Bayesian QNM search on black hole ringdown modes (applied to GW150914)

Reinhard Prix[1, a]

[1] *Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, D-30167 Hannover, Germany*

(Dated: 2016-04-06 14:40:28 +0200; commitID: b6c2caf-CLEAN; LIGO-T1500618-v4-)

We quantify the evidence for and estimate the parameters of QNM 'ringdown' in the GW150914 event. This is done by Bayesian hypothesis testing and parameter estimation using a QNM ringdown model $s(t \geq t_0) = A\, e^{-\frac{t-t_0}{\tau}} \cos\left(2\pi f_0 (t - t_0) + \phi_0\right)$ with unknown amplitude A, initial phase ϕ_0, frequency f_0 and decay time τ, as a function of the QNM start-time t_0. Using a Gaussian-isotropic prior on $\{\mathcal{A}_\mathrm{s} = -A \sin\phi_0,\, \mathcal{A}_\mathrm{c} = A\cos\phi_0\}$ we can approximate the Bayes factor by analytically marginalizing over $\{A, \phi_0\}$, leaving an explicit template search over $\{f_0,\, \tau\}$.

A. Changelog

A number of changes compared to the previous version (v3,v3+) that went into the submitted paper (arXiv:1602.03841).

- bug 1: previous results suffered from a small bug in (old version of) octapps `FourierTransform` wrapper, which contained an "off-by-one" type error, leading to slight de-phasing of the data time-series
- bug 2: previously the 'SNR-term' in the likelihood $\langle s|s\rangle$ was computed over the full SFT frequency band $[10, 2000]$ Hz, instead of the relevant narrow-banded data $[30, 1000]$ Hz actually used in the search-term $\langle x|s\rangle$
- expanded band-passed data-range from $[30, 1000]$ Hz to $[30, 1900]$ Hz (but avoid lower and upper boundaries near 10Hz and 2000Hz respectively, where pwelch() psd estimation trails off).
- PSD estimation: avoid using the data-segment of 8 s containing GW150914 (or any injections) to avoid affecting the PSD estimate
- slight inconsistency: previous results used $t_\mathrm{M} = 1126259462.42285$, while the paper stated $t_\mathrm{M} = 1126259462.423$ (rounded to ms). While small, this *does* change the numbers (and posteriors) slightly, especially towards the end (+7ms), due to the exponentiall fall-off of the Bayes-factor with offsets from t_0. The new results use the merger time rounded to ms-accuracy as quoted in the paper.
- Improved: use *maximum-posterior* (rather than partially maximum-likelihood) estimate of amplitude parameters and resulting SNR. Essentially makes no relevant numerical difference on actual signals, though.
- Improved: substantial speedup of code: larger sampling of "off-source" noise distributions on actual data and on Gaussian white noise
- Improved: added injection feature of QNM signals, allowing quantified test of accuracy of parameter-estimation and posterior coverage
- improved iso-probability contour estimation accuracy

All results shown in this version of the notes were produced using the ringdown pipeline in gitLab, version aebcd96a.

I. INTRODUCTION

Signal model: damped sinusoid starting at t_0:

$$s(t;\, A, \phi_0\,, t_0, \tau, f_0) = A\, w(t - t_0,\, \tau)\, \cos\left(2\pi\, f_0\, (t - t_0) + \phi_0\right)\,, \tag{1}$$

$$w(t,\tau) = \begin{cases} \exp\left(-\frac{t}{\tau}\right) & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases} \tag{2}$$

[a] Reinhard.Prix@ligo.org

Rewrite waveform in terms of two unknown amplitudes $\mathcal{A}_{\rm s} = -A\sin\phi_0$, $\mathcal{A}_{\rm c} = A\cos\phi_0$ in "JKS" factorization [1, 2]:

$$s(t;\theta) = \mathcal{A}_{\rm s}\, h_{\rm s}(t;\lambda) + \mathcal{A}_{\rm c}\, h_{\rm c}(t;\lambda)\,, \tag{3}$$

with *basis functions*

$$h_{\rm s}(t;\lambda) \equiv w(t-t_0,\tau)\,\sin(2\pi\, f_0\,(t-t_0))\,, \tag{4}$$
$$h_{\rm c}(t;\lambda) \equiv w(t-t_0,\tau)\,\cos(2\pi\, f_0\,(t-t_0))\,, \tag{5}$$

and the set of signal parameters separating into "amplitude parameters" $\vec{\mathcal{A}}$ and "evolution parameters" λ:

$$\theta \equiv \{\vec{\mathcal{A}},\lambda\}\,,\quad \vec{\mathcal{A}} \equiv (\mathcal{A}_{\rm s},\mathcal{A}_{\rm c})\,,\quad \lambda \equiv \{t_0,\tau,f_0\}\,. \tag{6}$$

The likelihoods for Gaussian (colored) noise $\mathcal{H}_{\rm G}$ and the ringdown model $\mathcal{H}_{\rm S}$ are

$$P\left(x|\mathcal{H}_{\rm G}\right) = c\,e^{-\frac{1}{2}\langle x|x\rangle}\,, \tag{7}$$
$$P\left(x|\mathcal{H}_{\rm S},\theta\right) = c\,e^{-\frac{1}{2}\langle x-s(\theta)|x-s(\theta)\rangle}\,, \tag{8}$$

with the multi-detector scalar product (over detector index X) defined as

$$\langle x|y\rangle \equiv \sum_X \langle x^X|y^X\rangle = \sum_X 2\int_{-\infty}^{\infty}\frac{\widetilde{x}^X(f)\,\widetilde{y}^{*X}(f)}{S_X(f)}\,df\,, \tag{9}$$

with the per-detector single-sided noise PSD $S_X(f)$ (which results in the prefactor of '2'). The (marginal) likelihood for the signal model can be expressed as

$$P\left(x|\mathcal{H}_{\rm S}\right) = \int P\left(x|\mathcal{H}_{\rm S},\theta\right)\,P\left(\theta|\mathcal{H}_{\rm S}\right)\,d\theta\,, \tag{10}$$

and the corresponding Bayes factor (or marginal likelihood ratio)

$$B_{\rm S/G}(x) \equiv \frac{P\left(x|\mathcal{H}_{\rm S}\right)}{P\left(x|\mathcal{H}_{\rm G}\right)} = \int \mathcal{L}(x;\theta)\,P\left(\theta|\mathcal{H}_{\rm S}\right)\,d\theta\,, \tag{11}$$

with the likelihood-ratio *function*

$$\mathcal{L}(x;\theta) \equiv \frac{P\left(x|\mathcal{H}_{\rm S},\theta\right)}{P\left(x|\mathcal{H}_{\rm G}\right)} = \exp\left(\langle x|s\rangle - \frac{1}{2}\langle s|s\rangle\right)\,. \tag{12}$$

We can further introduce a partially-marginalized (over unknown amplitude parameters $\vec{\mathcal{A}}$) Bayes factor $B_{\rm S/G}(x;\lambda)$ as

$$\begin{aligned}
B_{\rm S/G}(x) &= \int \mathcal{L}(x;\theta)\,P\left(\theta|\mathcal{H}_{\rm S}\right)\,d\theta \\
&= \int \mathcal{L}(x;\vec{\mathcal{A}},\lambda)\,P\left(\vec{\mathcal{A}}|\lambda,\mathcal{H}_{\rm S}\right)\,P\left(\lambda|\mathcal{H}_{\rm S}\right)\,d^2\mathcal{A}\,d\lambda \\
&= \int B_{\rm S/G}(x;\lambda)\,P\left(\lambda|\mathcal{H}_{\rm S}\right)\,d\lambda\,,
\end{aligned} \tag{13}$$
$$B_{\rm S/G}(x;\lambda) \equiv \int \mathcal{L}(x;\mathcal{A},\lambda)P\left(\mathcal{A}|\lambda,\mathcal{H}_{\rm S}\right)\,d^2\mathcal{A}\,. \tag{14}$$

