta t) = \sum_{(l+k_h)=k_a}^{(l+k_h)=k_b} e^{\mathrm{i}2\pi j\delta t \frac{(l+k_h)}{T_{\rm sft}}} \frac{z_{K(l+k_h)}}{T_{\rm sft}} \qquad \text{(B2)}
$$
$$
= e^{\mathrm{i}2\pi j\delta t \frac{k_h}{T_{\rm sft}}} \sum_{l=k_a-k_h}^{l=k_b-k_h} e^{\mathrm{i}2\pi j\delta t \frac{l}{T_{\rm sft}}} \frac{z_{K(l+k_h)}}{T_{\rm sft}}.
$$

The sum contains all information on $[k_a, k_b]$. Call it x_K^h:

$$
x_K^h(q_K \delta t) \equiv \sum_{l=k_a-k_h}^{l=k_b-k_h} e^{\mathrm{i}2\pi j\delta t \frac{l}{T_{\rm sft}}} \frac{z_{K(l+k_h)}}{T_{\rm sft}}, \qquad \text{(B3)}
$$

$$
x_{K,p}(q_K \delta t) = e^{\mathrm{i}2\pi j\delta t \frac{k_h}{T_{\rm sft}}} x_K^h(q_K \delta t), \qquad \text{(B4)}
$$

expressing bandpassed $x_{K,p}$ in terms of the desired, frequency-shifted x_K^h. In continuous time, Equation B4 is the expression $x_{K,p}(t) = \exp(\mathrm{i}2\pi f_h t) x_K^h(t)$, where $x_{K,p}$ is the bandpassed data (frequency content at f) and x_K^h is the heterodyned data (frequency content at $f - f_h$). Many derivations stop here, but we need the phase corrections for heterodyning a set of SFTs.

To represent *complex, downsampled* data in a frequency band $f_{\rm band}$ without aliasing, we need a total bandwidth of $\Delta f_{\rm load}$. Note that $\Delta f_{\rm load}$ must cover not only all frequencies of interest but also frequency modulation's Doppler wings, $\Delta f_{\rm drift}$, with additional bins to account for spectral leakage, including D 'Dirichlet terms'. The total width $\Delta f_{\rm load}$ is [65],

$$
\Delta f_{\rm load} = \left(1 + \frac{4}{2D + 1}\right)\left(f_{\rm band} + \Delta f_{\rm drift} + \frac{16}{T_{\rm sft}}\right). \tag{B5}
$$

Then we find the new sampling time interval is not δt but rather $\delta t' = 1/\Delta f_{\rm load}$. The old number of samples in an SFT is $N = T_{\rm sft}/(\delta t)$ and the new number is $N' = \Delta f_{\rm load} T_{\rm sft}$; $N' \leq N$. (N' can be rounded up to ensure it is an integer). Create the new coordinate q_K', so $t = q_K' \delta t'$. The Fourier transform kernel must contain an integer, and $q_K' \approx (t_K - T_{\rm sft}/2 + j\delta t)/(\delta t')$ is not generally an integer. Additional phase corrections thus arise. Note,

$$
q_K' = \Delta f_{\rm load} q_K \delta t, \tag{B6}
$$
$$
q_K' k/N' = (q_K \delta t) k/T_{\rm sft}, \tag{B7}
$$

Meanwhile we can choose k_a and k_b with a difference $k_b - k_a = \Delta f_{\rm load} T_{\rm sft}$, ergo $k_b - k_a = N'$:

$$
k_a = \left(f_h - \frac{1}{2}\Delta f_{\rm load}\right) T_{\rm sft}, \tag{B8}
$$

$$
k_b = \left(f_h + \frac{1}{2}\Delta f_{\rm load}\right) T_{\rm sft}. \tag{B9}
$$

In practice, we will use the minimum frequency of interest, $f_a = f_{\rm min}$, to choose a heterodyne frequency $f_h = f_{\rm min} + \frac{1}{2}f_{\rm band}$.

