SGNL/GstLAL Mass Model and Its Refinement for IR1

Victoria Niu, Chad Hanna, Jolien Creighton

09/02/2026

In Heather's generic description, the mass model is considered a "**source-dependent ranking statistic for CBC candidates**", which incorporates the astrophysical source population as a prior. She identifies several benefits of incorporating such a source-dependent ranking statistic into the likelihood ratio used by SGNL/GstLAL:

1. Increase BNS VT.
2. Enable the pipeline to perform source-specific searches using a single filtering product.
3. Direct mapping between likelihood-ratio $\mathcal{L}$ and FAR.

1 Heather's Mass-Model Formalism

In her dissertation [1], the likelihood ratio for the SGNL/GstLAL ranking statistic is derived as:

$$\mathcal{L}\left(\{D_{H_{H1}}, D_{H_{L1}}, \ldots\}, \{H1, L1, \ldots\}, \rho_{H1}, \chi^2_{H1}, \rho_{L1}, \chi^2_{L1}, \ldots, \vec{\theta}\right) \tag{1}$$

$$= \mathcal{L}(\ldots|\vec{\theta})\,\mathcal{L}(\vec{\theta}) \tag{2}$$

$$= \underbrace{\frac{P\left(\{D_{H_{H1}}, D_{H_{L1}}, \ldots\}, \{H1, L1, \ldots\}, \rho_{H1}, \chi^2_{H1}, \rho_{L1}, \chi^2_{L1}, \ldots \,\middle|\, \vec{\theta}, H_s\right)}{P\left(\{D_{H_{H1}}, D_{H_{L1}}, \ldots\}, \{H1, L1, \ldots\}, \rho_{H1}, \chi^2_{H1}, \rho_{L1}, \chi^2_{L1}, \ldots \,\middle|\, \vec{\theta}, H_n\right)}}_{\text{format of Kipp, Chad, \& Jacob's paper [3]}}\,\mathcal{L}(\vec{\theta}), \tag{3}$$

where $\{D_{H_{H1}}, D_{H_{L1}}, \ldots\}$ is the set of horizon distances for all instruments in the network at the time of the observed event, $\{H1, L1, \ldots\}$ is the set of instruments that observed the event, ρ_{IFO} and χ^2_{IFO} are template matched-filter SNRs and auto-correlations.

H_s represents the "signal hypothesis", and H_n represents the "noise hypothesis". $\vec{\theta}$ are the intrinsic parameters of the template that recovered the event. For $\mathcal{L}(\vec{\theta})$, Heather writes it as:

$$L(\vec{\theta}) = \frac{P(\vec{\theta}|H_s)}{P(\vec{\theta}|H_n)}. \tag{4}$$

The denominator, $P(\vec{\theta}|H_n)$, is the probability of recovering a template waveform with parameters $\vec{\theta}$ given that there is *no* signal in the data [1]. According to Heather, this probability is estimated as the total rate for a given $\vec{\theta}$, divided by the sum of the total rates over all possible values of $\vec{\theta}$ in the search. Confirm with Leo: In STRIKE/GstLAL, I believe $P(\vec{\theta}|H_n)$ is incorporated into the $P(t_{\mathrm{ref}}, \vec{\theta}|H_n)$, which quantifies the probability of an event with the template θ being caused by the random noise fluctuation at $t = t_{\mathrm{ref}}$ [2]. The occurrence of noise event is modeled with a Poisson process that is uniform across a template group $\{\vec{\theta}\}$.

The nominator, $P(\vec{\theta}|H_s)$, represents the probability that a template waveform with parameters $\vec{\theta}$ is recovered as the best-match template waveform, given that there *is* a GW signal in the data [1]. Without a mass-model, $P(\vec{\theta}|H_s)$ is effectively assumed to be constant, which does not accurately represent the astrophysical distribution of compact binaries in the Universe. **This term corresponds directly to the population-model term in Eq. (2) of Leo's likelihood-ratio formulation [2].**

1.1 Formulation of $P(\vec{\theta}|H_s)$

Heather's study focuses on deriving the weighting factor of $P(\vec{\theta}|H_s, \rho)$, which requires marginalizing over SNR ρ to obtain $P(\vec{\theta}|H_s)$:

$$P(\vec{\theta}|H_s) = \int_{\rho=4}^{\infty} P(\vec{\theta}|H_s, \rho) P(\rho|H_s) d\rho,$$

where she assumes an SNR distribution corresponding to sources uniformly distributed in volume, such that $P(\rho \mid H_s) = P(\rho) \propto \rho^{-4}$. The lower integration limit, $\rho = 4$, is set by the SNR threshold used in GW searches.

Meanwhile, for an actual detection, ρ is obtained directly from the matched-filter result. Thus,

$$P(\vec{\theta}|H_s) = P(\vec{\theta}|H_s, \rho = \rho_{\text{obs}}). \tag{5}$$

As Heather notes in her dissertation [1]: "Becuase the SNRs of event candidate are recorded by the pipeline, this allows us directly compute $P(\vec{\theta}|H_s, \rho)$ for each event candidate, which is more precise than computing the marginalized probability". Using Eq. 5 formulation, we briefly explain how the mass/population-model term is incorporated into the likelihood ratio within the software infrastructure.

Both STRIKE and GstLAL obtain $P(\vec{\theta}|H_s, \rho)$ from a specified mass/population-model file. For notes on loading this file in the ranking jobs, see App. B. The mass-model file, stored in HDF5 format, contains the coefficients needed to evaluate $P(\vec{\theta}|H_s, \rho)$ at the observed SNR, $\rho = \rho_{\text{obs}}$. Both codes therefore directly evaluate Eq. 5 using the observed ρ and the coefficients stored in the mass-model file (see gstlal-lr and STRIKE-lr). **Neither GstLAL nor STRIKE performs any marginalization over SNR.**

1.2 Formulation of $P(\vec{\theta}|H_s, \rho)$

After establishing its interface with the GstLAL/STRIKE infrastructure and Leo's likelihood-ratio notation, we revisit Heather's key quantity, $P(\vec{\theta}|H_s, \rho)$, which we refer to as the mass model throughout this writeup.

First, Heather defines the template/intrinsic coordinates that characterize the GW signal in her O1 study [1]:

$$\vec{\theta} = \{m_1, m_2, \chi_1, \chi_2\}$$

- Component mass m_i.
- Dimensionless component spin $\chi_i = \frac{\vec{S}_i \cdot \hat{L}}{m_i^2}$.
- Spin angular momentum of the i-th component $\vec{S}_i$
- Direction of the binary's orbital angular momentum $\hat{L}$

In her coordinate system, there are five degrees of freedom in total, including ρ. In **sgn-manifold**, as in O4, we place the templates in the coordinates of:

$$\vec{\theta} = \{m_1, m_2, \chi_{\rm eff}\}$$

where

$$\chi_{\rm eff} = \frac{m_1\chi_1 + m_2\chi_2}{m_1 + m_2}. \tag{6}$$

Our coordinate system has **four degrees of freedom**, including ρ.

