The $\mathcal{F}$-statistic and its implementation in `ComputeFStatistic_v2`

Reinhard Prix

21 June, 2011

LIGO-T0900149-v3

Abstract

These notes represent a somewhat high-level documentation of `ComputeFStatistic_v2`, starting from a derivation and general discussion of the $\mathcal{F}$-statistic, down to expressions that very closely resemble what is actually implemented in the code.

Contents

1 The signal $h(t)$ measured at the detector

1.1 General waveform

A plain gravitational wave $h_{\mu\nu}$ propagating along the unit-vector $-\hat{n}$ can be written in TT gauge as a purely spatial tensor $\underline{h}$, namely

$$\underline{h}(t, \vec{r}) = h_+(\tau)\, \underline{e}^+ + h_\times(\tau)\, \underline{e}^\times \,, \tag{1}$$

where $\tau = t + \hat{n} \cdot \vec{r}/c$ and the polarization tensors $\underline{e}^{\{+,\times\}}$ are defined as

$$\underline{e}^+ = \hat{u} \otimes \hat{u} - \hat{v} \otimes \hat{v}\,, \quad \text{and} \quad \underline{e}^\times = \hat{u} \otimes \hat{v} + \hat{v} \otimes \hat{u}\,, \tag{2}$$

in terms of unit vectors $\hat{u}, \hat{v}$ that form an orthonormal basis $\{\hat{u}, \hat{v}, -\hat{n}\}$ of the wave frame. The choice of basis $\{\hat{u}, \hat{v}\}$ in the transversal wave plane is arbitrary, but one often chooses preferred directions given either by the source-geometry or by the principal polarization axis of elliptically polarized waves. It is therefore convenient to re-express this in a source-independent basis that only depends on the propagation direction $-\hat{n}$ of the wave and the choice of an SSB-fixed reference frame $\{\hat{x}, \hat{y}, \hat{z}\}$. Such a frame is conventionally constructed using the unit basis vectors $\hat{\xi} \equiv \hat{n} \times \hat{z}/|\hat{n} \times \hat{z}|$, $\hat{\eta} \equiv \hat{\xi} \times \hat{n}$ and $-\hat{n}$. This definition is such that $\hat{\xi}$ lies in the equatorial plane and $\hat{\eta}$ points into the northern hemisphere. We now define the polarization angle ψ as the angle from $\hat{\xi}$ to $\hat{u}$, measured counter-clockwise in the plane with $-\hat{n}$ pointing *at* us[1], i.e. $\sin\psi = \hat{u} \cdot \hat{\eta}$.

This allows us to express the polarization basis $\{\hat{u}, \hat{v}\}$ in terms of the basis $\{\hat{\xi}, \hat{\eta}\}$ via a simple rotation by ψ around $-\hat{n}$, namely

$$\hat{u} = \hat{\xi} \cos\psi + \hat{\eta} \sin\psi \,, \tag{3}$$
$$\hat{v} = -\hat{\xi} \sin\psi + \hat{\eta} \cos\psi \,. \tag{4}$$

Introducing polarization-independent basis tensors in the wave-frame,

$$\underline{\varepsilon}^+ \equiv \hat{\xi} \otimes \hat{\xi} - \hat{\eta} \otimes \hat{\eta}\,, \tag{5}$$
$$\underline{\varepsilon}^\times \equiv \hat{\xi} \otimes \hat{\eta} + \hat{\eta} \otimes \hat{\xi}\,, \tag{6}$$

we can express the wave-basis $\underline{e}^{\{+,\times\}}$ as

$$\underline{e}^+ = \cos 2\psi\, \underline{\varepsilon}^+ + \sin 2\psi\, \underline{\varepsilon}^\times \tag{7}$$
$$\underline{e}^\times = -\sin 2\psi\, \underline{\varepsilon}^+ + \cos 2\psi\, \underline{\varepsilon}^\times \,. \tag{8}$$

[1] This is what is meant by the phrase "counter-clockwise around $-\hat{n}$" used in [8, 10]

In the long-wavelength limit (LWL), where the arm length L of the detector satisfies $L \ll \lambda/2\pi$ (where λ is the GW wavelength), the scalar response $h^X(t)$ of a detector X to an incident GW tensor $\underline{h}$ is expressible simply in terms of its detector tensor $\underline{d}^X$, namely

$$h^X(t) = \underline{d}^X(t) : \underline{h}(\tau^X) = d^X_{ij}\, h^{ij}(\tau^X)\,, \tag{9}$$

where $\tau^X(t) = t + \hat{n}\cdot\vec{r}^X(t)/c$ is (neglecting relativistic corrections) the arrival time of a wavefront at the SSB, which arrives at the detector X (at position $\vec{r}^X$) at time t. This timing relation accounts for the Doppler effect due to the motion of the detector relative to the source. The LWL detector tensor for an interferometer with arms along $\hat{l}_1$ and $\hat{l}_2$ is simply given by

$$\underline{d} = \frac{1}{2}\left(\hat{l}_1 \otimes \hat{l}_1 - \hat{l}_2 \otimes \hat{l}_2\right)\,. \tag{10}$$

Using (1), we can write (9) in the form

$$h^X(t) = F^X_+(t)\, h_+(\tau^X) + F^X_\times(t)\, h_\times(\tau^X)\,, \tag{11}$$

in terms of the so-called beam-pattern functions

$$F^X_+(t;\hat{n},\psi) \equiv \underline{d}^X(t) : \underline{e}^+\,, \quad F^X_\times(t;\hat{n},\psi) \equiv \underline{d}^X(t) : \underline{e}^\times\,. \tag{12}$$

Changing to the polarization-independent basis $\underline{\varepsilon}^{+,\times}$ using (7), we find

$$F^X_+(t;\hat{n},\psi) = a^X(t;\hat{n})\,\cos 2\psi + b^X(t;\hat{n})\,\sin 2\psi\,, \tag{13}$$

$$F^X_\times(t;\hat{n},\psi) = b^X(t;\hat{n})\,\cos 2\psi - a^X(t;\hat{n})\,\sin 2\psi\,, \tag{14}$$

where the antenna-pattern functions a^X, b^X are defined as

$$a^X(t;\hat{n}) \equiv \underline{d}^X(t) : \underline{\varepsilon}^+(\hat{n})\,, \quad b^X(t;\hat{n}) \equiv \underline{d}^X(t) : \underline{\varepsilon}^\times(\hat{n})\,. \tag{15}$$

This formulation has the advantage of explicitly factoring out the polarization angle ψ. The sky-position $\hat{n}$ of the source is expressible in standard equatorial (or ecliptic) coordinates α (right ascension), and δ (declination) as

$$\hat{n} = (\cos\delta\,\cos\alpha, \cos\delta\,\sin\alpha, \sin\delta)\,, \tag{16}$$

and by the above definitions, the corresponding polarization-independent wave-plane basis $\hat{\xi}, \hat{\eta}$ is therefore expressible as

$$\begin{aligned} \hat{\xi} &= (\sin\alpha,\, -\cos\alpha,\, 0)\,, \\ \hat{\eta} &= (-\cos\alpha\,\sin\delta,\, -\sin\alpha\,\sin\delta,\, \cos\delta)\,. \end{aligned} \tag{17}$$

The contractions (15) are explicitly given by

$$\underline{d} : \underline{\varepsilon} = d_{11}\varepsilon_{11} + d_{22}\varepsilon_{22} + d_{33}\varepsilon_{33} + 2\left(d_{12}\varepsilon_{12} + d_{13}\varepsilon_{13} + d_{23}\varepsilon_{23}\right)\,, \tag{18}$$

where $\underline{\varepsilon}^{\{+,\times\}}$ are easily computed in SSB coordinates from (17), and the problem of computing a, b is therefore reduced to computing the detector tensor $\underline{d}^X(t)$ as a function of time in this coordinate system.

1.2 Continuous-wave signals

The GW class of "continuous waves" is characterized by a signal model $h_{+,\times}(\tau)$ (in the SSB) of the form

$$h_{+}(\tau) = A_{+}\,\cos\Phi(\tau)\,, \quad h_{\times}(\tau) = A_{\times}\,\sin\Phi(\tau)\,. \tag{19}$$

Assuming a slowly varying intrinsic signal frequency $2\pi f(\tau) \equiv d\Phi(\tau)/d\tau$, the phase $\Phi(\tau)$ can be expanded around the reference time $\tau_{\rm ref}$, namely

$$\Phi(\tau) = \phi_0 + \phi(\Delta\tau)\,, \quad \text{where} \tag{20}$$

$$\phi_0 \equiv \Phi(\tau_{\rm ref})\,, \tag{21}$$

$$\phi(\Delta\tau) \equiv 2\pi \sum_{s=0} \frac{f^{(s)}(\tau_{\rm ref})}{(s+1)!}\,(\Delta\tau)^{s+1}\,. \tag{22}$$