As $t = q_K \delta t$, we can substitute $q_K' \delta t'$ into the argument of $x_{K,p}(q_K \delta t)$ as defined in Equation B2. Using N',

$$
x_{K,p}(q_K' \delta t') = \sum_{l=k_a-k_h}^{l=k_b-k_h} e^{\mathrm{i}2\pi q_K' \frac{l+k_h}{2\Delta f_{\rm load} T_{\rm sft}}} \frac{z_{K(l+k_h)}}{T_{\rm sft}}, \quad \text{(B10)}
$$
$$
= e^{\mathrm{i}2\pi q_K' k_h/N'} \sum_{l=k_a-k_h}^{l=k_b-k_h} e^{\mathrm{i}2\pi q_K' l/N'} \frac{z_{K(l+k_h)}}{T_{\rm sft}},
$$

We need to break apart q_K' in the sum:

$$
q_K'/N' = \frac{t_K - T_{\rm sft}/2}{T_{\rm sft}} + \frac{j\delta t}{T_{\rm sft}}, \tag{B11}
$$

where again, because $(t_K - T_{\rm sft}/2)$ is always an integer multiplied by integer l, the first term evaluates to unity

in the sum exponent. In $q'_K k_h/N'$, however, though k_h is also integer, we leave the term so we can see the effect of approximating $\bar{f}_h$. We find,

$$x_{K,p}(q'_K \delta t') = e^{i2\pi q_K \delta t \frac{k_h}{T_{\text{sft}}}} x_K^h(q_K \delta t). \quad (B12)$$

This result concords with Equation B4. Considering $\bar{f}_h$,

$$x_{K,p}(q'_K \delta t') = e^{i2\pi q_K \delta t(f_h - f_r)} x_K^h(q_K \delta t), \quad (B13)$$

where an approximation is used for this Appendix,

$$x_{K,p}(q'_K \delta t') \approx e^{i2\pi q_K \delta t f_h} x_K^h(q_K \delta t). \quad (B14)$$

Generally the code will have access to f_h but not k_h; the remainder f_r is fixed by later by rounding to the nearest bin (in the paper body, f_r^*).

Next, we seek x_K^h in downsampled time. Our goal is x^h covering all the observing time, but we must go through $x_{K,p}$ to preserve the phase shifts between SFTs. For the single-SFT case, we could just substitute $q'_K \delta t'$ into the argument for x_K^h and be done.

Notice that $k_a - k_h = -N'/2$, $k_b - k_h = +N'/2 - 1$ (for an even number of samples including 0). For any point in time, comparison with Equation B3 shows,

$$x_K^h(q_K \delta t) = \sum_{l=-N'/2}^{l=N'/2-1} e^{i2\pi j(\delta t/\delta t')\frac{l}{N'}} \frac{z_{K(l+k_h)}}{T_{\text{sft}}}, (B15)$$

we could define the generally non-integer $j' = j(\delta t/\delta t')$; fortunately, $j'/N' = j/N$. The exponent is then $\exp(i2\pi j'l/N)$.

For a detour, note that x_K^h is almost fit for a Fourier transform, but it requires an index shift. Periodicity in the Fourier transform means that any substitution $jk \to jk + QN$ for a transform with time steps j, frequency steps k, and number of samples N, by integer Q, leaves the result invariant. For half-integer Q, the substitution moves positive frequencies into negative frequencies (increasing in the same direction as before) and vice versa. Choose new index $m \equiv l + N'/2$, so $l = m - N'/2$:

$$x_K^h(q_K \delta t) = (-1)^{-j'} \sum_{m=0}^{m=N-1} B(0, N')$$
$$\times e^{i2\pi jm/N} \frac{z_{K(m+k_h-N'/2)}}{T_{\text{sft}}}, \quad (B16)$$

where, for illustration, $B(0, N')$ is the Boxcar function, acting as a bandpass. This x_K^h is at the full sampling rate and is only theoretical. The sum term is a straightforward inverse Fourier transform, from m to j, of the z_K data from frequency bins $k_h - N'/2$ to $k_h + N'/2 - 1$. In practice, the $(-1)^{-j'}$ factor (the move from positive

to negative frequencies) depends into the conventions of Fast Fourier Transform programs. Care is required to ensure the right convention. For us, the interface with the *FFTW* library absorbs this factor. We will use this Fourier transform after constructing the time series.