Next, Heather makes two assumptions for modeling the signal [1]: 1) The template bank is densely populated, such that a noise-free signal can be represented as:

$$\vec{s} \equiv \vec{s}(\vec{\theta}_j, \rho) = \rho\, \vec{t}_j(\vec{\theta}),$$

where $\vec{\theta}_j$ denotes the intrinsic parameters of a specific normalized template $\vec{t}_j(\vec{\theta})$. 2) The noise is stationary and Gaussian. In addition, the measured data, $\vec{d} = \vec{s} + \vec{n}$, is assumed to be described by a waveform template $\vec{t}_k(\vec{\theta})$ with intrinsic parameters $\vec{\theta}_k$ multiplied by the observed SNR $\rho_{{\rm obs},k}$:

$$\vec{d} \approx \rho_{{\rm obs},k}\, \vec{t}_k(\vec{\theta}).$$

In the following, using Heather's notation, we omit $\vec{\theta}$ from the template variable notation. We can interpret it as the noise in the data shifting the signal $\vec{t}_j$ such that it is recovered by the template $\vec{t}_k$. Fig. 1 from Heather's dissertation illustrates the relationship between $\vec{t}_k$ and $\vec{t}_j$.

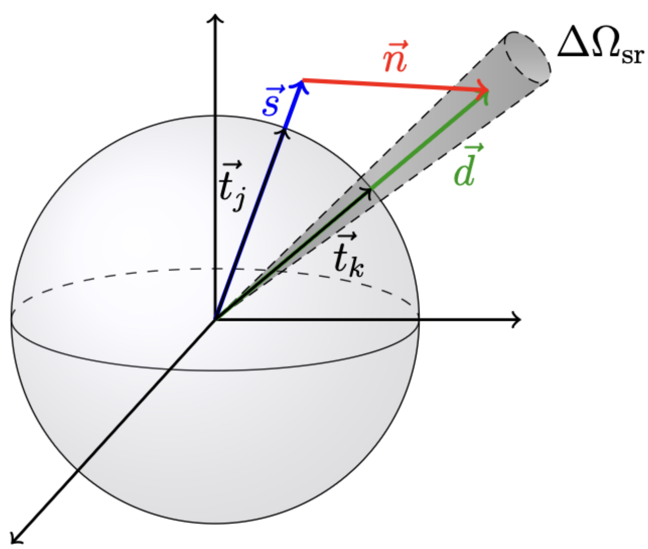

Figure 1: Figure 3.2 of [1]. Waveform space in which waveform templates lie on the unit sphere. $\vec{t}_j$ denotes the normalized template waveform that recovers the signal $\vec{s}$ in the absence of noise $\vec{n}$ (i.e. $\vec{s} = \rho\, \vec{t}_j$), while $\vec{t}_k$ denotes the template waveform that recovers the signal with the addition of noise (i.e. $\vec{d} = \vec{s} + \vec{n} = \rho_{{\rm obs},k}\, \vec{t}_k$). Signals found inside the solid angle $\Delta\Omega_{sr}$ are recovered by template $\vec{t}_k$, where $\Delta\Omega_{sr}$ is determined by the density of the template bank.

Given that a candidate known to look like $\vec{s} = \rho\,\vec{t_j}$, the probability that a signal from a specific source population is recovered by a template $\vec{t_k}$ is [1]:

$$P(\vec{t_k}|H_s,\rho) = \sum_j \left\{ \underbrace{P(\text{candidate in } \vec{t_k}|H_s = \rho\,\vec{t_j})}_{\text{template migration}} \times \underbrace{P(\vec{t_j}|H_s)}_{\text{source pop}} \right\}. \tag{7}$$

2 Source Population Term $P(\vec{t_j}|H_s)$

We refer $P(\vec{t_j}|H_s)$ as the source population term. Heather describes it as "The probability that a signal described by a template on the unit sphere, and depends on the astrophysical source model", which is [1]:

$$P(\vec{t_j}|H_s) = \int_{V_{\vec{t_j}}} p(\vec{\gamma})d\vec{\gamma} \tag{8}$$

(note: Heather writes the integrand as $P(\vec{\gamma})$, which we believe should instead be the PDF $p(\vec{\gamma})$.) Since the template bank is discretized, Heather further **discretizes the probability space using a probability mass function (PMF)** instead [1]:

$$P(\vec{t_j}|H_s) = \sum_{V_i \in \mathcal{V}} p_i(\vec{\gamma})V_i \tag{9}$$

where V_i **is the volume element of the i-th simplex of the Voronoi cell**. Heather turns the bank into a Voronoi diagram, with each template as the Voronoi point V_i enclosed in a Voronoi cell $\mathcal{V}$.

Section 3.4.1 of [1] explains how to compute the PMF of a template's Voronoi cell in stochastic placement. In sgn-manifold and O4, this computation is relatively straightforward because the templates are placed within a hyper-rectangle. V_i corresponds to the volume of the i-th hyper-rectangle:

$$V_i = \Delta m_{1,i} \times \Delta m_{2,i} \times \Delta\chi_i. \tag{10}$$

Victoria's notes: The $\Delta\Omega_{sr}$ shown in Fig. 1 does not, in general, correspond exactly to the volume V_i, since the Voronoi cell associated with a template may only partially overlap with the hypercube used to define the coordinate volume. In contrast, sgn-manifold, with its binary-tree approach, simplifies this geometry by partitioning the parameter space into non-overlapping hyper-rectangles, for which the volume V_i can be computed directly.

2.1 Heather's O1 Source Population Model

In O1, Heather considers two distinct models: an aligned-spin BNS population and an aligned-spin BBH population. She models the BNS population with a Gaussian distribution (informed by the Galactic BNS population), while modeling the BBH population as flat in $\log(m_1)$ and $\log(m_2)$. She assumes uniform distribution for all spins. Fig. 2 shows the fiducial population model adopted by Heather.

2.2 Anarya's O4 Mass Population Model

In O4, Anarya adopts a single fiducial mass-distribution model for all CBC sources. He models the **primary mass m_1 of CBC sources using the Salpeter initial mass function**:

$$p(m_1) \propto \ m_1^{-2.35} \ \text{ with } m_{\min} \leq m_1 \leq m_{\max} \tag{11}$$

	BNS	BBH
Mass limits	$1M_\odot \leq m_{1,2} \leq 3M_\odot$	$m_{1,2} \geq 5M_\odot$, $m_1 + m_2 \leq 100M_\odot$
Mass distribution	$\mathcal{N}(1.33M_\odot, 0.09)$	$\sim (1/m_1)(1/m_2)$
Spin limits	$-0.05 \leq \chi_{1,2} \leq 0.05$	$-0.99 \leq \chi_{1,2} \leq 0.99$
Spin distribution	$\mathcal{U}(-0.05, 0.05)$	$\mathcal{U}(-0.99, 0.99)$

Figure 2: Table 3.1 of [1]. A summary of the source distributions used by Heather. $\mathcal{N}(\mu, \sigma)$ denotes a normal distribution with mean μ and standard deviation σ. $\mathcal{U}(a, b)$ denotes a uniform distribution in the range $[a, b]$.

and the **secondary mass m_2 using a uniform distribution with $m_2 \leq m_1$ constraint**:

$$p(m_2|m_1) = \frac{1}{m_1 - m_{\text{min}}} \tag{12}$$

such that $\int_{m_{\text{min}}}^{m_1} p(m_2|m_1)dm_2 = 1$. m_{min} is a fiducial minimum mass introduced to normalize the Salpeter distribution. Currently, we set $m_{\text{min}} = 0.8M_\odot$ for the stellar-mass search.