The detector-specific timing relation for isolated neutron stars contains relativistic corrections for the light-travel in the solar system. These corrections are taken into account in the numerical $\mathcal{F}$-statistic computation in `CFS_v2`, but for simplicity we give here only the first order Newtonian timing model,

$$\Delta\tau^X(t;\hat{n}) \equiv \tau^X - \tau_{\rm ref} \approx t - \tau_{\rm ref} + \frac{\vec{r}^X(t)\cdot\hat{n}}{c}\,, \tag{23}$$

where τ^X is the arrival-time in the SSB of the GW phase reaching detector X at time t. The spin parameters $f^{(s)}(\tau_{\rm ref})$ are defined as

$$f^{(s)}(\tau_{\rm ref}) \equiv \left.\frac{d^s f(\tau)}{d\tau^s}\right|_{\tau_{\rm ref}}\,. \tag{24}$$

We denote the set of "Doppler parameters" affecting the time evolution of the phase $\phi(\Delta\tau^X)$ as $\lambda \equiv \{\hat{n}, f^{(s)}(\tau_{\rm ref})\}$. Combining (11), (13) (19), we find

$$h^X(t;\mathcal{A},\lambda) = \sum_{\mu=1}^{4} \mathcal{A}^\mu\, h^X_\mu(t;\lambda)\,, \tag{25}$$

with the four amplitude parameters $\mathcal{A}^\mu$ given by

$$\begin{aligned}
\mathcal{A}^1 &= A_+\,\cos\phi_0\,\cos 2\psi - A_\times\,\sin\phi_0\,\sin 2\psi\,,\\
\mathcal{A}^2 &= A_+\,\cos\phi_0\,\sin 2\psi + A_\times\,\sin\phi_0\,\cos 2\psi\,,\\
\mathcal{A}^3 &= -A_+\,\sin\phi_0\,\cos 2\psi - A_\times\,\cos\phi_0\,\sin 2\psi\,,\\
\mathcal{A}^4 &= -A_+\,\sin\phi_0\,\sin 2\psi + A_\times\,\cos\phi_0\,\cos 2\psi\,,
\end{aligned} \tag{26}$$

which is a re-parametrization of the (detector-independent) signal-parameters $A_+, A_\times, \phi_0, \psi$. The (detector-dependent) wave-components $h^X_\mu(t;\lambda)$ are

$$
\begin{aligned}
h_1^X(t) &= a^X(t)\,\cos\phi(\Delta\tau^X)\,, \quad h_2^X(t) = b^X(t)\,\cos\phi(\Delta\tau^X)\,,\\
h_3^X(t) &= a^X(t)\,\sin\phi(\Delta\tau^X)\,, \quad h_4^X(t) = b^X(t)\,\sin\phi(\Delta\tau^X)\,.
\end{aligned}
\tag{27}
$$

It is often useful to also consider the complex basis functions instead

$$
\begin{aligned}
h_{\mathrm{a}}^X(t) &\equiv h_1^X - i h_3^X = a^X\, e^{-i\phi^X}\,,\\
h_{\mathrm{b}}^X(t) &\equiv h_2^X - i h_4^X = b^X\, e^{-i\phi^X}\,.
\end{aligned}
\tag{28}
$$

We see from (26) that there is some gauge-freedom in the amplitude-parameters $\{A_+, A_\times, \psi, \phi_0\}$, namely

$$
\begin{aligned}
\text{(i)} &\quad \psi \to \psi + \pi/2,\ \phi_0 \to \phi_0 + \pi\\
\text{(ii)} &\quad \psi \to \psi + \pi/4,\ \phi_0 \to \phi_0 - \pi/2,\ A_+ \leftrightarrow A_\times\\
\text{(iii)} &\quad \phi_0 \to \phi_0 + \pi,\ A_+ \to -A_+,\ A_\times \to -A_\times
\end{aligned}
\tag{29}
$$

Applying (i) twice, and taking account of the trivial gauge-freedom by 2π, we also obtain the invariance $\psi \to \psi + \pi$.
In the case of a triaxial NS, the signal-amplitudes $A_{+/\times}$ are expressible explicitly in terms of the wave-amplitude h_0 and the inclination angle ι with respect to the line-of-sight, namely

$$
A_+ = \frac{1}{2} h_0\,\left(1 + \cos^2\iota\right)\,, \quad A_\times = h_0\,\cos\iota\,. \tag{30}
$$

where the overall GW amplitude h_0 is given by

$$
h_0 = \frac{4\pi^2 G}{c^4}\,\frac{\epsilon\, I_{zz}\, f^2}{d}\,, \tag{31}
$$

in terms of the triaxial ellipticity $\epsilon \equiv |I_{xx} - I_{yy}|/I_{zz}$, and the distance d. Note that this partially fixes the gauge, namely

$$
A_+ \geq |A_\times| \geq 0\,, \tag{32}
$$

which excludes gauge-transformations (ii) and (iii) in (29) In order to fix a unique gauge also for ψ, ϕ_0, we restrict the quadrant of ψ to be $\psi \in [-\pi/4, \pi/4)$ (in accord with the TDS convention), which can always be achieved by the gauge-transformation (i), while ϕ_0 remains unconstrained in $\phi_0 \in [0, 2\pi)$.

2 Noise and detection statistic

2.1 Theoretical framework

We follow the notation of [5, 1] by denoting vectors of detector-specific quantities in boldface, i.e. $\{\boldsymbol{x}\}^X = x^X$. We can now write the explicit dependencies of the signal-model (25) on the signal-parameters as

$$\boldsymbol{h}(t;\mathcal{A},\lambda) = \mathcal{A}^\mu\,\boldsymbol{h}_\mu(t;\lambda)\,, \tag{33}$$

where we implicitly sum over amplitude-indices $\mu,\nu \in \{1,2,3,4\}$. If the data $x^X(t)$ measured at different detectors X consists of stationary Gaussian noise $n^X(t)$ and a signal with parameters $\mathcal{A}$, λ, we can write

$$\boldsymbol{x}(t) = \boldsymbol{n}(t) + \boldsymbol{h}(t;\mathcal{A},\lambda)\,, \tag{34}$$

in terms of the signal-model (33). It is sometimes useful to consider the discrete-time formulation, as it more closely describes the actual measured data, which is sampled as discrete time-steps $t_j \equiv j\,\Delta t$, namely $x^X_j \equiv x^X(t_j)$. The noise samples $\{n^X_j\}$ are assumed to be drawn from a Gaussian distribution with zero mean, $E\big[n^X_j\big] = 0$, and covariance matrix

$$\kappa^{XY}_{jl} \equiv E\big[n^X_j\,n^Y_l\big]\,, \tag{35}$$

which allows us to write the noise probability distribution as

$$P\left(\boldsymbol{n}|\boldsymbol{\kappa}\right) = k\,e^{-\frac{1}{2}(\boldsymbol{n}|\boldsymbol{n})}\,, \tag{36}$$

where k is a normalization factor independent of the noise $\boldsymbol{n}$, and where we defined the discrete-time version of the multi-detector scalar product (40) as

$$(\boldsymbol{x}|\boldsymbol{y}) \equiv x^X_j\,\kappa^{jl}_{XY}\,y^Y_l\,, \tag{37}$$

with automatic summation over time-indices j,l and detector-indices X,Y, and κ^{jl}_{XY} defined as the inverse of the covariance matrix, namely

$$\kappa^{XY}_{jm}\,\kappa^{ml}_{YZ} = \delta^{Xl}_{Zj}\,. \tag{38}$$

For known functions of time g^X_j, h^X_j, and Gaussian noise n^X_j following (36), it is now easy to prove the general identity

$$\begin{aligned}
E[(\boldsymbol{n}|\boldsymbol{g})\,(\boldsymbol{n}|\boldsymbol{h})] &= E\Big[n^X_j\,\kappa^{jl}_{XY} g^Y_l\,n^Z_m \kappa^{mp}_{ZV} h^V_p\Big] \\
&= g^Y_l\,h^V_p\,\kappa^{jl}_{XY}\kappa^{mp}_{ZV}\,\kappa^{XZ}_{jm} \\
&= g^Y_l\,\kappa^{lp}_{YV}\,h^V_p \\
&= (\boldsymbol{g}|\boldsymbol{h})\,.
\end{aligned} \tag{39}$$

As shown in [3] (for the single-detector case), the natural discrete-time scalar product (37), which came directly from the Gaussian probability distribution (36), leads to the well-known continuous-time formulation in the appropriate limit, namely

$$(\boldsymbol{x}|\boldsymbol{y}) \to 4\,\Re \int_0^\infty \tilde{x}^X(f)\, S^{-1}_{XY}(f)\, \tilde{y}^{Y*}(f)\, df\,, \tag{40}$$

where $\Re$ denotes the real part, and $\tilde{x}(f)$ denotes the Fourier transformed

$$\tilde{x}(f) \equiv \int x(t)\, e^{-i2\pi\, ft}\, dt \approx \Delta t \sum_j x_j e^{-i2\pi\, f\, t_j}\,. \tag{41}$$