To construct the full time-series for the entire observing run, use the time-shift Equation B14 for $x_{K,p}$ and Equation B15 for x_K^h, noting that $\bar{f}_h \approx f_h$:

$$\exp(i2\pi f_h q_K \delta t) \approx e^{i2\pi f_h[t_K - T_{\text{sft}}/2 + j'k_h/N']}, (B17)$$

whereby the *frequency-shifted, heterodyned, downsampled* $x_K^h(q'_K \delta t')$ has the SFT start time phase shift with respect to $x_K^h(q_K \delta t)$:

$$x_K^h(q'_K \delta t') = e^{-i2\pi f_h[t_K - \frac{T_{\text{sft}}}{2} + \frac{j'k_h}{N'}]} x_{K,p}(q'_K \delta t')(B18)$$

The $j'k_h/N'$ can be absorbed into the bandpassing by a change of index, providing a quantity amenable to an FFT. Returning to Equation B2 for $x_{K,p}(q_K \delta t)$, which equals $x_{K,p}(q'_K \delta t')$ at equal times t:

$$x_K^h(q_K \delta t)i = e^{-i2\pi f_h[t_K - \frac{T_{\text{sft}}}{2}]} \quad (B19)$$
$$\times \sum_{l=-N'/2}^{N'/2+1} e^{i2\pi \left(\frac{j'(l+k_h)}{N'} - \frac{j'k_h}{N'}\right)} \frac{z_{K(l+k_h)}}{T_{\text{sft}}}.$$

Invoking Equation B16,

$$x_K^h(q'_K \delta t') = e^{-i2\pi f_h[t_K - \frac{T_{\text{sft}}}{2}]}(-1)^{-j'} \quad (B20)$$
$$\times \sum_{m=0}^{N'-1} e^{i2\pi j'm/N'} \frac{z_{K(m+k_h-N'/2)}}{T_{\text{sft}}}$$

While for arbitrary q_K, j' is not an integer, the downsampled time-series q'_K is specifically chosen for times where it is. Then the sum is indeed an inverse discrete Fourier transform of bandpassed data (which by itself is $x_{K,p}$), but also shifted by k_h. Including the negative-frequency sign convention with $(-1)^{-j'}$, call this $x_{K,s}$:

$$x_{K,s}(q'_K \delta t') \equiv (-1)^{-j'} \quad (B21)$$
$$\times \sum_{m=0}^{N'-1} e^{i2\pi j'm/N'} \frac{z_{K(m+k_h-N'/2)}}{T_{\text{sft}}}$$
$$x_K^h(q'_K \delta t') = e^{-i2\pi f_h[t_K - \frac{T_{\text{sft}}}{2}]} x_{K,s}(q'_K \delta t'). \quad (B22)$$

This result for the exponent depends on the heterodyne frequency f_h and SFT mid-time t_K but not the index q_K. In comparison with Equation B14, the index $j\delta t$ has been absorbed. So it is generally true of any time $t = q_K \delta t$, including $t = q'_K \delta t'$. Carefully note, however, that x_K^h is *still* heterodyned in the sense that a Fourier transform

will yield the spectrum shifted by f_h. All the correction has done is shift the phase so that different SFTs are in-phase. We now use this alignment to construct the complete time series from the SFTs.