II. COMPUTING THE BAYES FACTOR $B_{\rm S/G}$

A. Expressing the SNR^2: $\langle s|s\rangle$

We assume the data $x^X(t)$ from the different detectors has been time-shifted and corrected for antenna-pattern effects, in such a way that the expected signal $s(t)$ would be identical in all data streams, so we can assume the templates to be independent of detector, and write

$$\langle s|s\rangle = \sum_X 2\int_{-\infty}^{\infty} \frac{|\widetilde{s}(f)|^2}{S_X(f)}\,df \tag{15}$$

$$= 2N_{\rm det} \int \frac{|\widetilde{s}(f)|^2}{\mathcal{S}(f)}\,df\,, \tag{16}$$

where the multi-detector noise floor $\mathcal{S}(f)$ is defined as the harmonic mean

$$\mathcal{S}^{-1}(f) \equiv \frac{1}{N_{\rm det}} \sum_X S_X^{-1}(f)\,. \tag{17}$$

Using the factorization of Eq. (3), which in frequency domain yields

$$\widetilde{s}(f;\theta) = \mathcal{A}_{\rm s}\,\widetilde{h}_{\rm s}(f;\lambda) + \mathcal{A}_{\rm c}\,\widetilde{h}_{\rm c}(f;\lambda)\,, \tag{18}$$

we can further write this as

$$\langle s|s\rangle = \vec{\mathcal{A}}\cdot\mathcal{M}(\lambda)\cdot\vec{\mathcal{A}}, \tag{19}$$

with

$$\mathcal{M}(\lambda) \equiv 2N_{\rm det}\begin{pmatrix} I_{\rm s} & I_{\rm sc} \\ I_{\rm sc} & I_{\rm c}\end{pmatrix}, \tag{20}$$

$$I_{\rm s}(\lambda) = 2\int_0^{\infty} \frac{|\widetilde{h}_{\rm s}(f)|^2}{\mathcal{S}(f)}\,df\,, \tag{21}$$

$$I_{\rm c}(\lambda) = 2\int_0^{\infty} \frac{|\widetilde{h}_{\rm c}(f)|^2}{\mathcal{S}(f)}\,df\,, \tag{22}$$

$$I_{\rm sc}(\lambda) = 2\int_0^{\infty} \frac{\Re[\widetilde{h}_{\rm s}(f)\,\widetilde{h}_{\rm c}^{*}(f)]}{\mathcal{S}(f)}\,df\,. \tag{23}$$

The Fourier transforms $\widetilde{h}_{\rm s}, \widetilde{h}_{\rm c}$ of the signal basis functions can be computed analytically

$$\widetilde{h}_{\rm s}(f;\lambda) = \tau\,\frac{2\pi f_0\,\tau}{1 + i\,4\pi\,f\,\tau - 4\pi^2(f^2 - f_0^2)\tau^2}\,e^{-i2\pi f t_0}\,, \tag{24}$$

$$\widetilde{h}_{\rm c}(f;\lambda) = \tau\,\frac{1 + i\,2\pi\,f\,\tau}{1 + i\,4\pi\,f\,\tau - 4\pi^2(f^2 - f_0^2)\tau^2}\,e^{-i2\pi f t_0}\,. \tag{25}$$

B. Expressing the "matched filter" $\langle x|s(\theta)\rangle$

Note that in the scalar product involving the data x^X (assume time-shifted and antenna-pattern corrected) we can conveniently absorb the frequency-dependend noise-floors $S_X(f)$ by *over-whitening* the data, i.e. we define

$$\widetilde{y}^X \equiv \frac{\widetilde{x}^X(f)}{S_X(f)}\,,\quad \widetilde{y} \equiv \sum_X \widetilde{y}^X\,. \tag{26}$$

Here we define t to the arrival time in the 'H1' detector. We apply a detector-specific time-delay of adding 7 ms to L1 arrival time in the case of GW150914) and antenna-pattern corrections (a factor of -1 of L1 wrt H1) to the *data* $y^X(t)$, such that we can assume the putative signal waveform in the data to be in phase and of (approximately) same amplitude and phase. This means that we can assume a detector-independent template $s^X(t) = s(t)$, which allows us to write the scalar product in time-domain form

$$\langle x|s(\theta)\rangle = \sum_X 2\int_{-\infty}^{\infty} \widetilde{y}^X(f)\,\widetilde{s}^{*X}(f)\,df \tag{27}$$

$$= 2\int_{-\infty}^{\infty} \widetilde{y}(f)\,\widetilde{s}^{*}(f)\,df \tag{28}$$

$$= 2\int_{t_0}^{t_0+T} y(t)\,s(t)\,dt\,, \tag{29}$$

where $y(t)$ is the overwhitened summed-IFO timeseries, i.e. the inverse Fourier-transform of $\widetilde{y}(f)$, and where $T \gg \tau$ is some duration long enough so that $s(t_0+T) \approx 0$, e.g. $T = 5\tau$.

Using Eq. (3) we can further write

$$\langle x|s(\theta)\rangle = \vec{\mathcal{A}} \cdot \vec{x}(\lambda) = \mathcal{A}_{\mathrm{s}}\, x_{\mathrm{s}}(\lambda) + \mathcal{A}_{\mathrm{c}}\, x_{\mathrm{c}}(\lambda)\,, \tag{30}$$

with

$$x_{\mathrm{s}}(\lambda) \equiv 2\int_{t_0}^{t_0+T} y(t)\, h_{\mathrm{s}}(t;\lambda)\, dt\,, \tag{31}$$

$$x_{\mathrm{c}}(\lambda) \equiv 2\int_{t_0}^{t_0+T} y(t)\, h_{\mathrm{c}}(t;\lambda)\, dt\,. \tag{32}$$

Note that it will be convenient to write

$$h_{\mathrm{exp}} \equiv h_{\mathrm{c}}(t;\lambda) - i\, h_{\mathrm{s}}(t;\lambda) = e^{-\Delta t/\tau}\, e^{-i\,2\pi\, f_0\,\Delta t} = e^{-\Delta t\, \varpi}\,, \tag{33}$$

with $\Delta t \equiv t - t_0$ and complex frequency ϖ defined as

$$\varpi \equiv \frac{1}{\tau} + i\, 2\pi f_0\,, \tag{34}$$

and so we obtain the complex matched-filter as

$$F \equiv x_{\mathrm{c}} - i\, x_{\mathrm{s}} = 2\int_0^T y(t_0+\Delta t)\, e^{-\Delta t\,\varpi}\, d\Delta t\,, \tag{35}$$

which is the Laplace transform of the over-whitened data $y(t)$.