The matrix S^{-1}_{XY} satisfies $S^{-1}_{XY}\, S^{YZ} = \delta^Z_X$, where the (single-sided!) noise PSD matrix S^{XY} is defined as

$$S^{XY}(f) = 2\int_{-\infty}^\infty \kappa^{XY}(\tau)\, e^{-i2\pi\, f\tau}\, d\tau\,, \tag{42}$$

in terms of the correlation matrix (assuming stationary noise) $\kappa^{XY}(\tau) \equiv E\big[n^X(t+\tau)\, n^Y(t)\big]$. In the case of uncorrelated noises between detectors, i.e. $S^{XY} = S^X\,\delta^{XY}$, the scalar product (40) reduces to a sum over single-detector scalar products, namely

$$(\boldsymbol{x}|\boldsymbol{y}) = \sum_X^{N_{\mathrm{Det}}} (x^X|y^X) = \sum_X 4\,\Re\int_0^\infty \frac{\tilde{x}^X(f)\,\tilde{y}^{X*}(f)}{S^X(f)}\, df\,, \tag{43}$$

where N_{Det} is the number of detectors used. Assuming $\boldsymbol{x}(t)$ or $\boldsymbol{y}(t)$ is a narrow-band continuous-wave signal (25) at frequency f_{s}, we can approximate this scalar product as

$$(\boldsymbol{x}|\boldsymbol{y}) \approx 2\sum_X^{N_{\mathrm{Det}}} S^{-1}_X(f_{\mathrm{s}}) \int_0^T x^X(t)\, y^X(t)\, dt\,. \tag{44}$$

We can use the noise probability distribution (36) together with (34) to express the likelihood of observing data $\boldsymbol{x} = \boldsymbol{n} + \boldsymbol{h}$ in the presence of a signal $\boldsymbol{h}(t;\mathcal{A},\lambda)$, namely

$$P\left(\boldsymbol{x}|\mathcal{A},\lambda,\boldsymbol{S}\right) = k\, e^{-\frac{1}{2}(\boldsymbol{x}|\boldsymbol{x})}\, e^{(\boldsymbol{x}|\boldsymbol{h}) - \frac{1}{2}(\boldsymbol{h}|\boldsymbol{h})}\,, \tag{45}$$

while in the noise-only case $h_0 = 0$, i.e. $\mathcal{A}^\mu = 0$, the likelihood is simply

$$P\left(\boldsymbol{x}|0,\boldsymbol{S}\right) = k\, e^{-\frac{1}{2}(\boldsymbol{x}|\boldsymbol{x})}\,. \tag{46}$$

Therefore the likelihood ratio $\mathcal{L}(\boldsymbol{x};\mathcal{A},\lambda) \equiv P\left(\boldsymbol{x}|\mathcal{A},\lambda\right)/P\left(\boldsymbol{x}|0\right)$ is found as

$$\begin{aligned}\log\mathcal{L}(\boldsymbol{x};\mathcal{A},\lambda) &= (\boldsymbol{x}|\boldsymbol{h}) - \frac{1}{2}\,(\boldsymbol{h}|\boldsymbol{h})\\ &= \mathcal{A}^\mu x_\mu - \frac{1}{2}\mathcal{A}^\mu \mathcal{M}_{\mu\nu}\mathcal{A}^\nu\,,\end{aligned} \tag{47}$$

where we substituted the "JKS" signal factorization (33), and where we defined

$$x_\mu(\lambda) \equiv (\boldsymbol{x}|\boldsymbol{h}_\mu)\,, \tag{48}$$

$$\mathcal{M}_{\mu\nu}(\lambda) \equiv (\boldsymbol{h}_\mu|\boldsymbol{h}_\nu) = (\partial_\mu \boldsymbol{h}|\partial_\nu \boldsymbol{h})\,, \tag{49}$$

defining $\partial_\mu \equiv \partial/\partial\mathcal{A}^\mu$. From the last expression we see that $\mathcal{M}_{\mu\nu}$ is the Fisher matrix for the parameters $\mathcal{A}^\mu$. It is straightforward to analytically maximize the likelihood-ratio (47) with respect to the four amplitudes $\mathcal{A}^\mu$, and we obtain the so-called "$\mathcal{F}$-statistic", namely

$$\mathcal{F}(\boldsymbol{x};\lambda) \equiv \max_{\mathcal{A}} \log\mathcal{L}(\boldsymbol{x};\mathcal{A},\lambda) = \frac{1}{2}\,x_\mu\,\mathcal{M}^{\mu\nu}\,x_\nu\,, \tag{50}$$

where $\mathcal{M}^{\mu\nu} \equiv \{\mathcal{M}^{-1}\}^{\mu\nu}$, i.e. $\mathcal{M}_{\mu\sigma}\mathcal{M}^{\sigma\nu} = \delta_\mu^\nu$. The maximum-likelihood (ML) estimators for the four unknown amplitudes $\mathcal{A}^\mu$ are given by

$$\mathcal{A}^\mu_{\mathrm{ML}} = \mathcal{M}^{\mu\nu}\,x_\nu\,, \tag{51}$$

and alternatively we can also express the $\mathcal{F}$-statistic (50) in the form

$$2\mathcal{F}(\boldsymbol{x};\lambda) = \mathcal{A}^\mu_{\mathrm{ML}}\,\mathcal{M}_{\mu\nu}\,\mathcal{A}^\nu_{\mathrm{ML}}\,, \tag{52}$$

which can be interpreted as the "norm" of the ML amplitude $\mathcal{A}_{\mathrm{ML}}$ with respect to the "metric" $\mathcal{M}_{\mu\nu}$ [8, 10]

2.2 Non-stationary, non-complete data

In practice we will be computing the power-spectra $S_X(f)$ over shorter time-periods T_{SFT}, corresponding to the "Short Fourier Transforms" (SFT) that are used as input data to (most) CW codes. We therefore only need to assume approximately stationary noise $S_{X\alpha}(f)$ over each SFT α from detector X, allowing the noise-floor to vary from one SFT to the next. Furthermore, data might be available only for some of time during the time-span T, depending on the detector X, and we therefore base all our expressions on these SFTs as the elementary per-detector "data atoms", writing (44) as

$$(\boldsymbol{x}|\boldsymbol{y}) \approx 2\sum_{X=1}^{N_{\mathrm{Det}}}\sum_{\alpha=1}^{N^X_{\mathrm{SFT}}} S^{-1}_{X\alpha}(f)\int_0^{T_{\mathrm{SFT}}} x_{X\alpha}(t)\,y_{X\alpha}(t)\,dt\,, \tag{53}$$

using the convention $x_{X\alpha}(t) \equiv x^X(t_{X\alpha}+t)$, where $t_{X\alpha}$ is the start-time of the SFT $X\alpha$. The number of SFTs from detector X is N^X_{SFT}, i.e.

$$N_{\mathrm{SFT}} = \sum_{X=1}^{N_{\mathrm{Det}}} N^X_{\mathrm{SFT}} = \sum_{X\alpha} 1\,, \tag{54}$$

is the total number of SFTs from all detectors. Here and in the following we use the shorthand notation

$$\sum_{X\alpha} \ldots \equiv \sum_{X=1}^{N_{\mathrm{Det}}} \sum_{\alpha=1}^{N^X_{\mathrm{SFT}}} \ldots\,, \tag{55}$$

to denote the sum over all used SFTs from all detectors. It will be useful to re-normalize the noise factors $S^{-1}_{X\alpha}$ in (53), by introducing *noise weights*

$$w_{X\alpha}(f) \equiv \frac{S^{-1}_{X\alpha}(f)}{\mathcal{S}^{-1}}\,. \tag{56}$$

This will serve two purposes: (i) to make the weights numerically $\sim \mathcal{O}(1)$, and (ii) in order to allow factoring out the overall *scaling* of the scalar product with noise-floors and length of data, with the remaining factors being simple averages. Using these definitions, we can re-write the scalar product (53) as

$$(\boldsymbol{x}|\boldsymbol{y}) \approx 2\mathcal{S}^{-1} \sum_{X\alpha} w_{X\alpha} \int_0^{T_{\mathrm{SFT}}} x_{X\alpha}(t)\, y_{X\alpha}(t)\, dt\,, \tag{57}$$

which is a noise-weighted sum over single-SFT integrals. The noise-weights (56) depend on the frequency f at which they are computed, and in practice we assume $S_{X\alpha}(f)$ to be roughly constant over a small frequency band Δf around the template frequency f_0. The current code (in `LALComputeMultiNoiseWeights`) defines the weights in terms of the *arithmetic mean* of the PSD over Δf of the input SFTs, i.e.

$$w_{X\alpha}(f_0) \approx \frac{\langle S_{X\alpha}(f)\rangle^{-1}_{f_0\pm\Delta f/2}}{\mathcal{S}^{-1}}\,. \tag{58}$$

The normalization constant $\mathcal{S}^{-1}$ is in principle arbitrary and drops out from any physically meaningful result. For practical purposes, however, we choose it in such as way to achieve (i) and (ii) mentioned above, namely

$$\sum_{X\alpha} w_{X\alpha} = N_{\mathrm{SFT}}\,, \quad \text{therefore} \tag{59}$$

$$\mathcal{S}^{-1} \equiv \frac{1}{N_{\mathrm{SFT}}} \sum_{X\alpha} S^{-1}_{X\alpha}\,. \tag{60}$$