Since q_K is distinct for the entire time series, that series of $x_{q'}^h \equiv x^h(q'\delta t')$ is the sum (neglecting windowing),

$$x^h(q'\delta t') = \sum_{K=0}^{M} e^{-\mathrm{i}2\pi f_h[t_K - T_{\mathrm{sft}}/2]} x_{K,s}(q'_K \delta t'). \text{(B23)}$$

In practice, the quantity $x_{K,s}(q'_K \delta t')$ is computed from the inverse Fourier transform of a band of data centered around f_h with bandwidth f_{band}, so Equation B23 is the simplest construction of the complete downsampled time series. Again, any signal at frequency f_0 in x is at $f_0 - f_h$ in x'. Downsampling also reduces the computation cost of interpolating into the BB frame.

Appendix C: Interpretation and degeneracies of statistic

Several properties of the ρ statistic should be noted that do not neatly fit into the main text. In the fully-coherent limit, just as the $\mathcal{F}$-statistic is proportional to the log-likelihood ratio of a sinusoidal waveform hypothesis compared to Gaussian noise [13], so too should the ρ statistic be interpreted. In this limit, the set of output $\rho(f_0, \lambda)$ from a search constitutes a sampling of the likelihood surface. This likelihood surface is amenable to composite hypothesis testing, as well as Bayesian interpretation [56].

Locally, the 'likelihood surface' of ρ is well-described by the metric approximation [18]. Globally, long-range degeneracies appear. Degeneracies step mainly from surfaces of $d\Phi = 0$ in the phase model, Equation 14. In the $(f_0, a_p, T_{\mathrm{asc}})$ space, these degeneracies form a cone along the f_0 axis, with the vertex at the maximum ρ. The surface of the cone arises from the largest component of the set of sidebands from residual phase modulation when (a_p, T_{asc}) are offset from their true values. This surface has been noted elsewhere in cross-section as a 2-dimensional X shape, for example in the (f_0, a_p) plane [44, 46]. Because this extended surface correlates neighboring templates, naïve division by a trials factor equal to the number of templates (Bonferroni correction) may yield an overly-conservative p-value. The metric may also be too conservative for high values of mismatch [77].

Semicoherent statistics such as ρ grow proportionally to $(T_{\mathrm{obs}} T_{\mathrm{max}})^{1/4}$, and they also grow proportionally to h_0^2. This is in contrast to fully-coherent statistics, which take $T_{\mathrm{max}} = T_{\mathrm{obs}}$ and therefore grow proportionally to $T_{\mathrm{obs}}^{1/2}$. However, another class of power-based statistics, such as the 'TwoSpect' method [45], also grows as $T_{\mathrm{obs}}^{1/4}$ but, differently from the semicoherent case, as h_0^4. GW

phase coherence is not used over timescales longer than one SFT in these power-based statistics, and the final statistic depends on the power of a second FFT, over the orbital cycle.

The cross-correlation method's code must calculate ρ as efficiently as possible in a sample of the likelihood surface that does not miss its peak. Viewed as semi-coherent choice of the weights matrix $\mathbf{W}$, the goal is the calculate the largest number of elements of the weights matrix for the lowest cost. Skipping the auto-correlation in our code comes at the cost of the statistic contribution from that element. Avoidance of auto-correlation is natural from the standpoint of the *Radiometer*, which only permits same-time correlations and has no signal model. For the radiometer, auto-correlation would contaminate the search with the noise of the detector. From the standpoint of the $\mathcal{F}$-statistic, it is conversely natural to include the auto-correlation, because it fits in the middle of an FFT. Capturing the adjacent elements of the weights matrix from the cross-correlation method with an FFT requires additional overlap of a factor of $T_{\mathrm{coh}}/T_{\mathrm{short}} \geq 3$. It should be determined whether the cost of this overlap is worth the exclusion of noise (and signal) contributions from the auto-correlation.