C. Marginalizing over unknown amplitudes $\{\mathcal{A}_{\mathrm{s}}, \mathcal{A}_{\mathrm{c}}\}$

Combining these expressions in the likelihood-ratio function of Eq. (12), we can write this as

$$\ln\mathcal{L}(x;\theta) = \langle x|s\rangle - \frac{1}{2}\langle s|s\rangle \tag{36}$$

$$= -\frac{1}{2}\,\vec{\mathcal{A}}\cdot\mathcal{M}\cdot\vec{\mathcal{A}} + \vec{\mathcal{A}}\cdot\vec{x}\,, \tag{37}$$

i.e. a 2-dimensional Gaussian in $\{\mathcal{A}_{\mathrm{s}}, \mathcal{A}_{\mathrm{c}}\}$ with covariance matrix $\mathcal{M}^{-1}$. This can be marginalized analytically to yield $B_{\mathrm{S/G}}(x;\lambda)$ in Eq. (13) for a suitable choice of prior $P\left(\vec{\mathcal{A}}|\lambda,\mathcal{H}_{\mathrm{S}}\right)$.

First we assume that the amplitude prior is *logically* independent of the evolution parameters λ, which simply expresses ignorance about a possible dependence, not a claim about *physical* independence [3], i.e. $P\left(\vec{\mathcal{A}}|\lambda,\mathcal{H}_{\mathrm{S}}\right) = P\left(\vec{\mathcal{A}}|\mathcal{H}_{\mathrm{S}}\right)$

Further we use a simple isotropic Gaussian amplitude prior, which expresses ignorance about the initial phase ϕ_0, and posits an (unknown) characteristic scale H for the amplitude A, namely

$$P\left(\vec{\mathcal{A}}|\mathcal{H}_{\mathrm{S}},H\right) = \frac{1}{2\pi\, H^2}\, e^{-\frac{1}{2}\vec{\mathcal{A}}\cdot\vec{\mathcal{A}}/H^2}\,, \tag{38}$$

which implies a prior on the amplitude A (marginalized over ϕ_0):

$$P\left(A|\mathcal{H}_{\mathrm{S}},H\right) = \frac{A}{H^2}\, e^{-\frac{A^2}{2H^2}}\,. \tag{39}$$

Using this Gaussian amplitude prior we find the H-dependent Bayes factor:

$$B_{\mathrm{S/G}}(x;\lambda,H) \equiv \frac{P\left(x|\mathcal{H}_{\mathrm{S}},\lambda,H\right)}{P\left(x|\mathcal{H}_{\mathrm{G}}\right)} \tag{40}$$

$$= \frac{1}{2\pi H^2}\int e^{-\frac{1}{2}\,\vec{\mathcal{A}}\cdot\gamma^{-1}\cdot\vec{\mathcal{A}}+\vec{\mathcal{A}}\cdot\vec{x}}\, d^2\mathcal{A} \tag{41}$$

$$= \frac{\sqrt{\det\gamma}}{H^2}\, e^{\frac{1}{2}\,\vec{x}\cdot\gamma\cdot\vec{x}} \tag{42}$$

with

$$\gamma^{-1}(\lambda) \equiv \mathcal{M} + H^{-2}\,\mathbb{I} = \begin{pmatrix} I_{\mathrm{s}} + H^{-2} & I_{\mathrm{sc}} \\ I_{\mathrm{sc}} & I_{\mathrm{c}} + H^{-2}\,. \end{pmatrix}, \tag{43}$$

and determinant

$$\det\gamma^{-1} = \left(I_{\mathrm{s}} + H^{-2}\right)\left(I_{\mathrm{c}} + H^{-2}\right) - I_{\mathrm{sc}}^2 \tag{44}$$

$$= \det\mathcal{M} + H^{-2}\,\mathrm{tr}\mathcal{M} + H^{-4}\,, \tag{45}$$

inverse

$$\gamma(\lambda) = \frac{1}{\det\gamma^{-1}}\begin{pmatrix} I_{\mathrm{c}} + H^{-2} & -I_{\mathrm{sc}} \\ -I_{\mathrm{sc}} & I_{\mathrm{s}} + H^{-2} \end{pmatrix}, \tag{46}$$

and

$$\frac{\sqrt{\det\gamma}}{H^2} = \left[H^4\,\det\mathcal{M} + H^2\,\mathrm{tr}\mathcal{M} + 1\right]^{-1/2}\,. \tag{47}$$

We note that in the limit $H \to 0$ we have $\gamma^{-1} \to H^{-2}\mathbb{I}$, so $\gamma \to 0$, and $\sqrt{\det\gamma}/H^2 \to 1$, therefore $B_{\mathrm{S/G}} \to 1$. The signal hypothesis becomes indistinguishable from the noise hypothesis if signal amplitudes are assumed to be vanishingly small. In the opposite limit of $H \gg 1$, we find $\gamma \to \mathcal{M}^{-1}$, and $\sqrt{\det\gamma}/H^2 \to 1/(H^2\sqrt{\det\mathcal{M}})$, which is equivalent to the "$\mathcal{F}$-statistic" for finite H, but $B_{\mathrm{S/G}} \to 0$ for $H \to \infty$, as the prior volume gets increasingly thinly spread out, resulting in an "Occam factor" effect disfavoring the signal hypothesis.

1. *Marginalizing unknown scale H*

The most robust way to deal with the unknown scale parameter H is to marginalize this out using a Jeffreys prior $\propto 1/H$. Given that we *roughly* know the scale of H to fall somewhere in $H \in [2, 10] \times 10^{-22}$, we can simply discretize the corresponding marginalization integrals on a few points H_i, allowing us to normalize this discrete "hyper-prior" as

$$P\left(H_i|\mathcal{H}_{\mathrm{S}}\right) = c\,\frac{1}{H_i},, \quad \text{with} \quad c^{-1} = \sum_i H_i^{-1}\,. \tag{48}$$

We can therefore estimate the unknown H parameter from the data via Eq. (40), namely

$$P\left(H|x\right) = \int P\left(\lambda, H|x, \mathcal{H}_{\mathrm{S}}\right)\,d\lambda \tag{49}$$

$$= \left[\int P\left(x|\mathcal{H}_{\mathrm{S}}, \lambda, H\right)\,P\left(\lambda|\mathcal{H}_{\mathrm{S}}\right)\,d\lambda\right]\,P\left(H|\mathcal{H}_{\mathrm{S}}\right) \tag{50}$$

$$\propto B_{\mathrm{S/G}}(x; H)\,P\left(H|\mathcal{H}_{\mathrm{S}}\right) \tag{51}$$

$$= c\,\frac{1}{H_i}\,B_{\mathrm{S/G}}(x; H_i)\,. \tag{52}$$

Furthermore, we can compute the H-independent Bayes factor and posterior by marginalizing over H via

$$P\left(\lambda|\mathcal{H}_{\mathrm{S}}, x\right) \propto B_{\mathrm{S/G}}(x; \lambda) \tag{53}$$

$$= \int B_{\mathrm{S/G}}(x; \lambda, H) P\left(H|\mathcal{H}_{\mathrm{S}}\right)\,dH \tag{54}$$

$$= c\sum_i \frac{1}{H_i}\,B_{\mathrm{S/G}}(x; \lambda, H_i)\,. \tag{55}$$

III. PARAMETER ESTIMATION

The posterior is

$$\begin{aligned} P(\theta|x, \mathcal{H}_\mathrm{S}) &\propto P(x|\mathcal{H}_\mathrm{S}, \theta)\ P(\theta|\mathcal{H}_\mathrm{S}) \\ &\propto \mathcal{L}(x;\theta)\, P(\mathcal{A}|\mathcal{H}_\mathrm{S})\ P(\lambda|\mathcal{H}_\mathrm{S}) \\ &\propto \exp\left[-\frac{1}{2}\,\vec{\mathcal{A}}\cdot\gamma^{-1}\cdot\vec{\mathcal{A}} + \vec{\mathcal{A}}\cdot\vec{x}\right]\ P(\lambda|\mathcal{H}_\mathrm{S})\,. \end{aligned} \tag{56}$$

where in the first line we have dropped the normalization $1/P(x|\mathcal{H}_\mathrm{S})$, and in the second line we dropped $P(x|\mathcal{H}_\mathrm{G})$, as both are independent of θ. After marginalization over $\{\mathcal{A}_\mathrm{s}, \mathcal{A}_\mathrm{c}\}$ we recover again the Bayes factor $B_{\mathrm{S/G}}(x;\lambda)$ of (40), i.e.