Using this convention, $\mathcal{S}$ is defined as the *harmonic mean* over the per-SFT noise PSDs $S_{X\alpha}$ over all SFTs α from all detectors X. These weights have the property that ${N_{\mathrm{SFT}}}^{-1} \sum_{X\alpha} w_{X\alpha} = 1$, and so we can conveniently define a *total noise-weighted average* $\langle x\, y \rangle_w$, namely[2]

$$\langle x\, y \rangle_w \equiv \frac{1}{N_{\mathrm{SFT}}} \sum_{X\alpha} w_{X\alpha} \, \langle x_{X\alpha}\, y_{X\alpha} \rangle_t \,, \tag{61}$$

in terms of single-SFT time-averages $\langle Z_{X\alpha} \rangle_t$ of a function $Z_{X\alpha}(t)$ of time and detector, defined as

$$\langle Z_{X\alpha} \rangle_t \equiv \frac{1}{T_{\mathrm{SFT}}} \int_0^{T_{\mathrm{SFT}}} Z_{X\alpha}(t)\, dt \,. \tag{62}$$

Using this, the scalar product (57) can now be expressed as

$$(\boldsymbol{x}|\boldsymbol{y}) \approx 2\mathcal{S}^{-1} T_{\mathrm{data}} \, \langle x\, y \rangle_w \,, \tag{63}$$

where

$$T_{\mathrm{data}} \equiv N_{\mathrm{SFT}}\, T_{\mathrm{SFT}} \tag{64}$$

is the total time length of data used.
The scalar products involved in the $\mathcal{F}$-statistic contain slowly-varying (diurnal) antenna-pattern functions $\{a(t), b(t)\}$), and phase-functions $\{\sin\phi(t), \cos\phi(t)\}$ that are oscillatory on short timescales $1/f \ll T_{\mathrm{SFT}}$. Using these properties, the 4×4 matrix $\mathcal{M}_{\mu\nu}$ defined in Eq. (49), namely

$$\mathcal{M}_{\mu\nu} \equiv (\boldsymbol{h}_\mu | \boldsymbol{h}_\nu) = \mathcal{S}^{-1} T_{\mathrm{data}}\, m_{\mu\nu} \,, \tag{65}$$

can be approximated to yield the block-form

$$m_{\mu\nu} = 2\langle h_\mu\, h_\nu \rangle_w \approx \begin{pmatrix} A & C & 0 & 0 \\ C & B & 0 & 0 \\ 0 & 0 & A & C \\ 0 & 0 & C & B \end{pmatrix} \,, \tag{66}$$

with the 3 independent components

$$A \equiv \langle a^2 \rangle_w \,, \quad B \equiv \langle b^2 \rangle_w \,, \quad C \equiv \langle a\, b \rangle_w \,, \tag{67}$$

and we define the determinant $D \equiv A\, B - C^2$.

[2]Note that our definition of $\mathcal{S}^{-1}$ and averaging operator $\langle .. \rangle_w$ here differ from the conventions used in [6], which are less symmetric in time and detectors, and less suitable for generalization to varying noise-floors.

Introducing the complex matched filters

$$\begin{aligned} x_{\mathrm{a}} &\equiv x_1 - i x_3 = (\boldsymbol{x}|\boldsymbol{h}_{\mathrm{a}}) \ , \\ x_{\mathrm{b}} &\equiv x_2 - i x_4 = (\boldsymbol{x}|\boldsymbol{h}_{\mathrm{b}}) \ , \end{aligned} \tag{68}$$

in terms of the complex basis (28), and using (65), we can now write the $\mathcal{F}$-statistic (50) more explicitly as

$$2\mathcal{F} = \frac{D^{-1}}{\mathcal{S}^{-1} T_{\mathrm{data}}} \left[B|x_{\mathrm{a}}|^2 + A|x_{\mathrm{b}}|^2 - 2C\Re(x_{\mathrm{a}}\, x_{\mathrm{b}}^*) \right] \ . \tag{69}$$

2.3 $\mathcal{F}$-statistic of perfectly matched signal

Let us assume there is a signal $\boldsymbol{s}(t)$ in the data that is perfectly matched by the search-template, i.e.

$$\begin{aligned} \boldsymbol{x}(t) &= \boldsymbol{n}(t) + \boldsymbol{s}(t), \quad \text{where} \\ \boldsymbol{s}(t) &= \boldsymbol{h}(t; \mathcal{A}_{\mathrm{s}}, \lambda_{\mathrm{s}}) = \mathcal{A}_{\mathrm{s}}^{\mu}\, \boldsymbol{h}_{\mu}(t; \lambda_{\mathrm{s}})\,, \end{aligned} \tag{70}$$

and so the four amplitude-components x_μ, defined in (48), are

$$x_\mu(\mathcal{A}_{\mathrm{s}}, \lambda_{\mathrm{s}}) = n_\mu(\lambda_{\mathrm{s}}) + s_\mu(\mathcal{A}_{\mathrm{s}}, \lambda_{\mathrm{s}})\,, \tag{71}$$

where $n_\mu \equiv (\boldsymbol{n}|\boldsymbol{h}_\mu)$ and

$$s_\mu \equiv (\boldsymbol{s}|\boldsymbol{h}_\mu) = \mathcal{A}_{\mathrm{s}}^{\nu}\, \mathcal{M}_{\nu\mu}(\lambda_{\mathrm{s}})\,. \tag{72}$$

One can show the following identities for zero-mean Gaussian noise $\boldsymbol{n}$:

$$E[n_\mu] = 0\,, \quad \text{and} \quad E[n_\mu\, n_\nu] = \mathcal{M}_{\mu\nu}\,, \tag{73}$$

where in the second equation we used (39). This results in

$$E[x_\mu] = s_\mu\,, \quad \text{and} \quad E[x_\mu\, x_\nu] = \mathcal{M}_{\mu\nu} + s_\mu\, s_\nu\,, \tag{74}$$

which shows that the four random variables x_μ have means s_μ and covariance $\mathcal{M}_{\mu\nu}$ (independent of the signal strength). By applying these relations to Eq. (50), we find the expectation of $2\mathcal{F}$ in the perfectly-matched case as

$$E[2\mathcal{F}] = 4 + \rho^2(0)\,, \tag{75}$$

where we defined the "optimal" signal-to-noise ratio (SNR) $\rho(0)$ as

$$\rho^2(0) \equiv s_\mu\, \mathcal{M}^{\mu\nu}\, s_\nu = \mathcal{A}_{\mathrm{s}}^{\mu}\, \mathcal{M}_{\mu\nu}\, \mathcal{A}_{\mathrm{s}}^{\nu} = (\boldsymbol{s}|\boldsymbol{s}) \ . \tag{76}$$

Combining (26) and (65), (66) this can be written[3] more explicitly as

$$\rho^2(0) = h_0^2\,(\alpha_1\,A + \alpha_2\,B + 2\alpha_3\,C)\;\mathcal{S}^{-1}T_{\rm data}\,, \tag{77}$$

where the functions $\alpha_i(\eta,\psi)$ are defined as (with $\eta \equiv \cos\iota$):

$$\alpha_1(\eta,\psi) \equiv (\hat{\mathcal{A}}^1)^2 + (\hat{\mathcal{A}}^3)^2 = \frac{1}{4}(1+\eta^2)^2\cos^2 2\psi + \eta^2\sin^2 2\psi\,, \tag{78}$$

$$\alpha_2(\eta,\psi) \equiv (\hat{\mathcal{A}}^2)^2 + (\hat{\mathcal{A}}^4)^2 = \frac{1}{4}(1+\eta^2)^2\sin^2 2\psi + \eta^2\cos^2 2\psi\,, \tag{79}$$

$$\alpha_3(\eta,\psi) \equiv \hat{\mathcal{A}}^1\hat{\mathcal{A}}^2 + \hat{\mathcal{A}}^3\hat{\mathcal{A}}^4 = \frac{1}{4}(1-\eta^2)^2\sin 2\psi\,\cos 2\psi\,, \tag{80}$$

using the re-scaled amplitude parameters $\hat{\mathcal{A}}^\mu \equiv \mathcal{A}^\mu/h_0$.

2.4 Average SNR^2

It is often useful to compute averaged quantities over the amplitude parameters $\{\cos\iota,\psi\}$ and sky-position $\vec{n}$. Averaging a quantity Z over $\{\cos\iota,\psi\}$ with isotropic priors on the source-orientation, which translates into uniform priors [7] over $\cos\iota$ and ψ, namely

$$\langle Z\rangle_{\cos\iota,\psi} \equiv \frac{1}{2}\int_{-1}^{1} d\cos\iota\,\frac{1}{\pi/2}\int_{-\pi/4}^{\pi/4} d\psi\,Z(\cos\iota,\psi)\,, \tag{81}$$

yields

$$\langle\alpha_1\rangle_{\cos\iota,\psi} = \langle\alpha_2\rangle_{\cos\iota,\psi} = \frac{2}{5}\,,\quad \langle\alpha_3\rangle_{\cos\iota,\psi} = 0\,. \tag{82}$$