Therefore, as implemented in O4 software (code), $p(\vec{\gamma})$ in O4 is:

$$p(\vec{\gamma}) = p(m_1)p(m_2|m_1)$$

No spin distribution is considered in the O4 model. However, for the coordinate $\vec{\gamma} = \{m_1, m_2, \chi_{\text{eff}}\}$ used in the SGNL/GstLAL O4 bank, the full fiducial model must also account for the spin distribution:

$$p(\vec{\gamma}) = p(m_1)p(m_2|m_1)p(\chi_{\text{eff}}|m_1, m_2),$$

for which we address in the Sec. 2.3.

2.3 Victoria and Jolien's Spin Population Model

Overall, following Heather, we assume a uniform distribution for the dimensionless component spins χ_i. This requires marginalizing over χ_i to transform the distribution into the χ_{eff} bank coordinate. We refer to this procedure as χ_{eff} marginalization in our communications.

First, we revisit the constraints used in our stellar-mass template placement (Table 1). Since we place templates in aligned or anti-aligned directions with the orbital angular momentum, we only consider the z-direction component $\chi_i = \chi_{iz}$. **These constraints introduce a spin cut at the mass boundary between NSs and BHs**.

Class	Primary m_1	Secondary m_2	Primary spin χ_1	Secondary spin χ_2
BNS	$< 3.0\,M_\odot$ (NS)	$< 3.0\,M_\odot$ (NS)	$[-0.05,\ 0.05]$	$[-0.05,\ 0.05]$
NSBH	$\geq 3.0\,M_\odot$ (BH)	$< 3.0\,M_\odot$ (NS)	$[-0.99,\ 0.99]$	$[-0.05,\ 0.05]$
BBH	$\geq 3.0\,M_\odot$ (BH)	$\geq 3.0\,M_\odot$ (BH)	$[-0.99,\ 0.99]$	$[-0.99,\ 0.99]$

Table 1: Mass and Spin Constrain of O4 Bank, with the NS/BH boundary set at $3.0\,M_\odot$.

Next, to get $p(\chi_{\rm eff}|m_1, m_2)$, we can marginalize $p(\chi_{\rm eff}, \chi_2|m_1, m_2)$ over χ_2:

$$p(\chi_{\rm eff}|m_1, m_2) = \int p(\chi_{\rm eff}, \chi_2|m_1, m_2) d\chi_2 \tag{13}$$

Since χ_1 and χ_2 each separately depend on m_1 and m_2 (i.e. the cutoff is applied to the component spin of each mass in the binary), we have $p(\chi_1|m_1)$ and $p(\chi_2|m_2)$. The Jacobian then gives:

$$p(\chi_{\rm eff}, \chi_2|m_1, m_2) d\chi_{\rm eff} d\chi_2 = p(\chi_1|m_1) p(\chi_2|m_2) d\chi_1 d\chi_2 \tag{14}$$

From Eq. 6, we can get $d\chi_1 = \frac{m_1+m_2}{m_1} d\chi_{\rm eff}$ and $\chi_1 = \frac{(m_1+m_2)\chi_{\rm eff} - m_2\chi_2}{m_1}$. Plugging them into the Eq. 14 and cancel the $d\chi_2$ on both sides, we have

$$p(\chi_{\rm eff}, \chi_2|m_1, m_2) d\chi_{\rm eff} = p(\frac{(m_1+m_2)\chi_{\rm eff} - m_2\chi_2}{m_1}|m_1) p(\chi_2|m_2) \frac{m_1+m_2}{m_1} d\chi_{\rm eff} \tag{15}$$

Cancel $d\chi_{\rm eff}$ on both side, we have:

$$p(\chi_{\rm eff}, \chi_2|m_1, m_2) = p(\frac{(m_1+m_2)\chi_{\rm eff} - m_2\chi_2}{m_1}|m_1) p(\chi_2|m_2) \frac{m_1+m_2}{m_1}$$

Eq. 13 then becomes:

$$p(\chi_{\rm eff}|m_1, m_2) = \frac{m_1+m_2}{m_1} \int p(\frac{(m_1+m_2)\chi_{\rm eff} - m_2\chi_2}{m_1}|m_1) p(\chi_2|m_2) d\chi_2 \tag{16}$$

We now need to determine the $p(\chi_2 \mid m_2)$ distribution, including the mass-dependent spin cut, in order to evaluate this integral. For independent uniform priors without cuts, we have

$$p(\chi_1) = \frac{1}{\chi_{1,\rm max} - \chi_{1,\rm min}}, \qquad p(\chi_2) = \frac{1}{\chi_{2,\rm max} - \chi_{2,\rm min}},$$

Since $|\chi_{i,\rm max}| = |\chi_{i,\rm min}|$ in our constraint,

$$p(\chi_i) = \frac{1}{2\ \chi_{i,\rm max}},$$

To apply a spin cut at the NS/BH boundary, we consider a two-bin model in which the mass axis is divided at a single break-point of $3M_\odot$:

$$m_{\rm min} \le m_i < 3M_\odot \quad \text{(bin 1)}, \qquad 3M_\odot \le m_i \le m_{\rm max} \quad \text{(bin 2)}, \tag{17}$$

with spin ranges $\left[\chi_{i,\rm min}^{(1)}, \chi_{i,\rm max}^{(1)}\right]$ in bin 1 and $\left[\chi_{i,\rm min}^{(2)}, \chi_{i,\rm max}^{(2)}\right]$ in bin 2. The distribution then becomes:

$$p(\chi_i \mid m_i) = \begin{cases} \dfrac{1}{2\chi_{i,\rm max}^{(1)}} = \frac{1}{2\times 0.05} & m_{\rm min} \le m_i < 3M_\odot \\ \dfrac{1}{2\chi_{iz,\rm max}^{(2)}} = \frac{1}{2\times 0.99}, & 3M_\odot \le m_i \le m_{\rm max} \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

With Eq. 18, the solution of Eq. 16 now can be solved as a **trapezoid integration**:

$$\begin{aligned} p(\chi_{\rm eff} \mid m_1, m_2) = & \underbrace{\frac{m_1+m_2}{m_1}}_{\text{Jacobian}} \underbrace{\frac{1}{2\chi_{1,\rm max}} \frac{1}{2\chi_{2,\rm max}}}_{\text{probability density}} \int_{-\chi_{2,\rm max}}^{\chi_{2,\rm max}} {\rm d}\chi_2 \\ & \underbrace{\Theta\left(m_1\chi_{1,\rm max} - (m_1+m_2)\chi_{\rm eff} + m_2\chi_2\right) \Theta\left(m_1\chi_{1,\rm max} + (m_1+m_2)\chi_{\rm eff} - m_2\chi_2\right)}_{\text{trapezoid width}}. \end{aligned} \tag{19}$$

2.4 Code Integration

We integrate the analytical solution Eq. 19 of the spin distribution into sgn-manifold code, along with the mass distribution from Eq. 11 and Eq. 12, which together serves as the fiducial population of $p(\vec{\gamma})$ for $P(\vec{t_j}|H_s)$ in IR1.