$$P(\lambda|x, \mathcal{H}_\mathrm{S}) \propto B_{\mathrm{S/G}}(x;\lambda)\,. \tag{57}$$

Finding the maximum-posterior estimates (MPE) for $\mathcal{A}_\mathrm{s}, \mathcal{A}_\mathrm{c}$ at fixed λ from (56) yields

$$\vec{\mathcal{A}}' = \gamma(\lambda)\cdot\vec{x}\,. \tag{58}$$

Note that this depends on the unknown scale parameter H, but we will simplify this by using the MPE estimator for H from Eq. (49), i.e. H_{MPE} that maximizes $B_{\mathrm{S/G}}(x;H)/H$. We will further evaluate this at the MPE values found for λ in the Bayes-factor search using (57). From these we can obtain $A' = \sqrt{\mathcal{A}_\mathrm{s}^2 + \mathcal{A}_\mathrm{c}^2}$ and $\phi_0' = -\tan^{-1}\left(\frac{\mathcal{A}_\mathrm{s}}{\mathcal{A}_\mathrm{c}}\right)$.

From this can can also estimate an “SNR” in the MPE template, by substituting the MPE amplitude parameters into the SNR expression Eq. (19), i.e.

$$\rho_0^2 = \vec{\mathcal{A}}'\mathcal{M}\vec{\mathcal{A}}'\,. \tag{59}$$

IV. QNM SEARCH APPLIED TO GW150914

A. Prior choices

Isotropic 2D Gaussian amplitude prior (38) on $\{\mathcal{A}_\mathrm{s}, \mathcal{A}_\mathrm{c}\}$ with characteristic amplitudes $H = [2:10]\times 10^{-22}$, which corresponds to an isotropic prior in ϕ_0 and an A-prior of (39), as shown here:

B. Data preparation

Use band-passed 1800s-SFTs (covering GW150914) containing $[10, 2000]$ Hz from each detector, band-passed to $[30, 1900]$ Hz to avoid boundary regions where pwelch() PSD estimation trails off. From this we extracted a $T = 8\,\mathrm{s}$ timeseries $x^X(t)$ centered on the event to analyse. Time-shifted L1 data by (delaying it by $+7\,\mathrm{ms}$) and multiplied it by (-1) to account for the inverse detector response. We estimate the (single-sided) PSD S_X by a standard welch method using 8s windows over the 1800s SFT data.

The resulting original spectrum (top row), rms-whitened spectrum (middle row) and over-whitened spectrum (bottom row) are shown in the following plots, for H1 and L1, respectively.

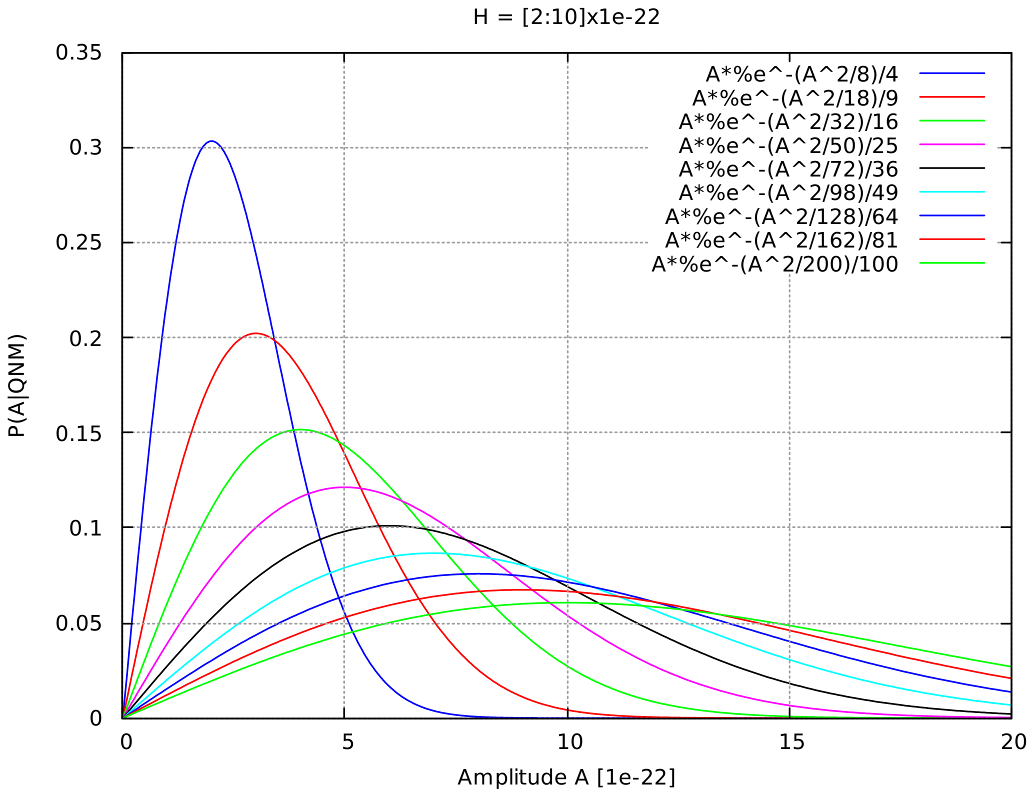
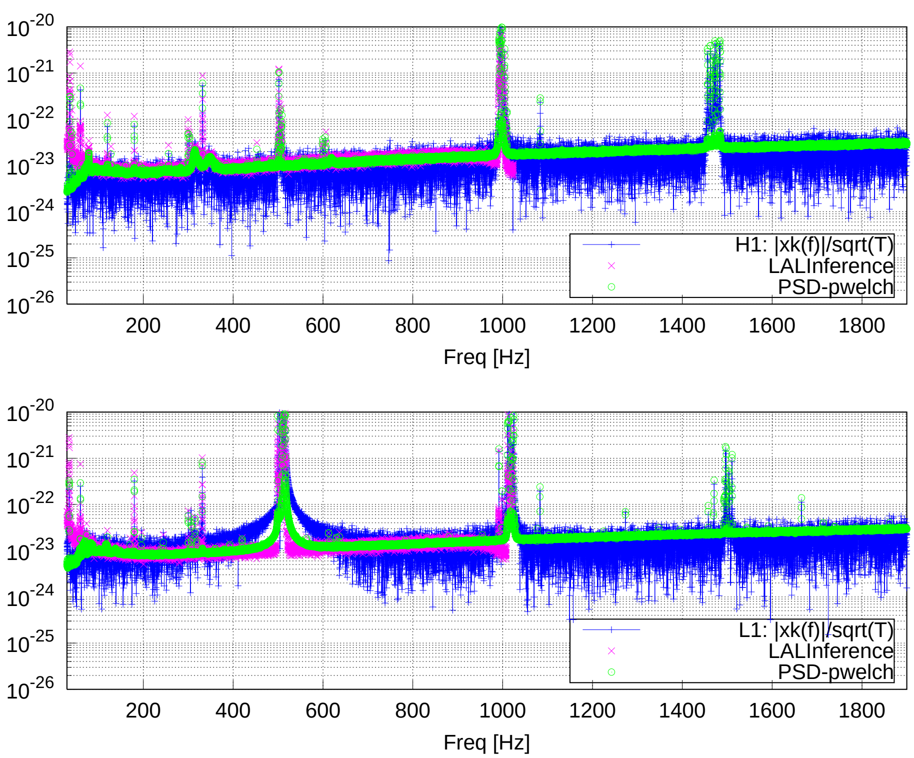
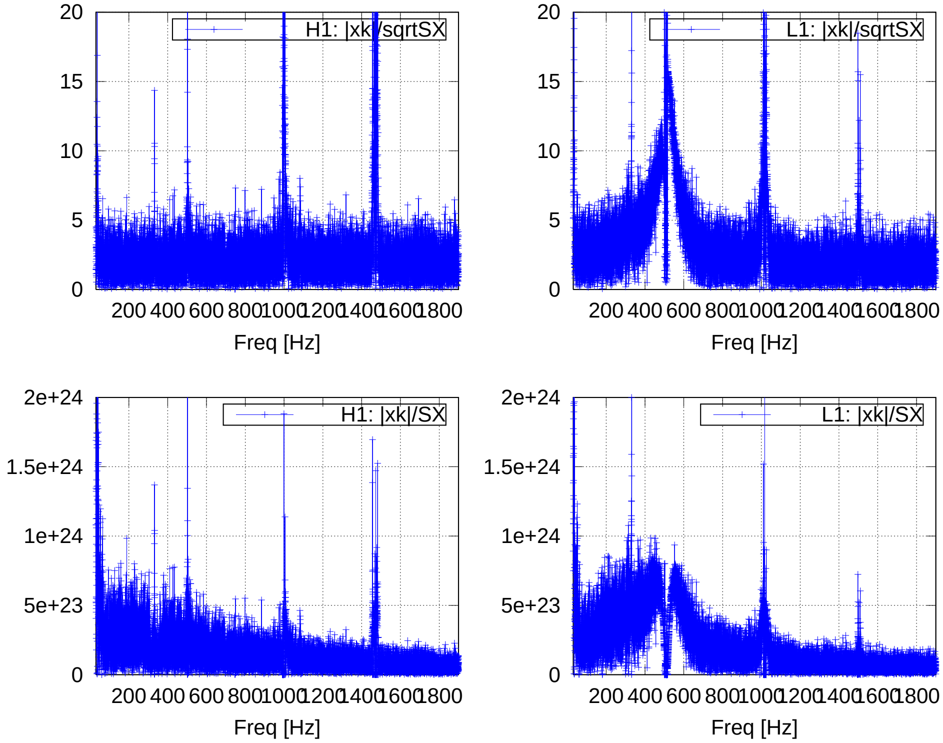


