Notes on likelihood heterodyning

A Preprint

Alex Leviyev[1] and Aaron Zimmerman[1]

[1]Center for Gravitational Physics, University of Texas at Austin

February 27, 2023

Abstract

This is a set of notes which explores in more detail the strategy of reducing the computational cost of evaluating the gravitational wave likelihood with a strategy known as heterodyning. Our aim is to provide concrete results and algorithms to supplement the ideas presented in the original paper.

1 Relative binning

1.1 Motivation

In GW data analysis, we work with a complex inner product space $(\mathbb{C}^n, \langle\cdot,\cdot\rangle_{\mathrm{GW}})$, where for $h, g \in \mathbb{C}^n$, $\langle\cdot,\cdot\rangle_{\mathrm{GW}} : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ is a sesquilinear inner product defined by:

$$\langle h, g\rangle_{\mathrm{GW}} \coloneqq 4\sum_{i=1}^{n} \frac{\tilde{h}(f_i)^* \tilde{g}(f_i)}{S(f_i)} \Delta f, \tag{1}$$

where $\Delta f = 1/T$, where T is the duration of the signal, $\tilde{h} \coloneqq \mathrm{DFT}(h)$, and $S(f)$ is a function which weighs each contribution to the sum (called the *power spectral density.*) Recall that a sesquilinear inner product satisfies the following three axioms:

- $\langle f, g\rangle = \langle g, f\rangle^*$, (conjugate symmetry)
- $\langle f, g + \alpha h\rangle = \langle f, g\rangle + \alpha\langle f, h\rangle$, (linearity in second argument)
- $\langle h, h\rangle \geq 0$ and $\langle h, h\rangle = 0 \iff h = 0_{\mathbb{C}^n}$. (postive definiteness)

Remark 1.1. In physics it is typical to require linearity in the second argument. Hence we suggest sticking to this convention.

The likelihood is defined in terms of the inner product by

$$\mathcal{L}(\theta) \coloneqq e^{-\frac{1}{2}\left\langle h(\cdot;\theta) - d(\cdot), h(\cdot;\theta) - d(\cdot)\right\rangle_{\mathrm{GW}}}, \tag{2}$$

where $d \in \mathbb{C}^n$ is the detector strain, and h is a waveform model parameterized by θ (e.g, IMRPhenomD). Using the linearity and conjugate symmetry of the inner product, we may show that

$$\begin{aligned} -\ln \mathcal{L}(\theta) &= \frac{1}{2}\big\langle h(\theta) - d, h(\theta) - d\big\rangle_{\mathrm{GW}} \\ &= \frac{1}{2}\overbrace{\langle h(\theta), h(\theta)\rangle_{\mathrm{GW}}}^{(a)} - \mathfrak{Re}\overbrace{\langle h(\theta), d\rangle_{\mathrm{GW}}}^{(b)} + \frac{1}{2}\overbrace{\langle d, d\rangle_{\mathrm{GW}}}^{(c)}. \end{aligned}$$

Hence, for every $\mathcal{L}(\theta)$ evaluation, we must evaluate (a) and (b), which require $\mathcal{O}(n)$ evaluations of the waveform. Note that (c) can be precomputed once and stored. Relative binning aims to reduce the number of evaluations to $\mathcal{O}(m)$, where $m \ll n$, providing significant computational advantage.

1.2 Derivatives and Fisher matrix

We may also investigate whether whether heterodyning produces accurate gradient and Fisher matrix evaluations. The expressions for this are given by

$$\begin{aligned} -\ln\mathcal{L}_{,i}(\theta) &= \mathfrak{Re}\langle h_{,i}(\cdot,\theta), h(\cdot,\theta)\rangle - \mathfrak{Re}\langle h_{,i}(\cdot,\theta), d(\cdot)\rangle && \text{(heterodyne)} \\ &= \mathfrak{Re}\langle h_{,i}(\cdot,\theta), h(\cdot,\theta) - d(\cdot)\rangle, && \text{(standard)} \end{aligned}$$

and

$$\begin{aligned} -\ln\mathcal{L}_{,ij}(\theta) &= \mathfrak{Re}\langle h_{,ij}(\cdot,\theta), h(\cdot,\theta) - d(\cdot)\rangle + \mathfrak{Re}\langle h_{,i}(\cdot,\theta), h_{,j}(\cdot,\theta)\rangle \\ &\approx \mathfrak{Re}\langle h_{,i}(\cdot,\theta), h_{,j}(\cdot,\theta)\rangle. \end{aligned}$$

1.3 Heterodyning strategy

Heterodyning is a technique in signal processing which combines two high frequency signals to yield another signal with lower frequency. If this can be successfully accomplished, then we can use a coarser grid in evaluating the inner product with minimal error accumulation. Suppose that θ_0 is a fiducial set of parameters, and that we are interested in a signal with support $[f_{\text{min}}, f_{\text{max}}]$. Then

$$\begin{aligned} r(f;\theta) &\coloneqq \frac{h(f;\theta)}{h(f;\theta_0)} \\ &= \frac{A}{A_0} e^{i\overbrace{\left(\psi(f;\theta) - \psi(f;\theta_0)\right)}^{:=\Psi(f;\theta)}} \end{aligned}$$

is successfully heterodyned if $\Psi(f;\theta)$ is a "slowly varying" function with respect to f for "sufficiently many" $\theta \in \Theta$. $\Psi(f;\theta)$ is referred to as the *differential phase*.