The relevant MRs to sgn-manifold:

- Victoria's initial MR: chi-marginalization
- Chad's later MR: hardcode fix

2.5 Validation: Prior Volume Ratios

One reason we identified a problem with the O4 mass model is that the prior volume assigned to each category does not properly reflect the observed rates of those populations. For GstLAL O4 Mass Model, we get:

$$\sum_{V_i \in \mathcal{V}_{\mathrm{BNS}}} p_i(\vec{\gamma}) V_i : \sum_{V_i \in \mathcal{V}_{\mathrm{NSBH}}} p_i(\vec{\gamma}) V_i : \sum_{V_i \in \mathcal{V}_{\mathrm{BBH}}} p_i(\vec{\gamma}) V_i = 1 : 2.6 : 3.7$$

with NS mass cut at $3M_{\odot}$. For the source population, using a rough estimate based on the joint population model of GWTC-5 [4], the ratios might be

$$\mathrm{BNS} : \mathrm{NSBH} : \mathrm{BBH} \approx 4.0 : 1.0 : 1.9.$$

To account for the "observable" volume set by the detector's horizon distance H_D, the observable ratios becomes:

$$\sum_{V_i \in \mathcal{V}_{\mathrm{BNS}}} p_i(\vec{\gamma}) V_i H_{D_i}^3(\vec{\gamma}) : \sum_{V_i \in \mathcal{V}_{\mathrm{NSBH}}} p_i(\vec{\gamma}) V_i H_{D_i}^3(\vec{\gamma}) : \sum_{V_i \in \mathcal{V}_{\mathrm{BBH}}} p_i(\vec{\gamma}) V_i H_{D_i}^3(\vec{\gamma}) = 1 : 10 : 1212,$$

where the actual number of events observed by GWTC-5 [4] is

$$\mathrm{BNS} : \mathrm{NSBH} : \mathrm{BBH} = 2 : 5 : 260.$$

The O4 Mass Model differs from the observed rates by **a factor of 10**. After the χ_{eff} marginalization fixed for the spin distribution, we see these ratios agree with the expected value:

$$\sum_{V_i \in \mathcal{V}_{\mathrm{BNS}}} p_i(\vec{\gamma}) V_i : \sum_{V_i \in \mathcal{V}_{\mathrm{NSBH}}} p_i(\vec{\gamma}) V_i : \sum_{V_i \in \mathcal{V}_{\mathrm{BBH}}} p_i(\vec{\gamma}) V_i = 4.0 : 1.5 : 1.0$$

and observable rates:

$$\sum_{V_i \in \mathcal{V}_{\mathrm{BNS}}} p_i(\vec{\gamma}) V_i H_{D_i}^3(\vec{\gamma}) : \sum_{V_i \in \mathcal{V}_{\mathrm{NSBH}}} p_i(\vec{\gamma}) V_i H_{D_i}^3(\vec{\gamma}) : \sum_{V_i \in \mathcal{V}_{\mathrm{BBH}}} p_i(\vec{\gamma}) V_i H_{D_i}^3(\vec{\gamma}) = 1 : 1 : 78$$

3 Template Migration Term $P(\text{candidate in } \vec{t_k} | H_s = \rho\, \vec{t_j})$

We refer to $P(\text{candidate in } \vec{t_k} \mid H_s = \rho, \vec{t_j})$ as the **template migration** term. Several models for this term have been proposed, which we discuss separately below. For IR1, we plan to use Chad's latest model, which yields prior-volume ratios that are consistent with the observed rates.

3.1 Heather's Model [1]

Using the formulation of Sec. 1.2,

$$\vec{d} = \vec{s} + \vec{n}$$
$$\rho_{\text{obs},k}\, \vec{t_k} = \rho\, \vec{t_j} + \vec{n},$$

Heather considers a Gaussian distribution for $\vec{n}$ with a constant variance of 1. The Gaussian PDF for N-dimensions is

$$f(\vec{n}) = \frac{1}{(2\pi)^{N/2} e^{-\frac{1}{2}|\vec{n}|^2}}$$

She solves for $|\vec{n}|^2$ as:

$$|\vec{n}|^2 = (\rho_{\text{obs, k}}\, \vec{t_k} - \rho\, \vec{t_j}) \cdot (\rho_{\text{obs, k}}\, \vec{t_k} - \rho\, \vec{t_j}) = \rho^2 + \rho_{\text{obs, k}}^2 - 2\rho_{\text{obs, k}}\, \rho(\vec{t_j} \cdot \vec{t_k}) \tag{20}$$

where $\vec{t_j} \cdot \vec{t_k}$ is the overlap between templates $\vec{t_j}$ and $\vec{t_k}$.

Heather then attempts to find the probability that $\vec{d} = \rho\, \vec{t_j} + \vec{n}$ lies inside the conic volume of some template $\vec{t_k}$ with solid angle $\Delta\Omega_{sr}$. If $\vec{d}$ is inside the conic volume, it means that it is recovered by the template $\vec{t_k}$ because $\vec{t_k}$ is the point of closest approach on the template bank. The probability that any signal is recovered by $\vec{t_k}$ can be expressed as a ratio of the probability density of one template over the sum of probability densities of all the templates in the bank:

$$P(\text{candidate in } \vec{t_k} | H_s = \rho\, \vec{t_j}) = \frac{\sum \int_0^\infty \frac{1}{(2\pi)^{N/2}} e^{-\frac{1}{2}|\vec{n}|^2} \rho_{\text{obs},k}^{N-1}\, \mathrm{d}\rho_{\text{obs},k}\, \Delta\Omega_{\text{sr}}}{\sum_{\vec{t_{k'}}} \int_0^\infty \frac{1}{(2\pi)^{N/2}} e^{-\frac{1}{2}|\vec{n}|^2} \rho_{\text{obs},k}^{N-1}\, \mathrm{d}\rho_{\text{obs},k}\, \Delta\Omega_{\text{sr}}} \tag{21}$$

$$= \frac{P(\vec{t_k})}{\sum_{\vec{t_{k'}}} P(\vec{t_{k'}})} \tag{22}$$

According to Heather, considering only the numerator in Eq. 21 and substituting $|\vec{n}|^2$ with Eq. 20, this becomes

$$P(\vec{t_k}) = \frac{\Delta\Omega_{\text{sr}}}{(2\pi)^{N/2}} \int_0^\infty e^{-\frac{1}{2}\left(\rho^2 + \rho_{\text{obs},k}^2 - 2\rho_{\text{obs},k}\rho(\vec{t_j}\cdot\vec{t_k})\right)} \rho_{\text{obs},k}^{N-1}\, \mathrm{d}\rho_{\text{obs},k} \tag{23}$$

$$= \frac{\Delta\Omega_{\text{sr}}}{(2\pi)^{N/2}} e^{-\frac{1}{2}\rho^2\left(1-(\vec{t_j}\cdot\vec{t_k})^2\right)} \int_0^\infty e^{-\frac{1}{2}\left(\rho_{\text{obs},k} - \rho(\vec{t_j}\cdot\vec{t_k})\right)^2} \rho_{\text{obs},k}^{N-1}\, \mathrm{d}\rho_{\text{obs},k}, \tag{24}$$