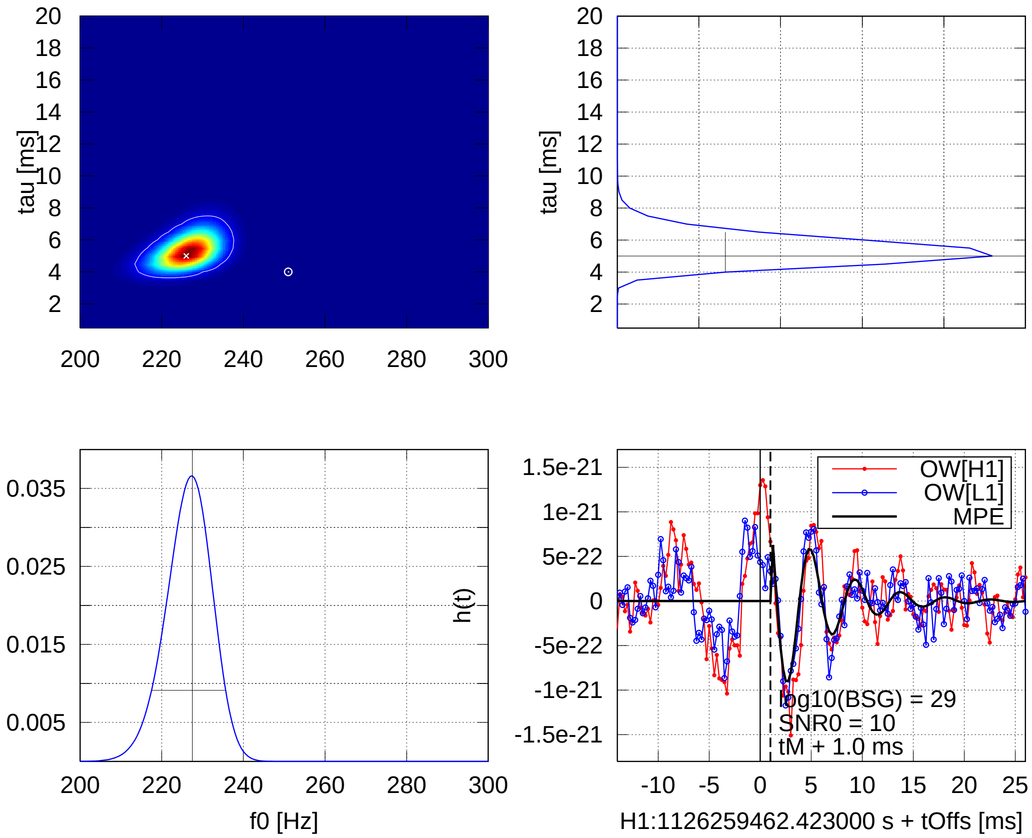
FIG. 1. Amplitude prior as a hyper-prior (superposition) of several 2D-Gaussian distributions with different scale parameters $H \in [2:10] \times 10^{-22}$