If this condition holds, then one may efficiently compress the real and imaginary parts of r with a *linear spline*.

Definition 1. Let $r : [f_{\text{min}}, f_{\text{max}}] \to \mathbb{C}$, and consider the cover $[f_{\text{min}}, f_{\text{max}}] = \cup_{i=1}^{m-1}[f_-^{(i)}, f_+^{(i)}) \cup [f_-^{(m)}, f_+^{(m)}]$, where $f_+^{(i)} = f_-^{(i+1)}$ for every $i \in \{1,\ldots,m-1\}$, with $f_-^{(1)} = f_{\text{min}}$ and $f_+^{(m)} = f_{\text{max}}$. Then the *linear spline approximation* of r is given by

$$r(f;\theta) \approx r_0^{(b)}(\theta) + r_1^{(b)}(\theta)(f - f_-^{(b)}), \quad f \in \text{bin } b \tag{3}$$

where

$$r_0^{(b)}(\theta) \coloneqq r(f_-^{(b)};\theta), \qquad r_1^{(b)}(\theta) \coloneqq \frac{r(f_+^{(b)};\theta) - r(f_-^{(b)};\theta)}{f_+^{(b)} - f_-^{(b)}},$$

Proposition 1.2. *Suppose $r(f;\theta)$ is successfully heterodyned with fiducial parameters θ_0. Then*

$$\begin{aligned} \langle h(\cdot,\theta), d(\cdot)\rangle &\approx \sum_b \left[A_0^{(b)} r_0^{(b)}(\theta)^* + A_1^{(b)} r_1^{(b)}(\theta)^*\right], \\ \langle h(\cdot,\theta), h(\cdot,\theta)\rangle &\approx \sum_b \Big[B_0^{(b)} |r_0^{(b)}(\theta)|^2 + 2B_1^{(b)} \mathfrak{Re}\big(r_0^{(b)}(\theta)^* r_1^{(b)}(\theta)\big)\Big], \end{aligned}$$

where

$$A_0^{(b)} \coloneqq 4 \sum_{i: f_i \in \textit{bin } b} \frac{d(f_i) h(f_i;\theta_0)^*}{S(f_i)} \Delta f, \qquad A_1^{(b)} = 4 \sum_{i: f_i \in \textit{bin } b} \frac{d(f_i) h(f_i;\theta_0)^* (f_i - f_-^{(b)})}{S(f_i)} \Delta f,$$

$$B_0^{(b)} \coloneqq 4 \sum_{i: f_i \in \textit{bin } b} \frac{|h(f_i;\theta_0)|^2}{S(f_i)} \Delta f, \qquad B_1^{(b)} \coloneqq 4 \sum_{i: f_i \in \textit{bin } b} \frac{|h(f_i;\theta_0)|^2 (f_i - f_-^{(b)})}{S(f_i)} \Delta f.$$

Proof. (i) We have

$$\langle h(\cdot,\theta), d(\cdot)\rangle \coloneqq 4 \sum_{i=1}^{n} \frac{h(f_i;\theta)^* d(f_i)}{S(f_i)} \Delta f$$

$$
\begin{aligned}
&= 4\sum_b \sum_{i:f_i\in \text{bin } b} \frac{d(f_i)h(f_i;\theta_0)^* r(f_i;\theta)^*}{S(f_i)}\Delta f \\
&\approx 4\sum_b \sum_{i:f_i\in \text{bin } b} \frac{d(f_i)h(f_i;\theta_0)^*}{S(f_i)}\big[r_0^{(b)}(\theta)^* + r_1^{(b)}(\theta)^*(f_i - f_-^{(b)})\big]\Delta f \\
&= \sum_b \big[A_0^{(b)} r_0^{(b)}(\theta)^* + A_1^{(b)} r_1^{(b)}(\theta)^*\big].
\end{aligned}
$$