Appendix D: Merits of the cross-correlation method in noisy data

The cross-correlation method, unlike the $\mathcal{F}$-statistic but like the *Radiometer* method, avoids auto-correlation by default. Consider the presence of some sine-Gaussian glitch in the data that might justify this avoidance:

$$g(t) = Ae^{-(t-t_0)^2/(2\sigma^2)} \sin \omega t - \phi_0. \quad (D1)$$

In the Fourier domain in which the cross-correlation method computes its statistic, the Fourier transform of $g(t)$, $\tilde{g}(f)$, is the convolution of the Fourier transforms of the Gaussian and sinusoidal terms, which are respectively also Gaussian and a Dirac delta function. The glitch does contribute noise in a Gaussian frequency distribution around the frequency ω, with amplitude proportional to A. By removing the auto-correlation, such glitches will never correlate with themselves. Assuming that ω and t_0 are randomly-distributed, they will be unlikely to correlate with other glitches at different times. Therefore, the noise background of the cross-correlation method could conceivably be lower.

Empirically, values of ρ and $\mathcal{F}$ appear similar for comparable noise and signal strength. Whether the theoretically lower background of the cross-correlation method holds in real data is an important test. If the two statistics recover signals comparably well for the same coherent integration time, then whichever calculates a given coherence time most efficiently is best. This paper has established a path between the two methods.

[1] B.P Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016).

[2] P.R. Brady, T. Creighton, C. Cutler, and B.F. Schutz, "Searching for periodic sources with LIGO," Phys. Rev. D **57**, 2101 (1998).

[3] B.P Abbott *et al.*, "Upper limits on gravitational waves from Scorpius X-1 from a model-based cross-correlation search in Advanced LIGO data," Astrophys J **847**, 47 (2017).

[4] J. Papaloizou and J.E. Pringle, "Gravitational radiation and the stability of rotating stars," MNRAS **184**, 501 (1978).

[5] R.V. Wagoner, "Gravitational radiation from accreting neutron stars," Ap. J. **278**, 345 (1984).

[6] L. Bildsten, "Gravitational radiation and rotation of accreting neutron stars," Astrophys. J. Lett. **501**, L89 (1998).

[7] R. Giacconi, H. Gursky, F.R. Paolini, and B.B. Rossi, "Evidence for X rays from sources outside the solar system," Phys. Rev. Lett. **9** (1962).

[8] P. Shawhan, "Gravitational-wave astronomy: observational results and their impact," Class. Quant. Grav. **27**, 084017 (2010).

[9] B.J. Owen, "How to adapt broad-band gravitational-wave searches for r-modes," Phys. Rev. D **82**, 104002 (2010).

[10] N. Andersson, "A new class of unstable modes of rotating relativistic stars," APJ **502**, 708 (1998).

[11] J.L Friedman and S.M. Morsink, "Axial instability of rotating relativistic stars," APJ **502** (1998).

[12] B.J. Owen, L. Lindblom, C. Cutler, B.F. Schutz, A. Vecchio, and N. Andersson, "Gravitational waves from hot young rapidly rotating neutron stars," Phys. Rev. D **58**, 084020 (1998).

[13] P. Jaranowski, A. Królak, and B.F. Schutz, "Data analysis of gravitational-wave signals from spinning neutron stars: the signal and its detection," Phys. Rev. D **58**, 063001 (1998).

[14] P.R. Brady and T. Creighton, "Searching for periodic sources with LIGO. II. Hierarchical searches," Phys. Rev. D **61**, 082001 (2000).

[15] Pinkesh Patel, Xavier Siemens, Rejean Dupuis, and Joseph Betzwieser, "Implementation of barycentric resampling for continuous wave searches in gravitational wave data," Phys. Rev. D **81**, 084032 (2010), arXiv:0912.4255 [gr-qc].

[16] Sanjeev Dhurandhar, Badri Krishnan, Himan Mukhopadhyay, and John T. Whelan, "Cross-correlation search for periodic gravitational waves," Phys. Rev. D **77**, 082001 (2008), arXiv:0712.1578.