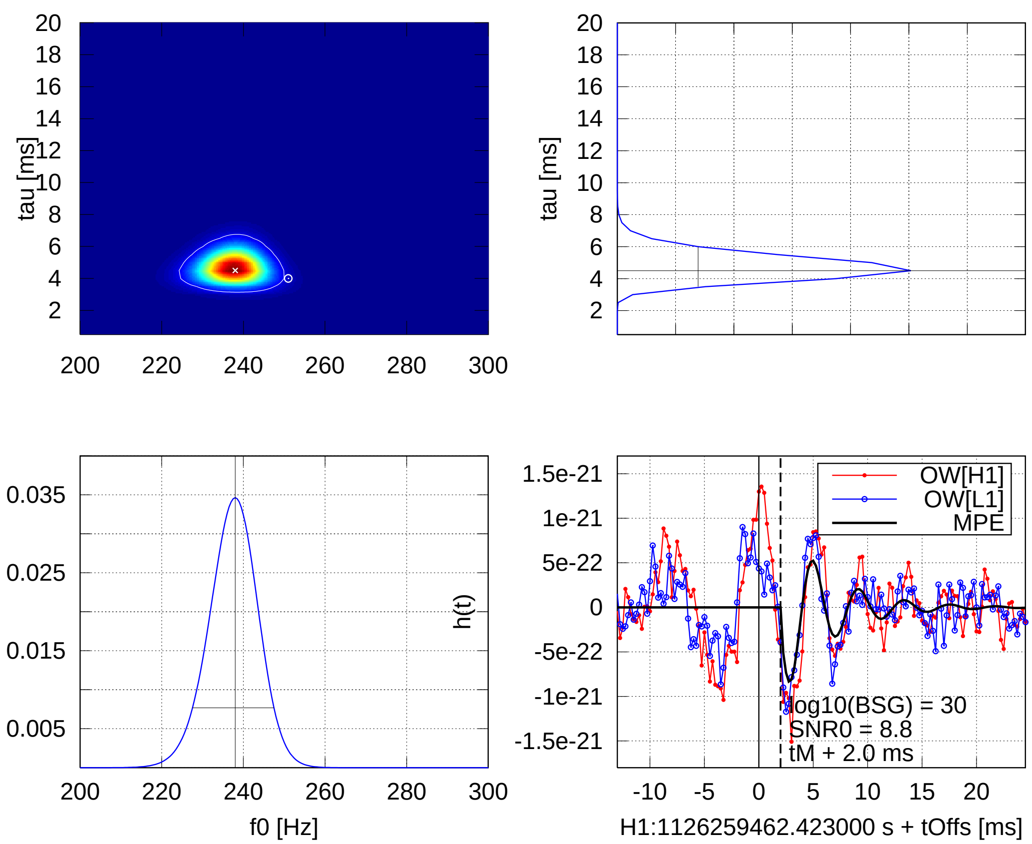


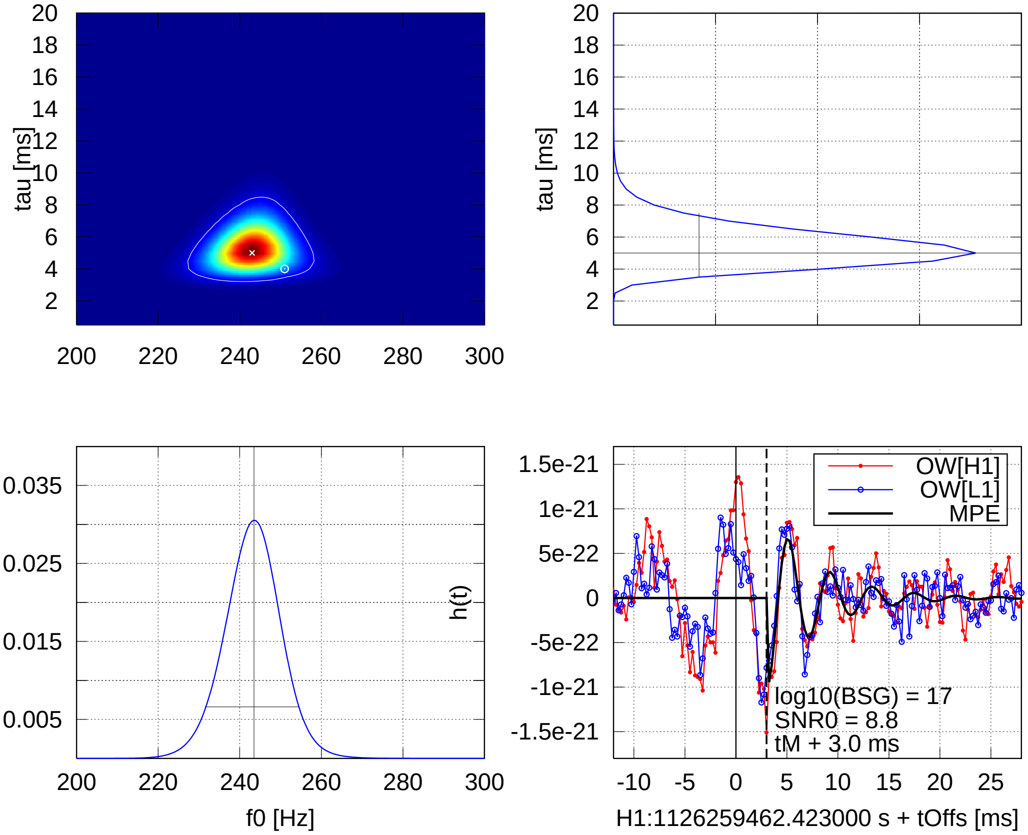


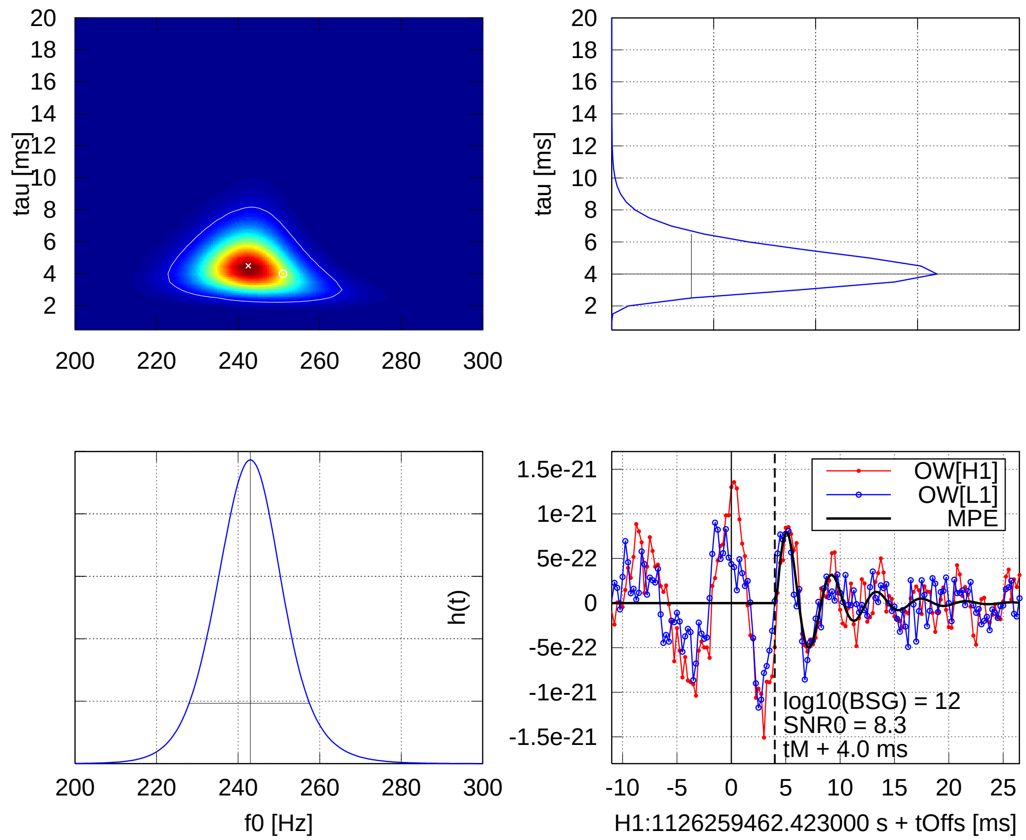
C. Search results on GW150914

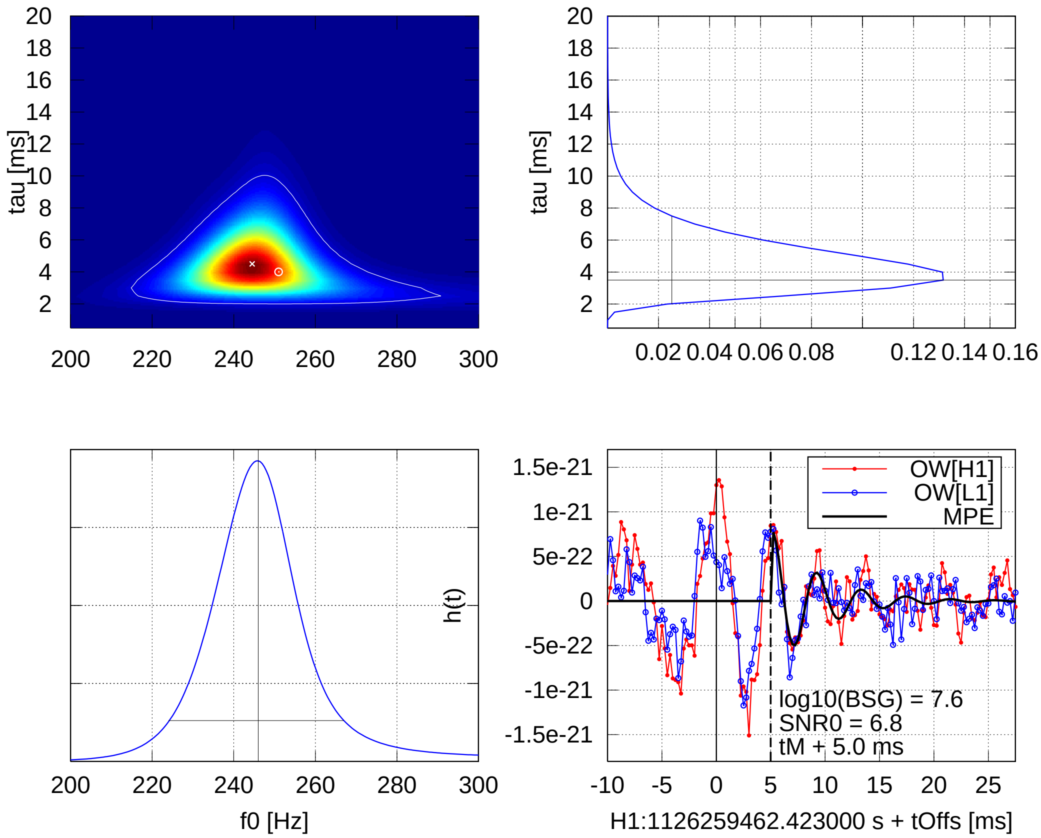
We search the $\{f_0, \tau\}$ range with uniform priors in $f_0 \in [200, 300]\,\mathrm{Hz}$ and $\tau \in [0.5, 20]\,\mathrm{ms}$, in steps of $df_0 = 0.5\,\mathrm{Hz}$ and $d\tau = 0.5\,\mathrm{ms}$, respectively. The following plots show snapshots of the posterior at different