The sky-average of A, B, C is a little more involved. From (67) we see that these antenna-pattern coefficients are time-averages and noise-weighted detector averages of a^2, b^2, and $a\,b$ respectively, with the antenna-pattern functions $a(t;\vec{n})$ and $b(t;\vec{n})$ defined in (15). We can therefore change the order of isotropic sky-averaging, defined as

$$\langle Z\rangle_{\vec{n}} \equiv \frac{1}{4\pi}\int_0^{2\pi} d\alpha\int_{-1}^{1} Z(\vec{n})\,d\sin\delta\,, \tag{83}$$

with time- and detector-averaging (61), i.e. $\langle A\rangle_{\vec{n}} = \langle\langle a^2\rangle_{\vec{n}}\rangle_w$. The all-sky antenna-pattern averages are independent of time and of detector, i.e. $\overline{a^2} \equiv \langle a^2_{X\alpha}\rangle_{\vec{n}}$, $\overline{b^2} \equiv \langle b^2_{X\alpha}\rangle_{\vec{n}}$ and $\overline{ab} \equiv \langle a_{X\alpha}b_{X\alpha}\rangle_{\vec{n}}$, and we therefore obtain

$$\langle A\rangle_{\vec{n}} = \langle\langle a^2\rangle_w\rangle_{\vec{n}} = \langle\langle a^2\rangle_{\vec{n}}\rangle_w = \overline{a^2}\,\langle 1\rangle_w = \overline{a^2}\,, \tag{84}$$

[3] The difference to Eq. (68) of [6] is the use of single-sided noise PSD, and the different definitions of $\mathcal{S}^{-1}$ and averaging operator

and similarly for B and C. For an interferometer with orthogonal arms we can simply choose $\hat{l}_1 = (1,0,0)$ and $\hat{l}_2 = (0,1,0)$, and combining (10), (5), (6) and (15), we find

$$2a_0 = \hat{\xi}^1\hat{\xi}^1 - \hat{\xi}^2\hat{\xi}^2 - \hat{\eta}^1\hat{\eta}^1 + \hat{\eta}^2\hat{\eta}^2\,, \quad b_0 = \hat{\xi}^1\hat{\eta}^1 - \hat{\xi}^2\hat{\eta}^2\,. \tag{85}$$

Inserting the explicit expressions (17) for $\hat{\xi}, \hat{\eta}$ as a function of skyposition, we obtain

$$\begin{aligned} a_0^2 &= \frac{1}{4}\left(\sin^2\alpha - \cos^2\alpha\right)^2\left(1 + \sin^2\delta\right)^2\,, \\ b_0^2 &= \sin^2 2\alpha\,\sin^2\delta\,, \end{aligned} \tag{86}$$

which are easily integrated, yielding

$$\begin{aligned} \langle A\rangle_{\vec{n}} &= \langle a_0^2\rangle_{\vec{n}} = \frac{7}{30}\,, \\ \langle B\rangle_{\vec{n}} &= \langle b_0^2\rangle_{\vec{n}} = \frac{1}{6}\,, \\ \langle C\rangle_{\vec{n}} &= \langle a_0\,b_0\rangle_{\vec{n}} = 0\,. \end{aligned} \tag{87}$$

and so $\langle A\rangle_{\vec{n}} + \langle B\rangle_{\vec{n}} = \frac{2}{5}$.
Equipped with these averages, we now obtain from (77)

$$\langle\rho^2\rangle_{\cos\iota,\psi} = \frac{2}{5}h_0^2\,(A+B)\;\mathcal{S}^{-1}T_{\rm data}\,, \tag{88}$$

$$\langle\rho^2\rangle_{\vec{n}} = h_0^2\left(\frac{7}{30}\alpha_1 + \frac{1}{6}\alpha_2\right)\;\mathcal{S}^{-1}T_{\rm data}\,, \tag{89}$$

$$\langle\rho^2\rangle_{\vec{n},\cos\iota,\psi} = \frac{4}{25}\,h_0^2\,\mathcal{S}^{-1}T_{\rm data}\,, \tag{90}$$

in agreement with Eq.(93) in [4].
It is sometimes convenient to express the instantaneous "strength" of a signal in the detectors, independently of the observation time and noise floor, and following [2] we define $h_{\rm rms}$ as the root-mean-square (rms) of the signal strain, averaged over time and detectors, i.e.

$$h_{\rm rms}^2 \equiv \langle s^2\rangle_w = \mathcal{A}_{\rm s}^\mu\langle h_\mu h_\nu\rangle_w\mathcal{A}_{\rm s}^\nu = \frac{1}{2}\mathcal{A}_{\rm s}^\mu\,m_{\mu\nu}\,\mathcal{A}_{\rm s}^\nu\,, \tag{91}$$

in terms of the antenna-pattern matrix $m_{\mu\nu}$ defined in (66). Using this definition and (65), the optimal SNR (76) can now also be written as

$$\rho^2 = 2\mathcal{S}^{-1}T_{\rm data}\,h_{\rm rms}^2\,, \tag{92}$$

and comparing this to (77) we obtain

$$h^2_{\rm rms} = \frac{1}{2} h_0^2 \, \left(\alpha_1 \, A + \alpha_2 \, B + 2\alpha_3 \, C \right) \, . \tag{93}$$

Averaging this over all sky-positions $\vec{n}$ and polarization angles $\cos\iota, \psi$ at fixed amplitude h_0, we find

$$\langle h^2_{\rm rms} \rangle_{\cos\iota,\psi,\vec{n}} = \frac{2}{25} \, h_0^2 \, , \tag{94}$$

in agreement with the expression found in [2].

3 Parameter estimation of the signal

3.1 Estimating amplitude parameters $\{h_0, \cos\iota, \psi, \phi_0\}$

From the expression (51) for the maximum-likelihood amplitudes $\mathcal{A}^\mu_{\rm ML}$ in terms of the measured x_μ, we can infer the signal-parameters $A_+, A_\times$ (or equivalently h_0, $\cos\iota$) and ψ, ϕ_0, by using (26) and (30), mostly following Yousuke's notes. We want to invert the four relations $\mathcal{A}^\mu(h_0, \cos\iota, \psi, \phi_0)$ in Eq. (26), and we start by computing the two quantities

$$A_s^2 \;\equiv\; \sum_{\mu=1}^{4} (\mathcal{A}^\mu)^2 = A_+^2 + A_\times^2 \, , \tag{95}$$

$$D_a \;\equiv\; \mathcal{A}^1 \mathcal{A}^4 - \mathcal{A}^2 \mathcal{A}^3 = A_+ \, A_\times \, , \tag{96}$$

which can easily be solved for A_+, $A_\times$, namely

$$2 A_{+,\times}^2 = A_s^2 \pm \sqrt{A_s^4 - 4 D_a^2} \, , \tag{97}$$

where our convention here is $|A_+| \geq |A_\times|$, cf. (30), and therefore the '+' solution is A_+, and the $'-'$ is $A_\times$. The sign of A_+ is always positive by convention (30), while the sign of $A_\times$ is given by the sign of D_a, as can be seen from (96). Note that the discriminant in (97) is also expressible as

$${\rm disc} \equiv \sqrt{A_s^4 - 4 \, D_a^2} = A_+^2 - A_\times^2 \geq 0 \, . \tag{98}$$

Having computed $A_+, A_\times$, we can now also obtain ψ and ϕ_0, namely defining $\beta \equiv A_\times / A_+$, and

$$b_1 \;\equiv\; \mathcal{A}^4 - \beta \mathcal{A}^1 \, , \tag{99}$$

$$b_2 \;\equiv\; \mathcal{A}^3 + \beta \mathcal{A}^2 \, , \tag{100}$$

$$b_3 \;\equiv\; \beta \mathcal{A}^4 - \mathcal{A}^1 \, , \tag{101}$$

we easily find

$$\psi = \frac{1}{2}\text{atan}\left(\frac{b_1}{b_2}\right) . \tag{102}$$

and

$$\phi_0 = \text{atan}\left(\frac{b_2}{b_3}\right) . \tag{103}$$

Note that there is still an overall sign-ambiguity in the amplitudes $\mathcal{A}^\mu$, which can be determined as follows: compute a 'reconstructed' $\mathcal{A}^1_r$ from (26) using the estimates $A_{+,\times}$ and ψ, ϕ_0, and compare its sign to the original estimate $\mathcal{A}^1$ of (130). If the sign differs, the correct solution is simply found by replacing $\phi_0 \rightarrow \phi_0 + \pi$.
Converting $A_+, A_\times$ into h_0 and $\mu \equiv \cos\iota$ is done by solving (30), which yields

$$h_0 = A_+ + \sqrt{A_+^2 - A_\times^2} , \tag{104}$$

where we only kept the '+' solution, as we must have $h_0 \geq A_+$ (which can be seen from (30)). Finally, $\mu = \cos\iota$ is simply given by $\cos\iota = A_\times / h_0$.

3.2 Errors in amplitude-parameter estimation

Let us define the error $\Delta\mathcal{A}^\mu$ in maximum-likelihood parameter estimation on the four amplitude $\mathcal{A}^\mu$ simply as

$$\Delta\mathcal{A}^\mu \equiv \mathcal{A}^\mu_{\text{ML}} - \mathcal{A}^\mu_{\text{s}} . \tag{105}$$

Given (51), (71) and (72), we have

$$\mathcal{A}^\mu_{\text{ML}} = \mathcal{M}^{\mu\nu}\, n_\nu + \mathcal{A}^\nu_{\text{s}} , \tag{106}$$

and therefore

$$\Delta\mathcal{A}^\mu = \mathcal{M}^{\mu\nu}\, n_\nu , \tag{107}$$

and so we directly obtain using (73)

$$E[\mathcal{A}^\mu_{\text{ML}}] = \mathcal{A}^\mu_{\text{s}} , \quad \text{i.e.} \quad E[\Delta\mathcal{A}^\mu] = 0 , \tag{108}$$

namely the ML estimators for the $\mathcal{A}^\mu$ are *unbiased*. Furthermore, the covariance matrix of the errors $\Delta\mathcal{A}^\mu$ is found as

$$E[\Delta\mathcal{A}^\mu\, \Delta\mathcal{A}^\nu] = \mathcal{M}^{\mu\nu} , \tag{109}$$

which corresponds to the Cramér-Rao bound, where $\mathcal{M}^{\mu\nu}$ is the inverse of the Fisher matrix (49). Note that we have not made any assumptions about the errors $\Delta\mathcal{A}^\mu$ being "small", the Fisher-matrix relation (109) is strictly true here for any deviations and SNR, *provided* the $\mathcal{A}^\mu_{\mathrm{ML}}$ were measured at exactly the right signal Doppler location λ_{s}, such that $\mathcal{M}_{\mu\nu} = \mathcal{M}_{\mu\nu}(\lambda_{\mathrm{s}})$. Any parameter-estimation error in λ would complicate the picture, which is why these error-estimates strictly only apply in a perfectly-matched ("targeted") search case.