where in the final equality, the exponential term is rewritten by completing the square and the constant terms are moved outside the integral. Letting the number of dimensions $N = 5$ (it is the degree of freedom in Heather's coordinate), Eq. 24 is integrated and becomes

$$P(\vec{t_k}) = \frac{\Delta\Omega_{\text{sr}}}{(2\pi)^2} e^{-\frac{1}{2}\rho^2\left(1-(\vec{t_j}\cdot\vec{t_k})^2\right)} \Bigg\{ \sqrt{\frac{\pi}{2}}\, \text{erf}\left(\frac{\rho(\vec{t_j}\cdot\vec{t_k})}{\sqrt{2}}\right) \left[(\rho(\vec{t_j}\cdot\vec{t_k}))^4 + 6(\rho(\vec{t_j}\cdot\vec{t_k}))^2 + 3\right] + e^{-\frac{1}{2}(\rho(\vec{t_j}\cdot\vec{t_k}))^2} \left[(\rho(\vec{t_j}\cdot\vec{t_k}))^3 + 5\rho(\vec{t_j}\cdot\vec{t_k})\right] \Bigg\}, \tag{25}$$

which therefore gives

$$P(\text{candidate in } \vec{t_k} | H_s = \rho\, \vec{t_j}) \sim e^{\frac{1}{2}\rho^2(\vec{t_j}\cdot\vec{t_k})^2} \sqrt{\frac{\pi}{2}}\, \text{erf}\left(\frac{\rho(\vec{t_j}\cdot\vec{t_k})}{\sqrt{2}}\right) \left[(\rho(\vec{t_j}\cdot\vec{t_k}))^4 + 6(\rho(\vec{t_j}\cdot\vec{t_k}))^2 + 3\right] + (\rho(\vec{t_j}\cdot\vec{t_k}))^3 + 5\rho(\vec{t_j}\cdot\vec{t_k}). \tag{26}$$

Eq. 26 is normalized by by the numerator summed over all templates, which is the denominator of Eq. 22. In her dissertation, Heather then discusses how to compute the overlap $\vec{t_k} \cdot \vec{t_j}$, which should be calculated over the entire bank, in a computationally feasible way [1].

3.2 Chad's O4 Model

In O4, Anarya and Chad consider a formulation similar to Heather's:

$$\vec{d} = \rho\,\vec{t_j} + \vec{n}$$

$$\rho_{\mathrm{obs},k} = \rho(\vec{t_j} \cdot \vec{t_k}) + \underbrace{\vec{n} \cdot \vec{t_k}}_{n_k}$$

where $n_k \sim \mathcal{N}(0,1)$. Meanwhile, **Chad's model assumes that the observed SNR is equal to the true SNR of the signal**:

$$\max_{k'} \left\{\rho_{obs,k'}\right\} \approx \rho.$$

Here, $\max_{k'}\{\rho_{\mathrm{obs},k'}\}$ indicates that $\vec{t_k}$ is the recovered template with the highest SNR, implying that n_k must satisfy

$$n_k = \rho(1 - \vec{t_j} \cdot \vec{t_k}),$$

up to a window accounting for uncertainty and SNR loss. Therefore, the template migration term in O4 can be written as:

$$P(\text{candidate in } \vec{t_k} | H_s = \rho\,\vec{t_j}) \propto e^{-\frac{1}{2}\rho^2(1-(\vec{t_j}\cdot\vec{t_k}))^2}\Delta n \tag{27}$$

and marginalized over all templates:

$$P(\text{candidate in } \vec{t_k} | H_s = \rho\,\vec{t_j}) = \frac{e^{-\frac{1}{2}\rho^2(1-(\vec{t_j}\cdot\vec{t_k}))^2}\Delta n}{\sum_k e^{-\frac{1}{2}\rho^2(1-(\vec{t_j}\cdot\vec{t_k}))^2}\Delta n} \tag{28}$$

Anarya's notes: $n_k \sim \mathcal{N}(0,1)$ is probably not true due to correlations between k and k'.

3.3 Anarya's Two-Phase Model [5]

The following is adapted from Anarya's notes on the mass model, with minor edits for notational consistency.

In Anarya's model, the shift in the recovered template due to noise can be interpreted as a phase shift on the unit sphere of Figure 1. Therefore, modeling $\vec{s}$ and $\vec{d}$ in the two-phase space $(\sin\phi, \cos\phi)$ gives:

$$\vec{s} \;=\; (\rho\cos(\phi_0)\,\vec{t_j}, \rho\sin(\phi_0)\,\vec{t_j}) \tag{29}$$

$$\vec{d} \;=\; \vec{s} + \vec{n} = (\rho_{\mathrm{obs},k}\cos(\phi)\,\vec{t_k}, \rho_{\mathrm{obs},k}\sin(\phi)\,\vec{t_k}) \tag{30}$$

where noise $\vec{n} = (\vec{n_1}, \vec{n_2})$ and each $\vec{n_i}$ are **large**-dimensional independent Gaussian function (The bank is whitened). Let θ_k be the intrinsic parameters corresponding to template t_k.

$$p(\rho_{\mathrm{obs},k}, \vec{t_k}, \phi | \rho, \vec{t_j}, \phi_0) \;\propto\; e^{-\frac{1}{2}\{(\rho_{\mathrm{obs},k}\cos(\phi)\vec{t_k} - \rho\cos(\phi_0)\vec{t_j})^2 + (\rho_{\mathrm{obs},k}\sin(\phi)\vec{t_k} - \rho\sin(\phi_0)\vec{t_j})^2\}} \tag{31}$$

$$\propto\; e^{-\frac{1}{2}\{\rho^2_{\mathrm{obs},k} + \rho^2 - 2\rho\rho_{\mathrm{obs},k}\vec{t_j}\cdot\vec{t_k}\cos(\phi-\phi_0)\}} \tag{32}$$

He then marginalizes it over unknown phase:

$$p(\rho_{\text{obs},k}, \vec{t_k}|\rho, \vec{t_j}, \phi_0) = \int_0^{2\pi} d\phi\, e^{-\frac{1}{2}\{\rho^2_{\text{obs},k}+\rho^2-2\rho\rho_{\text{obs},k}\, \vec{t_j}\cdot\vec{t_k}\, \cos(\phi-\phi_0)\}} \quad (33)$$

$$= e^{-\frac{1}{2}(\rho^2+\rho^2_{\text{obs},k})}\mathrm{I}_0(\rho\rho_{\text{obs,k}}\, \tilde{\mathrm{t_j}}\cdot\tilde{\mathrm{t_k}}), \quad (34)$$

where I_0 is the modified Bessel function of the first kind of order 0. Following Heather's method, he computes $P(\text{candidate in } \vec{t_k}|H_s = \rho\,\vec{t_j})$ as the integral of $p(\rho_{\text{obs},k}, \vec{t_k}|\rho, \vec{t_j})$ over an N+1 dimensional conical volume around the k-th template which resides in an N-parameter template bank. For O4 bank, N=3 for $\vec{\theta} = \{m_1, m_2, \chi_{eff}\}$.

He notes that the bank has curvature and hence the angular volume of the cone is taken to be the solid angle of the cone (volume element of the bank in mismatch space) at the location of the k-th template: $\Delta\Omega_k$. **This is different from Heather's approach since she took the volume element to be constant**.