(ii) On the other hand, we have

$$
\begin{aligned}
\langle h(\cdot,\theta), h(\cdot,\theta)\rangle &= 4\sum_{i=1}^n \frac{|h(f_i;\theta)|^2}{S(f_i)}\Delta f \\
&= 4\sum_{i=1}^n \frac{|h(f_i;\theta_0)r(f_i;\theta)|^2}{S(f_i)}\Delta f \\
&\approx 4\sum_b \sum_{i:f_i\in \text{bin } b} \frac{|h(f_i;\theta_0)|^2\big|r_0^{(b)}(\theta) + r_1^{(b)}(\theta)(f_i - f_-^{(b)})\big|^2}{S(f_i)}\Delta f \\
&= 4\sum_b \sum_{i:f_i\in \text{bin } b} \frac{|h(f_i;\theta_0)|^2\Big(|r_0^{(b)}(\theta)|^2 + |r_1^{(b)}(\theta)(f - f_-^{(b)})|^2 + 2(f - f_-^{(b)})\mathfrak{Re}\big(r_0^{(b)}(\theta)^* r_1^{(b)}(\theta)\big)\Big)}{S(f_i)}\Delta f \\
&\approx \sum_b \Big[B_0^{(b)}|r_0^{(b)}(\theta)|^2 + 2B_1^{(b)}\mathfrak{Re}\big(r_0^{(b)}(\theta)^* r_1^{(b)}(\theta)\big)\Big],
\end{aligned}
$$

where in the second to last line we've used the fact that for $z_1, z_2 \in \mathbb{C}$, $|z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + 2\mathfrak{Re}(z_1^* z_2)$, and in the last line we've ignored terms of $\mathcal{O}(|f - f_-^{(b)}|^2)$. □

Proposition 1.3. *The terms we need to calculate for the derivative and Fisher may be approximated as follows:*

$$
\begin{aligned}
\langle h_{,j}(\theta), d\rangle &\approx \sum_b \big[A_0^{(b)} r_{0,j}^{(b)}(\theta)^* + A_1^{(b)} r_{1,j}^{(b)}(\theta)^*\big], \\
\langle h_{,j}(\theta), h(\theta)\rangle &\approx \sum_b \Big\{B_0^{(b)} r_{0,j}^{(b)}(\theta)^* r_0^{(b)}(\theta) + B_1^{(b)}\big[r_{0,j}^{(b)}(\theta)^* r_1^{(b)}(\theta) + r_{1,j}^{(b)}(\theta)^* r_0^{(b)}(\theta)\big]\Big\}, \\
\langle h_{,j}(\theta), h_{,k}(\theta)\rangle &\approx \sum_b \Big\{B_0^{(b)} r_{0,j}^{(b)}(\theta)^* r_{0,k}^{(b)}(\theta) + B_1^{(b)}\big[r_{0,j}^{(b)}(\theta)^* r_{1,k}^{(b)}(\theta) + r_{1,j}^{(b)}(\theta)^* r_{0,k}^{(b)}(\theta)\big]\Big\}.
\end{aligned}
$$

Proof. □

2 Bin selection algorithm

Definition 2. Let $h(f;\theta)$ be a complex valued signal. Then the *post-Newtonian ansatz* asserts that $\psi(f) \coloneqq \arg(h(f))$ takes the form

$$
\psi(f;\theta) \coloneqq \sum_i \alpha_i(\theta) f^{\gamma_i}, \tag{4}
$$

where $\gamma \coloneqq (-5/3, -2/3, 1, 5/3, 7/3)$.

We now have that

$$
\begin{aligned}
\Psi(f;\theta) &\coloneqq \psi(f;\theta) - \psi(f;\theta_0) \\
&= \sum_i \big(\alpha_i(\theta) - \alpha_i(\theta_0)\big) f^{\gamma_i} \\
&= \sum_i \delta\alpha_i(\theta) f^{\gamma_i}
\end{aligned}
$$

The heterodyne approximation is valid if Ψ varies slowly within a bin, which translates to a constraint on the $\delta\alpha_i$ terms. This motivates the following definition:

Proposition 2.1. *Suppose that the parameters θ are restricted to*

$$\Theta \coloneqq \left\{\theta \in \chi : \forall i \; |\delta\alpha_i(\theta)| \le 2\pi\chi f_{*,i}^{-\gamma_i}\right\}, \tag{5}$$

where

$$f_{*,i} \coloneqq \begin{cases} f_{max} & \text{if } i: \; \gamma_i > 0 \\ f_{min} & \text{else}. \end{cases} \tag{6}$$

Then the differential phase change over the interval $[f_-, f_+]$ obeys the following bound:

$$|\Psi(f_+;\theta) - \Psi(f_-;\theta)| \le 2\pi\chi \sum_i \left| \left(\frac{f_+}{f_{*,i}}\right)^{\gamma_i} - \left(\frac{f_-}{f_{*,i}}\right)^{\gamma_i} \right|. \tag{7}$$

Proof.

$$\begin{aligned} |\Psi(f_+;\theta) - \Psi(f_-;\theta)| &= \left| \sum_i \delta\alpha_i(\theta) \left[f_+^{\gamma_i} - f_-^{\gamma_i} \right] \right| \\ &\le \sum_i |\delta\alpha_i(\theta)| |f_+^{\gamma_i} - f_-^{\gamma_i}| && \text{(triangle inequality)} \\ &\le 2\pi\chi \sum_i \left| \left(\frac{f_+}{f_{*,i}}\right)^{\gamma_i} - \left(\frac{f_-}{f_{*,i}}\right)^{\gamma_i} \right|. \end{aligned}$$

□