fixed start-times t_0. The offset from merger assumes a merger time $t_\mathrm{M} = 1126259462.423$ (in H1 arrival time), as taken from Ian's wiki and rounded to ms accuracy. We show results for QNM start-times $t_0 - t_\mathrm{M} \in \{1, 2, 3, 4, 5, 6, 6.5, 7\}\,\mathrm{ms}$ (referring to H1 arrival times).

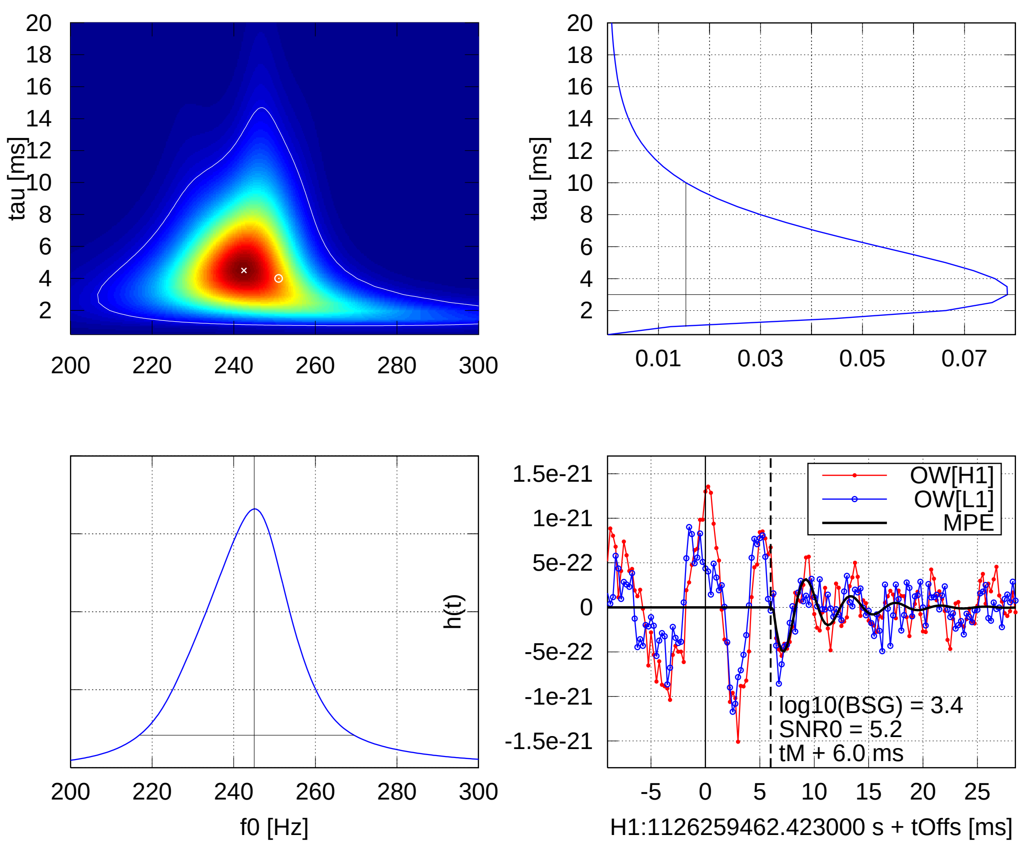

tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 29
SNR0 = 10
tM + 1.0 ms
H1:1126259462.423000 s + tOffs [ms]