Therefore, the integral discussed above can be written as:

$$P(\text{candidate in } \vec{t_k}|H_s = \rho\,\vec{t_j}) \propto \Delta\Omega_k \int_0^\infty \rho^N_{\text{obs},k} d\rho_{\text{obs},k} e^{-\frac{1}{2}(\rho^2+\rho^2_{\text{obs},k})}\mathrm{I}_0(\rho\rho_{\text{obs,k}}\tilde{\mathrm{t_j}}\cdot\tilde{\mathrm{t_k}}) \quad (35)$$

$$\propto \Delta\Omega_k e^{-\frac{1}{2}\rho^2} \sum_{n=0}^{\infty} \frac{(\rho\vec{t_k}\cdot\vec{t_j})^{2n}}{(n!)^2 4^n} \int_0^\infty d\rho_{\text{obs},k} e^{-\frac{1}{2}\rho^2_{\text{obs}}} \rho^{2n+N}_{\text{obs},k} \quad (36)$$

$$\propto \Delta\Omega_k e^{-\frac{1}{2}\rho^2} \sum_{n=0}^{\infty} \frac{(\rho\vec{t_k}\cdot\vec{t_j})^{2n}}{(n!)^2 4^n} \times \Gamma(n+\frac{N}{2}+\frac{1}{2})2^{n+\frac{N}{2}-\frac{1}{2}} \quad (37)$$

$$\propto \Delta\Omega_k 2^{\frac{N-1}{2}}\Gamma\left(\frac{N+1}{2}\right) e^{-\frac{1}{2}\rho^2} \sum_{n=0}^{\infty} \frac{\left(\frac{N+1}{2}\right)^{(n)} \times \left(\frac{(\rho\,\vec{t_k}\cdot\vec{t_j})^2}{2}\right)^n}{1^{(n)} n!} \quad (38)$$

where $a^{(n)} = \frac{\Gamma(a+n)}{\Gamma(a)}$ is the rising factorial (Note, $1^{(n)} = n!$). This series is the expansion of Kummer's confluent hypergeometric function of the first kind, which can be evaluated as

$$P(\text{candidate in } \vec{t_k}|H_s = \rho\,\vec{t_j}) \propto 2^{\frac{N-1}{2}}\Delta\Omega_k\Gamma(\frac{1+N}{2})e^{-\frac{1}{2}\rho^2}{}_1\mathrm{F}_1\left(\frac{1+\mathrm{N}}{2}, 1, \frac{(\rho\tilde{\mathrm{t_k}}\cdot\tilde{\mathrm{t_j}})^2}{2}\right) \quad (39)$$

where ${}_1\mathrm{F}_1$ is Kumer's confluent hypergeometric function of the first kind, using `scipy.special.hyp1f1`.

For N-dimensional banks:

- N=1: one dimensional circular bank for toy model,

$$P(\text{candidate in } \vec{t_k}|H_s = \rho\,\vec{t_j}) \propto \Delta\Omega_k e^{-\frac{1}{2}\rho^2(1-(\vec{t_k}\cdot\vec{t_j})^2)}.$$

- N=3: O4 Bank,

$$P(\text{candidate in } \vec{t_k}|H_s = \rho\,\vec{t_j}) \propto \Delta\Omega_k e^{-\frac{1}{2}\rho^2(1-(\vec{t_k}\cdot\vec{t_j})^2)}(1+\rho^2(\vec{t_k}\cdot\vec{t_j})^2).$$

In his study, the results from the two-phase method for both $N = 1$ and $N = 3$ lie between Chad's and Heather's formulas.

Although Anarya developed his work during O4, it has not yet been incorporated into the O4 mass model. He also conducted a series of studies comparing the two-phase model with Heather's and Chad's models, as well as with a toy-model result (see [5] for details).

3.4 Chad's IR1 Model FIXME!!

Chad's notes on mass-model recovery reform MR:

- Likelihood recovery kernel + per-signal Z normalization (new default):

$$\exp(-\rho^2(1-k^2)/2) \text{ with } Z_j = \sum_k w(k_j k),$$

  so every signal is recovered with total probability 1.

- The historical model (squared kernel, no normalization — measured non-uniform by up to 10x on refined banks) is exactly reproducible via legacy-kernel.

- match-dispersion epsilon: spherical-Gaussian marginalization of an irreducible mismatch scale, modeling bank discreteness + out-of-manifold waveform physics; keeps near-boundary templates honestly ambiguous at any SNR.

For IR1, we are adopting this model for the template migration term in allsky mass-model.

3.5 Code Integration

Chad integrates the analytical solution from Sec. 3.4 of template migration term into sgn-manifold code. The relevant MR is: mass-model recovery reform (Chad refers it as first-order recovery kernel + per-template Z normalization.)

Notes on normalization: Overall, the sgn-manifold code applies a normalization factor to the mass model $P(\vec{t_k} \mid H_s, \rho)$ when outputting the coefficients, ensuring consistency in the units of our calculation.

3.6 Validation: Prior Volume Ratios

As evaluated in Sec. 2.5, we consider the overall prior-volume ratios, including both the source-population and template-migration terms, and compare them with the observed rates.

We estimate the total prior volume associated with the likelihood term $P(\vec{t_k}|H_s, \rho)$ by summing over the template bank:

$$\sum_k P(\vec{t_k}|H_s, \rho),$$

partitioned according to the mass region occupied by each template k.

Since $P(\text{candidate in } \vec{t_k} \mid H_s = \rho\, \vec{t_j})$ is constrained by two conditions:

- (Heather) when $\rho = 0$, both numerator and denominator goes to 0, then $P(\text{candidate in } \vec{t_k}|H_s = \rho\, \vec{t_j}) \to 1$
- (Jolien) when $\rho \to \infty$, the noise has negligible effect, $P(\text{candidate in } \vec{t_k}|H_s = \rho\, \vec{t_j}) \to 1$.

We then consider **a high-SNR scenario with $\rho = 30$ for the given SNR at the evaluation**. For GstLAL O4 Mass Model, in source frame, we get:

$$\sum_{k\in\text{BNS}} P(\vec{t_k} \mid H_s, \rho) : \sum_{k\in\text{NSBH}} P(\vec{t_k} \mid H_s, \rho) : \sum_{k\in\text{BBH}} P(\vec{t_k} \mid H_s, \rho) = 1 : 4 : 15.5$$

With IR1 template migration and χ_{eff} marginalization, in source frame, we have:

$$\sum_{k\in\text{BNS}} P(\vec{t_k} \mid H_s, \rho) : \sum_{k\in\text{NSBH}} P(\vec{t_k} \mid H_s, \rho) : \sum_{k\in\text{BBH}} P(\vec{t_k} \mid H_s, \rho) = 4.8 : 4.0 : 1$$

As Sec. 2.5, we also consider the observable volume set by the detector's horizon distance H_D. Our prior volume then become:

$$\sum_k P(\vec{t_k}|H_s, \rho) H_{D_k}^3,$$

For GstLAL O4 Mass Model, in observable frame, we get:

$$\sum_{k\in\text{BNS}} P(\vec{t_k} \mid H_s, \rho) H_{D_k}^3 : \sum_{k\in\text{NSBH}} P(\vec{t_k} \mid H_s, \rho) H_{D_k}^3 : \sum_{k\in\text{BBH}} P(\vec{t_k} \mid H_s, \rho) H_{D_k}^3 \approx 1 : 20 : 50000$$

With IR1 template migration and χ_{eff} marginalization, in observable frame, we have:

$$\sum_{k\in\text{BNS}} P(\vec{t_k} \mid H_s, \rho) H_{D_k}^3 : \sum_{k\in\text{NSBH}} P(\vec{t_k} \mid H_s, \rho) H_{D_k}^3 : \sum_{k\in\text{BBH}} P(\vec{t_k} \mid H_s, \rho) H_{D_k}^3 = 1 : 3.5 : 108$$

The O4 mass model deviates from the expected rates in both the source and observable frames, while the fixes described in Secs. 2.3 and 3.4 bring the ratios into consistency with the observed rates from GWTC-5. Both fixes are introduced in the IR1 mass-model, and is currently running with MDC50. In the following, I am going to refer this mass-model version as refined IR1 Mass-Model.