Let us now consider arbitrary functions $f_i(\mathcal{A}^\mu)$ of the four amplitudes $\mathcal{A}^\mu$, where for *small* errors df_i we have

$$df_i = \partial_\mu f_i \, d\mathcal{A}^\mu \,, \tag{110}$$

and therefore we can find the error covariances

$$E[df_i \, df_j] = \partial_\mu f_i \, \mathcal{M}^{\mu\nu} \, \partial_\nu f_j \,. \tag{111}$$

We can consider different more "physical" amplitude-parameter coordinates such as $\mathcal{A}^{\hat{i}} \equiv (A_+, A_\times, \phi_0, \psi)$ or $\mathcal{A}^i \equiv (h_0, \cos\iota, \phi_0, \psi)$. From (26) one easily obtains the explicit Jacobian

$$J^\mu{}_{\hat{i}} \equiv \frac{\partial \mathcal{A}^\mu}{\partial \mathcal{A}^{\hat{i}}} = \begin{pmatrix} \cos\phi_0 \, \cos 2\psi & -\sin\phi_0 \, \sin 2\psi & \mathcal{A}^3 & -2\,\mathcal{A}^2 \\ \cos\phi_0 \, \sin 2\psi & \sin\phi_0 \, \cos 2\psi & \mathcal{A}^4 & 2\,\mathcal{A}^1 \\ -\sin\phi_0 \, \cos 2\psi & -\cos\phi_0 \, \sin 2\psi & -\mathcal{A}^1 & -2\,\mathcal{A}^4 \\ -\sin\phi_0 \, \sin 2\psi & \cos\phi_0 \, \cos 2\psi & -\mathcal{A}^2 & 2\,\mathcal{A}^3 \end{pmatrix} \tag{112}$$

and by (numerical) inversion we can obtain $\partial_\mu \mathcal{A}^{\hat{i}} = J^{-1\,\hat{i}}{}_\mu$. We therefore can compute the covariance matrix of errors $d\mathcal{A}^{\hat{i}}$ from (111), namely

$$E\left[d\mathcal{A}^{\hat{i}} \, d\mathcal{A}^{\hat{j}}\right] = J^{-1\,\hat{i}}{}_\mu \, J^{-1\,\hat{j}}{}_\nu \, \mathcal{M}^{\mu\nu} \,. \tag{113}$$

Similarly, for the choice of output-variables $\mathcal{A}^i$, using (30) we find

$$\frac{\partial \mathcal{A}^\mu}{\partial h_0} = \frac{\mathcal{A}^\mu}{h_0} \,, \quad \frac{\partial \mathcal{A}^\mu}{\partial \cos\iota} = B^\mu \,, \tag{114}$$

where we defined

$$B^\mu \equiv A_\times \frac{\partial \mathcal{A}^\mu}{\partial A_+} + h_0 \frac{\partial \mathcal{A}^\mu}{\partial A_\times} = \{\mathcal{A}^\mu \,|\text{replace}\, A_\times \mapsto h_0, \, A_+ \mapsto A_\times\} \,, \tag{115}$$

and so we obtain the corresponding Jacobian

$$J^\mu{}_i \equiv \frac{\partial \mathcal{A}^\mu}{\partial \mathcal{A}^i} = \begin{pmatrix} \mathcal{A}^1/h_0 & B^1 & \mathcal{A}^3 & -2\,\mathcal{A}^2 \\ \mathcal{A}^2/h_0 & B^2 & \mathcal{A}^4 & 2\,\mathcal{A}^1 \\ \mathcal{A}^3/h_0 & B^3 & -\mathcal{A}^1 & -2\,\mathcal{A}^4 \\ \mathcal{A}^4/h_0 & B^4 & -\mathcal{A}^2 & 2\,\mathcal{A}^3 \end{pmatrix} \tag{116}$$

and we can obtain the covariance matrix of small errors $d\mathcal{A}^i$ as

$$E\left[d\mathcal{A}^i\, d\mathcal{A}^j\right] = {J^{-1}}^i{}_\mu\, {J^{-1}}^j{}_\nu\, \mathcal{M}^{\mu\nu}\,. \tag{117}$$

Note, however, that both (113) and (117) are only valid in the limit of *small* errors d (i.e. the high-SNR limit), and are potentially subject to singularities in the coordinate transformations, i.e. (112) (116). The formulation (109) in "canonical" coordinates $\mathcal{A}^\mu$ is generally true at any SNR and is always well-defined.

4 Practical computation in CFS_v2

4.1 Data normalization and antenna weighting

The expectation value of the $\mathcal{F}$-statistic is $E[2\mathcal{F}] = 4 + \mathrm{SNR}^2$. For practical and numerical convenience, we want to make all quantities involved in computing $\mathcal{F}$ of order $\mathcal{O}(1)$. This is already the case for the antenna-pattern functions $\{A, B, C\}$, defined in (67). However, the scale of the input data $x^X(t)$ is vastly different, namely from the Wiener-Khinchin theorem we can estimate[4] the (single-sided) PSDs $S_{X\alpha}(f)$ as

$$E\left[|\widetilde{x}_{X\alpha}(f)|^2\right] \approx \frac{1}{2} T_{\mathrm{SFT}}\, S_{X\alpha}(f) \sim \mathcal{O}\left(10^{-44} s^2\right)\,, \tag{118}$$

where $\widetilde{x}_{X\alpha}(f)$ is the "Short Fourier transform" (SFT), defined as

$$\widetilde{x}_{X\alpha}(f) = \int_0^{T_{\mathrm{SFT}}} x_{X\alpha}(t)\, e^{-i2\pi f t}\, dt = T_{\mathrm{SFT}}\, \langle x_{X\alpha}(t)\, e^{-i2\pi f t}\rangle_t\,. \tag{119}$$

Therefore, if we re-normalize[5] the data as (`LALNormalizeMultiSFTVect()`[6]):

$$\widetilde{y}_{X\alpha}(f) \equiv \frac{\widetilde{x}_{X\alpha}(f)}{\sqrt{\frac{1}{2} T_{\mathrm{SFT}} S_{X\alpha}(f)}} \approx \frac{\widetilde{x}_{X\alpha}(f)}{\sqrt{E[|\widetilde{x}_{X\alpha}(f)|^2]}}\,, \tag{120}$$

then $E[|\widetilde{y}_{X\alpha}(f)|^2] = 1$ and therefore $\widetilde{y}_{X\alpha} \sim \mathcal{O}(1)$. Note, however, that in practice we *estimate* $E[|\widetilde{x}_{X\alpha}(f)|^2]$ from the median of a finite number of

[4] This is the basis for estimating the noise PSD in the function `LALNormalizeSFT()`.

[5] In the special `--SignalOnly` case the `CFS_v2` code does not try to normalize the data and instead *assumes* the (single-sided) noise-power to be $S_{X\alpha} = 1$. The "missing" normalization-factor of $\sqrt{T_{\mathrm{SFT}}/2}$ is then applied to $F_{\{a,b\}}$ a-posteriori.