[17] C.T.Y. Chung, A. Melatos, B. Krishnan, and J.T. Whelan, "Designing a cross-correlation search for continuous-wave gravitational radiation from a neutron star in the supernova remnant SNR 1987A," MNRAS **414**, 2650 (2011).

[18] John T. Whelan, Santosh Sundaresan, Yuanhao Zhang, and Prabath Peiris, "Model-based cross-correlation search for gravitational waves from Scorpius X-1," Phys. Rev. D **91**, 102005 (2015).

[19] 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, "Gravitational waves from Scorpius X-1: A comparison of search methods and prospects for detection with advanced detectors," Phys. Rev. D **92**, 023006 (2015).

[20] B. Abbott *et al.*, "Searches for periodic gravitational waves from unknown isolated sources and Scorpius X-1: results from the second LIGO science run," Phys. Rev. D **76**, 082001 (2007).

[21] J. Abadie *et al.*, "Directional limits on persistent gravitational waves using LIGO S5 science data," Phys. Rev. Lett. **107**, 271102 (2011).

[22] J. Aasi *et al.*, "First all-sky search for continuous gravitational waves from unknown sources in binary systems," Phys. Rev. D **90**, 062010 (2014).

[23] L. Sammut, C. Messenger, A. Melatos, and B.J. Owen, "Implementation of the frequency-modulated sideband search method for gravitational waves from low mass x-ray binaries," Phys. Rev. D **89**, 043001 (2014).

[24] J. Aasi *et al.*, "Directed search for gravitational waves from Scorpius X-1 with initial LIGO data," Phys. Rev. D **91**, 062008 (2015).

[25] G.D. Meadors, E. Goetz, K. Riles, T. Creighton, and F. Robinet, "Searches for continuous gravitational waves from Scorpius X-1 and XTE J1705-305 in LIGO's sixth science run," Phys Rev D **95**, 042005 (2017).

[26] B.P. Abbott *et al.*, "Directional limits on persistent gravitational waves from Advanced LIGO's first observing run," Phys Rev Lett **118**, 121102 (2017).

[27] B.P. Abbott *et al.*, "Search for gravitational waves from Scorpius X-1 in the first Advanced LIGO observing run with a hidden Markov model," Phys Rev D **95**, 122003 (2017).

[28] D. Chakrabarty *et al.*, "Nuclear-powered millisecond pulsars and the maximum spin frequency of neutron stars," Nature **424**, 42 (2003).

[29] A.L. Watts, B. Krishnan, L. Bildsten, and B.F. Schutz, "Detecting gravitational wave emission from the known accreting neutron stars," MNRAS **389**, 839 (2008).

[30] J. Aasi *et al.* (LIGO Scientific), "Advanced LIGO," Class. Quant. Grav. **32**, 074001 (2015), arXiv:1411.4547 [gr-qc].

[31] F. Acernese *et al.* (VIRGO), "Advanced Virgo: a second-generation interferometric gravitational wave detector," Class. Quant. Grav. **32**, 024001 (2015), arXiv:1408.3978 [gr-qc].

[32] Yoichi Aso, Yuta Michimura, Kentaro Somiya, Masaki Ando, Osamu Miyakawa, Takanori Sekiguchi, Daisuke Tatsumi, and Hiroaki Yamamoto (The KAGRA Collaboration), "Interferometer design of the KAGRA gravitational wave detector," Phys. Rev. D **88**, 043007 (2013).

[33] B. Behnke, M.A. Papa, and R. Prix, "Postprocessing methods used in the search for continuous gravitational-wave signals from the Galactic Center," PRD **91**, 064007 (2015).

[34] Paola Leaci and Reinhard Prix, "Directed searches for continuous gravitational waves from binary systems: Parameter-space metrics and optimal Scorpius X-1 sensitivity," Phys. Rev. D **91**, 102003 (2015).

[35] K. Riles, "Gravitational waves: sources, detectors and searches," Prog. in Particle & Nucl. Phys. **68**, 1 (2013).