4 Summary

In summary, Anarya's O4 Mass-Model have several distinct differences from Heather's O1 Mass-Model:

1. Introduce the NSBH source
2. Adopt a single mass distribution for all CBC sources
3. Adopt a different model for template migration term

The refined IR1 mass-model addresses two issues based on Anarya's O4 Mass-Model:

1. Add a spin distribution for CBC sources tailored to the O4 Bank boundary
2. address errors in template migration term

In Heather's O1 evaluation, she estimated a 6-16% VT improvement by introducing the mass-model term than not, and 2-7% more compared with a BNS-only bank search.

We did not evaluate Anarya's VT improvement in compare to Heather's. However, there is a clear improvement of refined IR1 mass-model compared with O4 mass-model. See Sec. 5 for details.

The latest code of mass-model in sgn-manifold can be found at: line

5 Results

To evaluate our IR1 refined mass-model, We rerank the O4a/GWTC-4.0 data using it. Here, we only discuss the sensitivity improvement with our fixes. We put the one-to-one comparison with GWTC-4.0 GstLAL Catalog into App. A. Compared with the published GWTC-4.0 VT (Figure 4), the result shows **a 28% VT improvement on BNS recovery in O4a**:

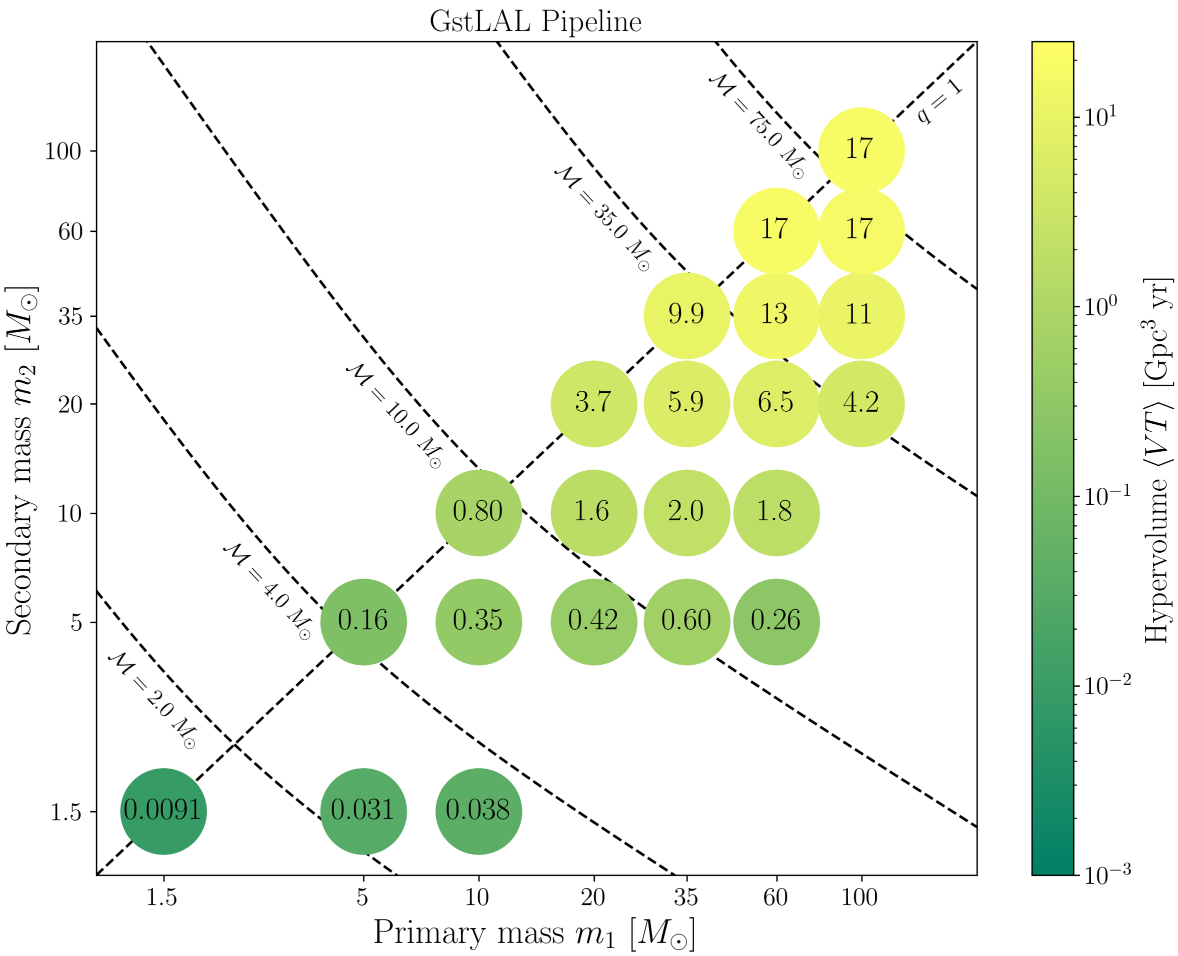

Figure 3: R&P VT estimated from reranked O4a result with refined mass-model, using the same injection sample and weighting distribution.

We also find that, **under this refined mass model, the significance of GW231109_235456 corresponds to a FAR of 2 per year**.

6 Next Step in Progress: MDC50

If possible, we would like to evaluate how the refined IR1 mass model performs in the MDC50, particularly in terms of BNS injection recovery. Victoria would strongly suggest writing this as a paper and put it on arXiv, which clarify some misunderstanding we have regarding event

GW231109_235456. We should also present this result with the collaboration and consider a published GWTC-6.1 with a rerank of O4a, O4b, and O4c data.

References

[1] H. Fong, "From simulations to signals: Analyzing gravitational waves from compact binary coalescences," PhD thesis, University of Toronto, 2009.

[2] L. Tsukada et al., "Improved ranking statistics of the GstLAL inspiral search for compact binary coalescences," Phys. Rev. D **108**, 043004 (2023), arXiv:2305.06286.

[3] K. Cannon, C. Hanna, and J. Peoples, "Likelihood-Ratio Ranking Statistic for Compact Binary Coalescence Candidates with Rate Estimation," arXiv:1504.04632 [astro-ph.IM] (2015).