[6] There is a small inconsistency here: in the definition of the noise-weights (58), we used the frequency-averaged $\langle S_{X\alpha}\rangle_{\Delta f}$ over the Band Δf of the SFT, while in the data-normalization (120) we use the per-bin values of $S_{X\alpha}(f)$.

neighboring bins. The fluctuations in this noise-floor estimator introduce a bias in (120), namely $E[|\widetilde{y}_{X\alpha}(f)|^2] \gtrsim 1$, resulting in a bias in $\mathcal{F}$, namely $E[2\mathcal{F}] \gtrsim 4$ in pure noise. Substituting (120) into (68) using the scalar product (57), we find

$$x_{\rm a} = \sqrt{2\mathcal{S}^{-1}T_{\rm SFT}}\,\sum_{X\alpha}\sqrt{w_{X\alpha}}\int_0^{T_{\rm SFT}} y_{X\alpha}(t)\,a_{X\alpha}(t)\,e^{-i\phi_{X\alpha}(t)}\,dt\,, \tag{121}$$

and similarly for $x_{\rm b}$. Furthermore, expanding (67) into

$$A \equiv \langle a^2\rangle_w = \frac{1}{N_{\rm SFT}}\sum_{X\alpha} w_{X\alpha}\,\langle a^2_{X\alpha}\rangle_t\,, \tag{122}$$

we see that we can completely absorb the noise-weights $w_{X\alpha}$ into $\{a_{X\alpha}(t), b_{X\alpha}(t)\}$, namely by defining *noise-weighted* antenna-pattern functions

$$\widehat{a}_{X\alpha}(t) \equiv \sqrt{w_{X\alpha}}\;a_{X\alpha}(t)\,,\quad \widehat{b}_{X\alpha}(t) \equiv \sqrt{w_{X\alpha}}\;b_{X\alpha}(t)\,. \tag{123}$$

We can now write

$$x_{\{\rm a,b\}} = \sqrt{2\mathcal{S}^{-1}T_{\rm SFT}}\;F_{\{\rm a,b\}}\,, \tag{124}$$

$$\{A,B,C\} = \frac{1}{N_{\rm SFT}}\,\{\hat{A},\hat{B},\hat{C}\}\,, \tag{125}$$

introducing the quantities $F_{\{\rm a,b\}}$ and $\{\hat{A},\hat{B},\hat{C}\}$ that are used in the `CFS_v2` code, and which are defined as

$$F_{\{\rm a,b\}} \equiv \sum_{X\alpha} F^{X\alpha}_{\{\rm a,b\}}\,, \tag{126}$$

$$F^{X\alpha}_{\rm a} \equiv \int_0^{T_{\rm SFT}} y_{X\alpha}(t)\,\widehat{a}_{X\alpha}(t)\,e^{-i\phi_{X\alpha}(t)}\,dt\,,\quad F^{X\alpha}_{\rm b} = \ldots(\widehat{a}\mapsto\widehat{b}) \tag{127}$$

$$\hat{A} \equiv \sum_{X\alpha}\langle\widehat{a}^2_{X\alpha}\rangle_t\,,\quad \hat{B} \equiv \sum_{X\alpha}\langle\widehat{b}^2_{X\alpha}\rangle_t\,,\quad \hat{C} \equiv \sum_{X\alpha}\langle\widehat{a}_{X\alpha}\widehat{b}_{X\alpha}\rangle_t\,, \tag{128}$$

Inserting (124)(125) into (69), we obtain

$$2\mathcal{F} = \frac{2}{\hat{D}}\left[\hat{B}\,|F_{\rm a}|^2] + \hat{A}\,|F_{\rm b}|^2 - 2\hat{C}\,\Re\left(F_{\rm a}\,F^*_{\rm b}\right)\right]\,, \tag{129}$$

with $\hat{D} \equiv \hat{A}\,\hat{B} - \hat{C}^2$. We can express the maximum-likelihood estimators (51) for the amplitudes $\mathcal{A}^\mu$ explicitly as

$$\mathcal{A}^\mu_{\rm ML} = \mathcal{M}^{\mu\nu}\,x_\mu = \frac{\sqrt{2}\hat{D}^{-1}}{\sqrt{\mathcal{S}^{-1}T_{\rm SFT}}}\begin{pmatrix} \hat{B}\,F^\Re_{\rm a} - \hat{C}\,F^\Re_{\rm b} \\ -\hat{C}\,F^\Re_{\rm a} + \hat{A}\,F^\Re_{\rm b} \\ -\hat{B}\,F^\Im_{\rm a} + \hat{C}\,F^\Im_{\rm b} \\ \hat{C}\,F^\Im_{\rm a} - \hat{A}\,F^\Im_{\rm b}\end{pmatrix}\,, \tag{130}$$

with $F^{\Re}_{\{a,b\}} \equiv \Re F_{\{a,b\}}$, and $F^{\Im}_{\{a,b\}} \equiv \Im F_{\{a,b\}}$.
We see from (126)–(129) that the $\mathcal{F}$-statistic is computed completely from the set of per-SFT "$\mathcal{F}$-atoms"

$$\{F^{X\alpha}_{\{a,b\}},\, \langle\widehat{a}^2_{X\alpha}\rangle_t,\, \langle\widehat{b}^2_{X\alpha}\rangle_t,\, \langle\widehat{a}_{X\alpha}\widehat{b}_{X\alpha}\rangle_t\}\,. \tag{131}$$

These "$\mathcal{F}$-atoms" are also the primary input to `CFS_v2` for the transient-CW search over different start-times and durations, as described in [9].

4.2 The Williams-Schutz approximation ("LALDemod")

This section is originally based on Xavie's LALDemod-notes[7], and the method is largely based on [11].
With the convention introduced in (53), the (normalized) data time-series corresponding to an SFT $X\alpha$ of duration $T_{\rm SFT}$ is written as

$$y_{X\alpha j} = y_{X\alpha}(t_j) = y(t_{X\alpha} + j\Delta t)\,, \tag{132}$$

where $j = 0, ..., N-1$ such that $T_{\rm SFT} = N\Delta t$, and $t_{X\alpha}$ is the start-time of the SFT $X\alpha$. As noted above, all components of $\mathcal{F}$ entering (129), namely $F_{\{a,b\}}$ and $\{\hat{A}, \hat{B}, \hat{C}\}$ are *sums* over per-SFT "$\mathcal{F}$-atoms" (131). Here we focus on the calculation of the atoms $F^{X\alpha}_{\{a,b\}}$, and in order simplify the notation we drop the SFT-index $X\alpha$ from most of the following expressions, which refer to quantities evaluated for a single SFT $X\alpha$. The frequency-domain SFT data is computed from the discretized version of (119), namely

$$\widetilde{y}_k \equiv \Delta t \sum_{j=0}^{N-1} y_j\, e^{-i2\pi\, kj/N}\,, \tag{133}$$

which is exactly what is stored in an SFT-file according to the "SFT-v2" specification (LIGO-T040164-01-Z), where in practice we only store the first $\lfloor N/2 \rfloor$ frequency-bins, as for real y_j we have $\widetilde{y}_{N-k|N} = \widetilde{y}^*_k$. The inverse operation to (133) is

$$y_j = \Delta f \sum_{k=0}^{N-1} \widetilde{y}_k\, e^{i2\pi\, kj/N}\,. \tag{134}$$

We write the discretized version of (127) as

$$F^{X\alpha}_{a} = \Delta t \sum_{j=0}^{N-1} y_j\, \widehat{a}_j\, e^{-i2\pi\,\varphi_j}\,, \tag{135}$$

[7] www.lsc-group.phys.uwm.edu/˜siemens/demod.pdf

where we defined $2\pi\varphi_j \equiv \phi(t_j)$ for later convenience.
The typical SFT-duration (e.g. half an hour) is chosen to be short compared to the variability of the signal, and so we can approximate the antenna-pattern functions as nearly constant over this period. Writing the SFT-midpoint as $t_{\frac{1}{2}} \equiv T_{\rm SFT}/2$, we approximate $\widehat{a}_j \approx \widehat{a} \equiv \widehat{a}(t_{\frac{1}{2}})$. Using this and the inverse DFT (134), we can write (135) as

$$F_{\rm a}^{X\alpha} \approx \widehat{a}\,\Delta f \Delta t \sum_{j=0}^{N-1} e^{-i2\pi\,\varphi_j} \sum_{k=0}^{N-1} \widetilde{y}_k\, e^{i2\pi\, jk/N}\,. \tag{136}$$

The phase-evolution of a typical continuous pulsar-signal is dominated by the linear term $\phi(t) \approx 2\pi f\, t$, and we approximate it by a first-order expansion around the SFT-midpoint, namely

$$\varphi_j = \varphi_{\frac{1}{2}} + \dot{\varphi}_{\frac{1}{2}}\, T_{\rm SFT} \left(\frac{j}{N} - \frac{1}{2}\right) + \mathcal{O}(2)\ . \tag{137}$$

Using this expansion, (136) now reads as

$$F_{\rm a}^{X\alpha} \approx \widehat{a}\,\Delta t\,\Delta f\, e^{-i2\pi\,\lambda} \sum_{k=0}^{N-1} \widetilde{y}_k \sum_{j=0}^{N-1} e^{-i2\pi\,\kappa(k)\,j/N}\,, \tag{138}$$

where we defined

$$\begin{aligned} \lambda &\equiv \varphi_{\frac{1}{2}} - \frac{1}{2}\dot{\varphi}_{\frac{1}{2}}\, T_{\rm SFT}\,,\\ \kappa(k) &\equiv \dot{\varphi}_{\frac{1}{2}}\, T_{\rm SFT} - k\,. \end{aligned} \tag{139}$$

The last sum in (138) is simply a geometrical series, and so we find

$$\begin{aligned} \sum_{j=0}^{N-1} e^{-i2\pi\,\kappa\, j/N} &= \frac{1 - e^{-i2\pi\,\kappa}}{1 - e^{-i2\pi\,\kappa/N}} \\ &\overset{N\gg 1}{\approx} \frac{N}{2\pi}\left(\frac{\sin 2\pi\kappa}{\kappa} + i\,\frac{\cos 2\pi\kappa - 1}{\kappa}\right) \\ &\equiv \frac{N}{2\pi}\, P\left(\kappa(k)\right) = \frac{N}{2\pi}\, P_k\,. \end{aligned} \tag{140}$$