[36] Badri Krishnan, Alicia M. Sintes, Maria Alessandra Papa, Bernard F. Schutz, Sergio Frasca, and Cristiano Palomba, "Hough transform search for continuous gravitational waves," Phys. Rev. D **70**, 082001 (2004).

[37] B. Abbott, Phys. Rev. D **77**, 022001 (2008).

[38] B. Abbott *et al.*, "All-sky LIGO search for periodic gravitational waves in the early fifth-science-run data," Phys. Rev. Lett **102**, 111102 (2009).

[39] V. Dergachev, "On blind searches for noise dominated signals: a loosely coherent approach," Class. Quant. Grav. **27**, 205017 (2010).

[40] J. Abadie *et al.*, "All-sky search for periodic gravitational waves in the full S5 LIGO data," Phys. Rev. D **85**, 022001 (2012).

[41] R.J. Dupuis and G. Woan, "Bayesian estimation of pulsar parameters from gravitational wave data," Phys. Rev. D **72**, 102002 (2005).

[42] J. Aasi *et al.*, "Gravitational-waves from known pulsars: results from the initial detector era," Astrophys. J **785**, 119 (2014).

[43] C. Messenger and G. Woan, "A fast search strategy for gravitational waves from low-mass x-ray binaries," Classical and Quantum Gravity **24**, S469 (2007).

[44] S. Suvorova, L. Sun, A. Melatos, W. Moran, and R.J. Evans, "Hidden Markov model tracking of continuous gravitational waves from a neutron star with wandering spin," Phys Rev D **93**, 123009 (2016).

[45] E. Goetz and K. Riles, "An all-sky search algorithm for continuous gravitational waves from spinning neutron stars in binary systems," Class. Quant. Grav. **28**, 215006 (2011).

[46] G.D. Meadors, E. Goetz, and K. Riles, "Tuning into Scorpius X-1: adapting a continuous gravitational-wave search for a known binary system," Class. Quant. Grav. **33**, 105017 (2016).

[47] Stefan W Ballmer, "A radiometer for stochastic gravitational waves," Classical and Quantum Gravity **23**, S179 (2006).

[48] S. van der Putten, H. J. Bulten, J. F. J. van den Brand, and M. Holtrop, "Searching for gravitational waves from pulsars in binary systems: An all-sky search," Journal of Physics Conference Series **228**, 012005 (2010).

[49] M. F. Skrutskie *et al.*, "The Two Micron All Sky Survey (2MASS)," The Astronomical Journal **131**, 1163–1183 (2006).

[50] C.F. Bradshaw, E.B. Fomalont, and B.J. Geldzahler, "High-resolution parallax measurements of Scorpius X-1," ApJ **512**, L121 (1999).

[51] D. K. Galloway, S. Premachandra, D. Steeghs, T. Marsh, J. Casares, and R. Cornelisse, "Precision Ephemerides for Gravitational-wave Searches. I. Sco X-1," Ap J **781**, 14 (2014), arXiv:1311.6246 [astro-ph.HE].

[52] L. Wang, D. Steeghs, and D. Galloway, "Sco X-1 parameters summary," private communication (2016).

[53] D. Steeghs and J. Casares, "The Mass Donor of Scorpius X-1 Revealed," Astrophys. J. **568**, 273–278 (2002), astro-ph/0107343.

[54] B. Allen and J.D. Romano, "Detecting a stochastic background of gravitational radiation: signal processing strategies and sensitivities," Phys. Rev. D **59**, 102001 (1999).

[55] C. Cutler and B.F. Schutz, "Generalized F-statistic: Multiple detectors and multiple gravitational wave pulsars," Phys Rev D **72**, 063006 (2005).

[56] R. Prix and B. Krishnan, "Targeted search for continuous gravitational waves: Bayesian versus maximum-likelihood statistics," Class Quant Grav **26**, 204013 (2009).