[4] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration, "GWTC-5.0: Population Properties of Merging Compact Binaries," arXiv:2605.27226 [astro-ph.HE] (2026).

[5] A. Ray, "Two-Phase Mass Model Notes," unpublished notes (2025). `https://www.overleaf.com/read/pccsgkmzdqyz#ad7f5c`

[6] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration, "GWTC-4.0: Updating the Gravitational-Wave Transient Catalog with Observations from the First Part of the Fourth LIGO-Virgo-KAGRA Observing Run," arXiv:2508.18082 [gr-qc] (2025).

Thanks Claude for checking the math of this notes.

A Reranked Candidate List

The open box for the reranked result using the refined mass model can be found here: link. We perform a one-to-one Comparison with GWTC-4 GstLAL catalog and find:

Table 2: Catalog of newly significant events (FAR < 1 per 5 months).

Rank	GPS time	GraceDB ID	FAR (Hz)	Previous FAR (Hz)	Network SNR	Total mass ($M_\odot$)	Chirp mass ($M_\odot$)	IFOs	Notes
1	1 384 482 081.03	S231120e	2.45×10^{-9}	1.13×10^{-7}	9.59	21.05	9.03	H1,L1	
2	1 387 398 397.01	S231223bg	4.29×10^{-9}	2.11×10^{-7}	10.04	22.45	9.73	H1	GstLAL-only
3	1 387 353 073.32	S231223aa	5.11×10^{-9}	2.88×10^{-7}	9.31	30.75	9.25	H1,L1	GstLAL-only
4	1 377 524 799.74	S230831bn	2.54×10^{-8}	–	8.78	23.34	10.15	H1,L1	
5	1 381 707 055.02	S231018by	5.94×10^{-8}	–	8.69	23.84	10.11	H1,L1	
6	1 383 609 314.06	S231109ci	5.95×10^{-8}	–	9.64	3.01	1.31	H1,L1	
7	1 369 017 515.86	S230525a	6.90×10^{-8}	–	10.74	33.01	13.45	L1	GstLAL-only

Table 3: Catalog of insignificant events recovered significantly in the previous open box (FAR < 1 per 5 months).

Num.	GPS time	GraceDB ID	FAR (Hz)	Previous FAR (Hz)	Network SNR	Total mass ($M_\odot$)	Chirp mass ($M_\odot$)	IFOs	Notes
1	1 374 568 606.88	S230728ap	–	4.51×10^{-8}	13.12	315.39	136.84	H1,L1	GstLAL-only
2	1 372 835 957.44	S230708z	1.30×10^{-7}	4.52×10^{-8}	8.11	107.81	46.62	H1,L1	
3	1 376 780 635.50	S230822bm	1.83×10^{-7}	4.56×10^{-8}	8.17	170.15	73.89	H1,L1	

B Rerank workflow with Mass-Model

In the workflow/DAG, the mass-model file should be the file specified by the given "prior.mass-model" in a complete ranking workflow. However, some hard-coded logic in the 'ProcessParameter' table forces the mass-model file to be loaded from the 'ProcessParameter' table rather than from the configuration file, presumably to ensure workflow consistency. This makes it inconvenient to rerank events with a new mass model using low-latency data product. It does not seem to affect the rerank workflow using offline data product. Leo and Prathamesh are working on addressing this issue in the SGNL workflow and STRIKE.

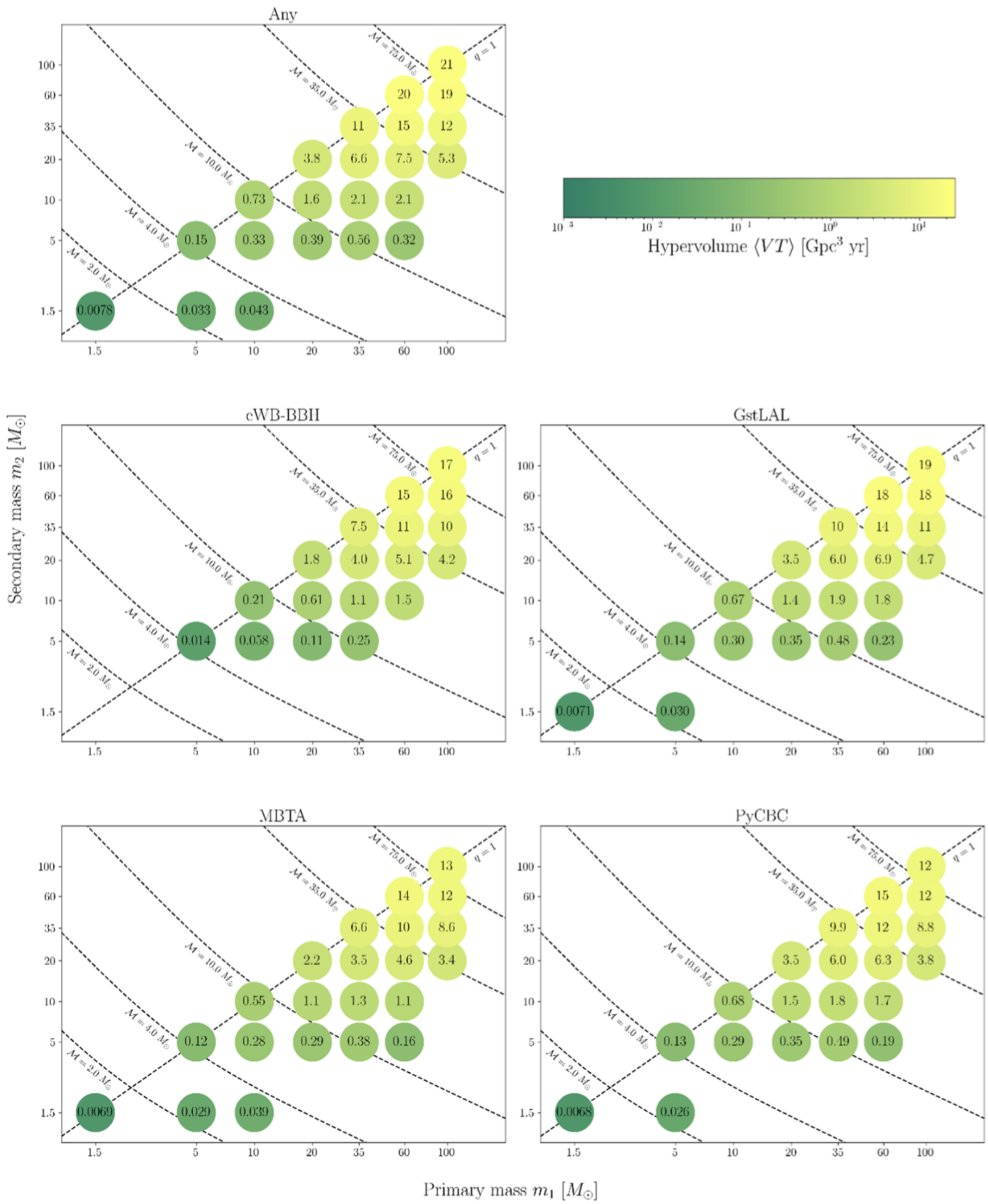

Figure 4: Published GWTC-4.0 VT result [6], with O4 Mass-Model for GstLAL result.