The function $P(\kappa)$ is sometimes called "Dirichlet kernel", and it has the property of being strongly peaked around $\kappa = 0$, and so we can truncate the sum over k in (138) to a few terms Δk (referred to as `Dterms` in the code) on either side of k^*, corresponding the to the frequency bin closest to the maximum of $P(\kappa)$, i.e. the bin closest to the solution of $\kappa(k) = 0$, namely

$$k^* \equiv \text{round}\left[\dot{\varphi}_{\frac{1}{2}}\, T_{\text{SFT}}\right] = \text{round}\left[\hat{f}(t_{\frac{1}{2}})/\Delta f\right]\,, \tag{141}$$

where $\hat{f}(t)$ is the "effective" signal-frequency in the detector frame at time t (the time-derivative $\dot{\varphi}$ refers to the time in the detector-frame!), which shows that generally we'll have $k^* \gg 1$. With this approximation we finally find

$$F_{\text{a}}^{X\alpha} \approx \frac{1}{2\pi}\,\widehat{a}\,e^{-i2\pi\,\lambda} \sum_{k=k^*-\Delta k}^{k^*+\Delta k} \widetilde{y}_k\, P_k\,. \tag{142}$$

We'll also need explicit expressions for $\varphi_{\frac{1}{2}}$ and $\dot{\varphi}_{\frac{1}{2}}$ in order to compute λ and $\kappa(k)$, defined in (139). For this we need the timing-function $\tau(t)$, which translates detector arrival times t to the SSB τ. In the purely Newtonian approximation this would be given by (23), but in general the code uses a full ephemeris-based relativistic timing model $\tau(t)$ (in `LALBarycenter()`). Given this function, we define

$$\Delta\tau_{\frac{1}{2}} \equiv \tau(t_{\frac{1}{2}}) - \tau_{\text{ref}}\,, \tag{143}$$

$$\dot{\tau}_{\frac{1}{2}} \equiv \left.\frac{d\tau}{dt}\right|_{t_{\frac{1}{2}}} \quad (\approx 1 + \vec{v}_{\frac{1}{2}}\cdot\vec{n}/c)\,, \tag{144}$$

(which are referred to as `(Multi)SSBtimes` in the code), and so the (full) phase-model (22) yields

$$\varphi_{\frac{1}{2}} = \sum_s \frac{f^{(s)}}{(s+1)!}\,\Delta\tau_{\frac{1}{2}}^{s+1}\,, \tag{145}$$

$$\dot{\varphi}_{\frac{1}{2}} = \dot{\tau}_{\frac{1}{2}} \sum_s \frac{f^{(s)}}{s!}\Delta\tau_{\frac{1}{2}}^{s}\,. \tag{146}$$

4.3 Efficient computation of the "atoms" $F_{\{\text{a,b}\}}^{X\alpha}$

The computation of (142) will be the most time-consuming part in this code, in particular the "hot loop" which is the sum over k. It is therefore important to compute these sums in the most efficient way possible.
First it will be convenient to relabel this sum using $l(k) \equiv k - k_0$ with $k_0 \equiv k^* - \Delta k$ being the leftmost bin in the sum, and so we write

$$\kappa_l \equiv \kappa\left(k(l)\right) = \kappa_0 - l\,, \tag{147}$$

where

$$\kappa_0 \equiv \text{rem}\left(\dot{\varphi}_{\frac{1}{2}} T_{\text{SFT}}\right) + \Delta k \,, \tag{148}$$

and where we defined the "remainder"

$$\text{rem}(x) \equiv x - \text{round}\,[x] \ . \tag{149}$$

Next we note that

$$\begin{aligned} \sin 2\pi\kappa_l = \sin 2\pi\kappa_0 &\equiv s & (150)\\ \cos 2\pi\kappa_l - 1 = \cos 2\pi\kappa_0 - 1 &\equiv c\,, & (151) \end{aligned}$$

and so the Dirichlet-kernel (140) has the form

$$P_{k(l)} = \frac{s}{\kappa_0 - l} + i\frac{c}{\kappa_0 - l} \,. \tag{152}$$

Now let us look at the "hot loop" in (142), which we can express as

$$\chi \equiv \sum_{k=k_0}^{k_0+\mathcal{N}} \widetilde{y}_k P_k = [s\,U - c\,V] + i\ [c\,U + s\,V] \ , \tag{153}$$

where $\mathcal{N} \equiv 2\Delta k - 1$, and the two sums we need to evaluate are

$$U \equiv \sum_{l=0}^{\mathcal{N}} \frac{u_l}{p_l}\,, \quad V \equiv \sum_{l=0}^{\mathcal{N}} \frac{v_l}{p_l}\,, \tag{154}$$

with the further definitions

$$\begin{aligned} p_l &\equiv \kappa_0 - l\,, & (155)\\ u_l &\equiv \Re(\widetilde{y}_{k_0+l})\,, & (156)\\ v_l &\equiv \Im(\widetilde{y}_{k_0+l})\,. & (157) \end{aligned}$$

The above sums (154) are numerically not efficient as they consist of many divisions, which are slower than multiplications. This can be remedied with a clever algorithm suggested by Fekete Ãkos: bringing the sums on a common denominator $q_{\mathcal{N}}$, we get

$$U = \frac{S_{\mathcal{N}}}{q_{\mathcal{N}}}\,, \quad V = \frac{T_{\mathcal{N}}}{q_{\mathcal{N}}}\,, \tag{158}$$

where

$$\begin{aligned} S_{\mathcal{N}} &= u_0\,p_1 p_2 ... p_{\mathcal{N}} + p_0\,u_1\,p_1 ... p_{\mathcal{N}} + ... + p_0 p_1 ... p_{\mathcal{N}-1}\,u_{\mathcal{N}}\,, & (159)\\ T_{\mathcal{N}} &= v_0\,p_1 p_2 ... p_{\mathcal{N}} + p_0\,v_1\,p_1 ... p_{\mathcal{N}} + ... + p_0 p_1 ... p_{\mathcal{N}-1}\,v_{\mathcal{N}}\,, & (160)\\ q_{\mathcal{N}} &= p_0 p_1 ... p_{\mathcal{N}}\,, & (161) \end{aligned}$$

reducing the $2\mathcal{N}+2$ divisions to only 2. The required three components $S_\mathcal{N}$, $T_\mathcal{N}$ and $q_\mathcal{N}$ can be computed efficiently using the following recurrence:

$$\begin{aligned} S_n &= p_n S_{n-1} + q_{n-1} u_n \,, \qquad (162)\\ T_n &= p_n T_{n-1} + q_{n-1} v_n \,, \qquad (163)\\ p_n &= p_{n-1} - 1 \,, \qquad (164)\\ q_n &= p_n \, q_{n-1} \,, \qquad (165) \end{aligned}$$

and the starting conditions

$$\begin{aligned} S_0 &= u_0 \,, \qquad (166)\\ T_0 &= v_0 \,, \qquad (167)\\ p_0 &= \kappa_0 \,, \qquad (168)\\ q_0 &= p_0 \,. \qquad (169) \end{aligned}$$

The number of floating-point operations per iteration is 8, so in total we need $8\mathcal{N}+8$ operations (not counting one sin/cos), of which only 2 are divisions. In the previous "LALDemod" algorithm (e.g `ComputeFstat.c:1.19`) χ was computed more directly resulting in $12\mathcal{N}$ floating point operations, of which $\mathcal{N}$ are divisions!

References

[1] C. Cutler and B. F. Schutz. The generalized $\mathcal{F}$-statistic: multiple detectors and multiple gravitational wave pulsars. *Phys. Rev. D.*, 72:063006, 2005.

[2] C. Cutler, I. Gholami, and B. Krishnan. Improved stack-slide searches for gravitational-wave pulsars. *Phys. Rev. D.*, 72:042004, Aug. 2005. doi: 10.1103/PhysRevD.72.042004.

[3] L. S. Finn. Detection, measurement, and gravitational radiation. *Phys. Rev. D.*, 46:5236–5249, 1992.

[4] 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.

[5] A. Królak, M. Tinto, and M. Vallisneri. Optimal filtering of the LISA data. *Phys. Rev. D.*, 70:022003, 2004.

[6] R. Prix. Search for continuous gravitational waves: metric of the multi-detector $\mathcal{F}$-statistic. *Phys. Rev. D.*, 75:023004, 2007.

[7] R. Prix and B. Krishnan. Targeted search for continuous gravitational waves: Bayesian versus maximum-likelihood statistics. *Class. Quant. Grav.*, 26(20):204013–+, Oct. 2009.

[8] R. Prix and J. T. Whelan. $\mathcal{F}$-statistic search for white-dwarf binaries in the first Mock LISA Data Challenge. *Class. Quant. Grav.*, 24:565, 2007.

[9] R. Prix, S. Giampanis, and C. Messenger. Search method for long-duration gravitational-wave transients from neutron stars. *Phys. Rev. D.*, 2011. (eprint arXiv:1104.1704).

[10] J. T. Whelan, R. Prix, and D. Khurana. Searching for galactic white-dwarf binaries in mock LISA data using an $\mathcal{F}$-statistic template bank. *Class. Quant. Grav.*, 27(5):055010–+, Mar. 2010.

[11] P. R. Williams and B. F. Schutz. An Efficient Matched Filtering Algorithm for the Detection of Continuous Gravitational Wave Signals. 1999. (eprint gr-qc/9912029).