[57] A. Mukherjee, C. Messenger, and K. Riles, "Accretion-induced spin-wandering effects on the neutron star in Scorpius X-1: implications for continuous gravitational wave searches," (2017-10-17), 1710.06185 [gr-qc].

[58] R. Prix, "Search for continuous gravitational waves: metric of the multidetector F-statistic," Phys Rev D **75**, 023004 (2007).

[59] J.T. Whelan, R. Prix, C.J. Cutler, and J.L Willis, "New coordinates for the amplitude parameter space of continuous gravitational waves," Class Quant Grav **31**, 065002 (2014).

[60] R. Blandford and S. Teukolsky, "Arrival-time analysis for a pulsar in a binary system," Astrophys J **205**, 580 (1976).

[61] B. Allen and G. Mendell, "SFT Data Format Version 2 Specification," LIGO DCC **T040164** (2004), https://dcc.ligo.org/LIGO-T040164/public.

[62] B Allen, M.A. Papa, and B.F. Schutz, "Optimal strategies for sinusoidal signal detection," Phys Rev D **66**, 102003 (2002).

[63] B. Schutz, private communication (2017).

[64] The LIGO Scientific Collaboration, "LALApps repository," Web: http://www.lsc-group.phys.uwm.edu/daswg/.

[65] R. Prix, "Characterizing timing and memory-requirements of the F-statistic implementations in LALSuite," (2017), https://dcc.ligo.org/LIGO-T1600531/public.

[66] R. Prix, "The F-statistic and its implementation in ComputeFStatistic_v2," LIGO DCC **T0900149-v3** (2011), https://dcc.ligo.org/LIGO-T0900149-v3/public.

[67] J. Ming, B. Krishnan, M.A. Papa, C. Aulbert, and H. Fehrmann, "Optimal directed searches for continuous gravitational waves," Phys Rev D **93**, 064011 (2016).

[68] S. Suvorova, P. Clearwater, A. Melatos, L. Sun, W. Moran, and R.J. Evans, "Hidden Markov model tracking of continuous gravitational waves from a binary neutron star with wandering spin. II. Binary orbital phase tracking," Phys Rev D **96**, 102006 (2017).

[69] C. Cutler, I. Gholami, and B. Krishnan, "Improved stack-slide searches for gravitational-wave pulsars," Phys Rev D **72**, 042004 (2005).

[70] R. Prix and M. Shaltev, "Search for continuous gravitational waves: optimal StackSlide method at fixed computing cost," Phys Rev D **85**, 084010 (2012).

[71] J. Aasi *et al.*, "Einstein@home all-sky search for periodic gravitational waves in LIGO S5 data," Phys Rev D **87**, 042001 (2013).

[72] E. Thrane, S. Ballmer, J.D. Romano, S. Mitra, D. Talukder, S. Bose, and V. Mandic, "Probing the anisotropies of a stochastic gravitational-wave background using a network of ground-based laser interferomters," Phys Rev D **80**, 122002 (2009).

[73] N. Christensen, "Measuring the stochastic gravitational-radiation background with laser-interferometric antennas," Phys Rev D **46**, 5250 (1992).

[74] E.E. Flanagan, "Sensitivity of the Laser Interferomter Gravitational Wave Observatory to a stochastic background, and its dependence on the detector orientations," Phys Rev D **48**, 2389 (1993).

[75] S. Mitra, S. Dhurandhar, T. Souradeep, A. Lazzarini, V. Mandic, S. Bose, and S. Ballmer, "Gravitational wave radiometry: mapping a stochastic gravitational wave background," Phys Rev D **77**, 042002 (2008).

[76] D. Keitel, R. Prix, M.A. Papa, P. Leaci, and M. Siddiqi, "Search for continuous gravitational waves: improving robustness versus instrumental artifacts," Phys Rev D **89**, 064023 (2014).

[77] K. Wette, "Empirically extending the range of validity of parameter-space metrics for all-sky searches for gravitational-wave pulsars," Phys Rev D **94**, 122002 (2016).