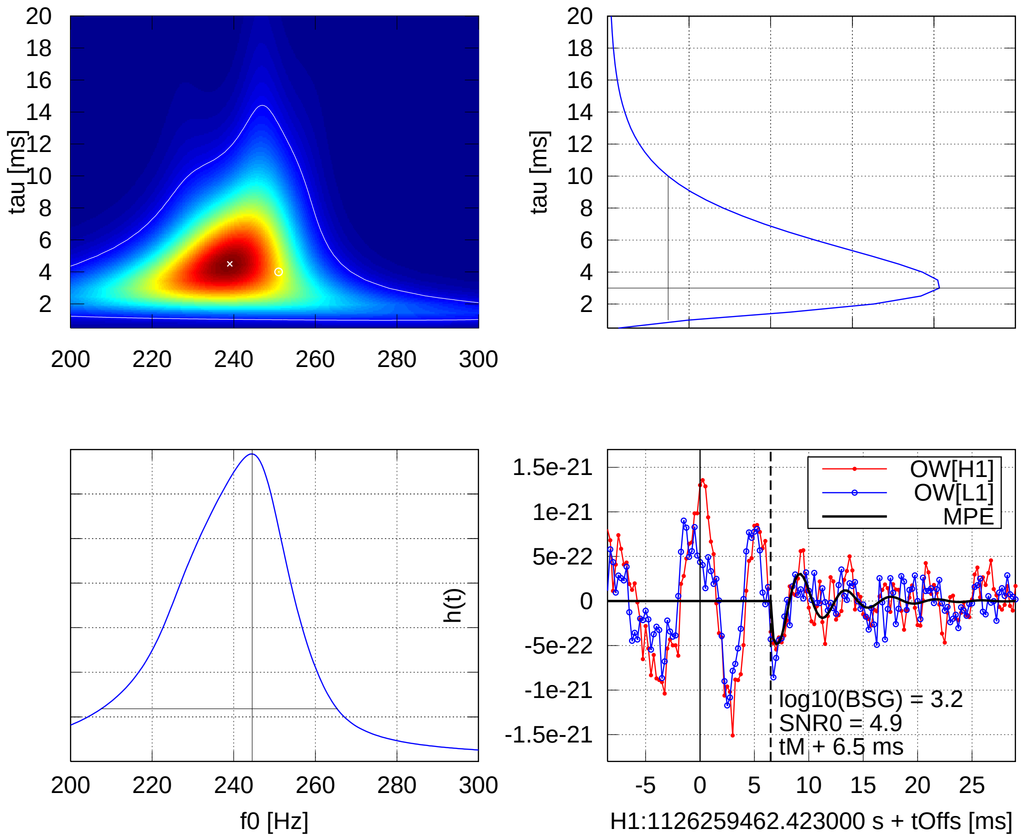

tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 30
SNR0 = 8.8
tM + 2.0 ms
H1:1126259462.423000 s + tOffs [ms]


tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 17
SNR0 = 8.8
tM + 3.0 ms
H1:1126259462.423000 s + tOffs [ms]

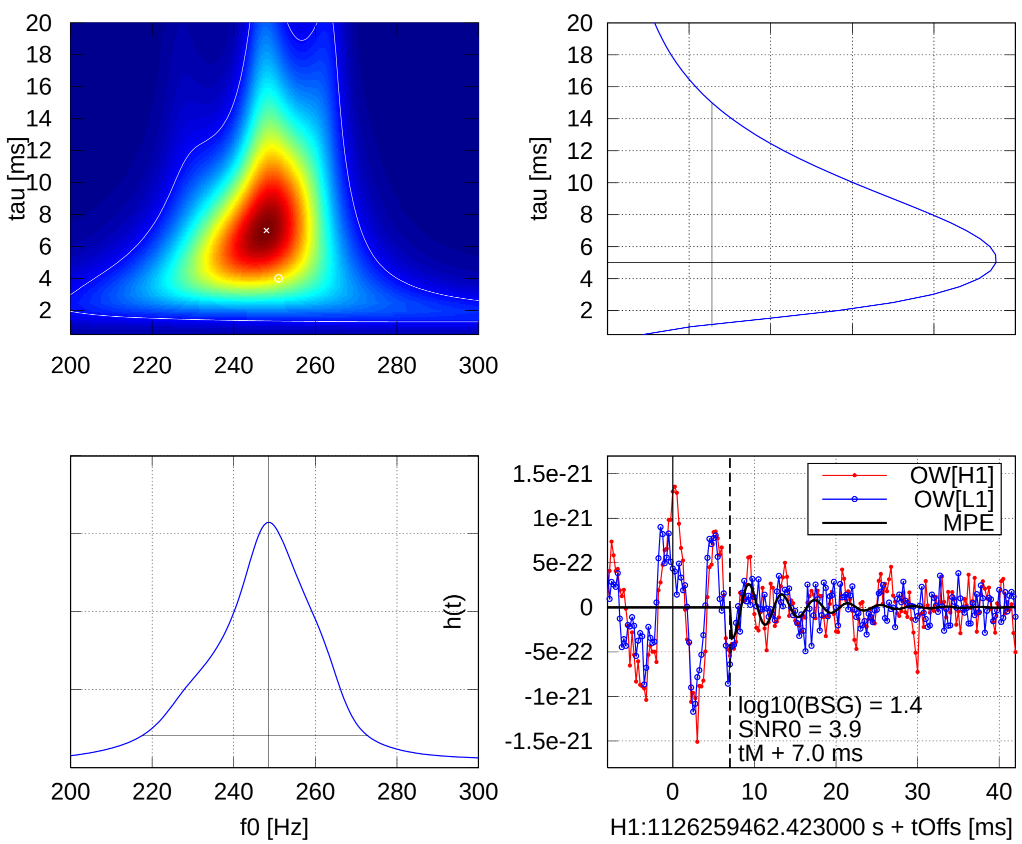

tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 12
SNR0 = 8.3
tM + 4.0 ms
H1:1126259462.423000 s + tOffs [ms]


tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 7.6
SNR0 = 6.8
tM + 5.0 ms
H1:1126259462.423000 s + tOffs [ms]


tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 3.4
SNR0 = 5.2
tM + 6.0 ms
H1:1126259462.423000 s + tOffs [ms]

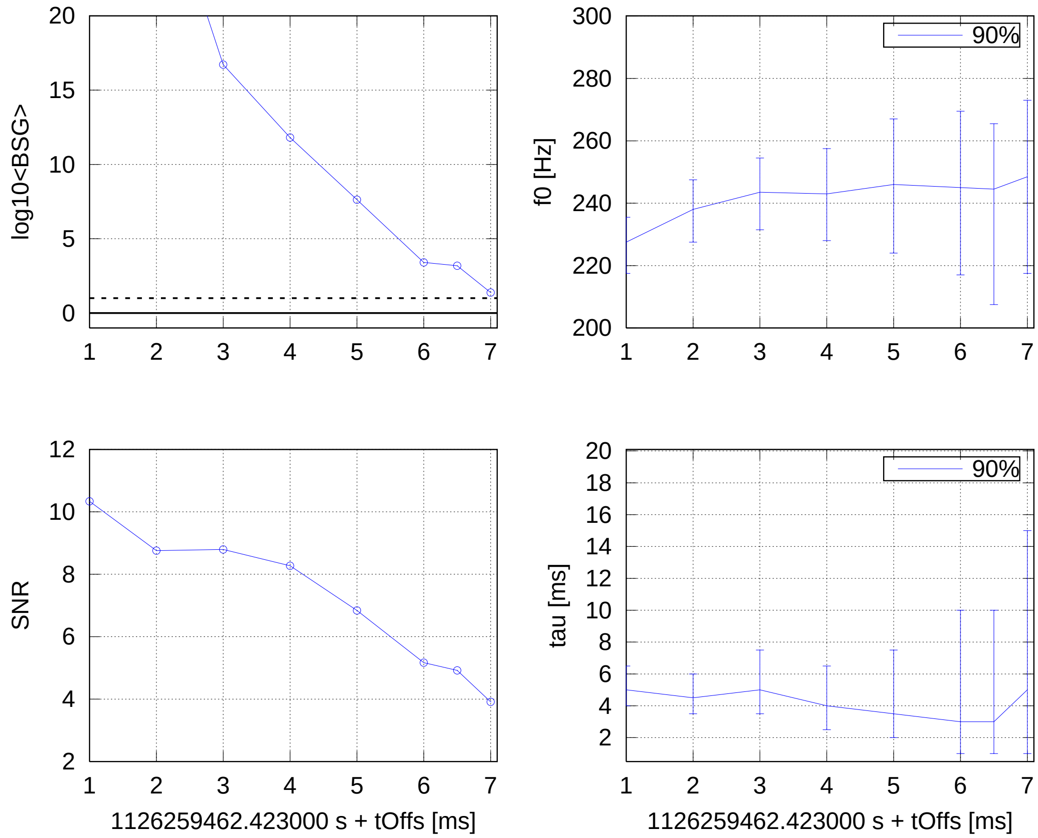


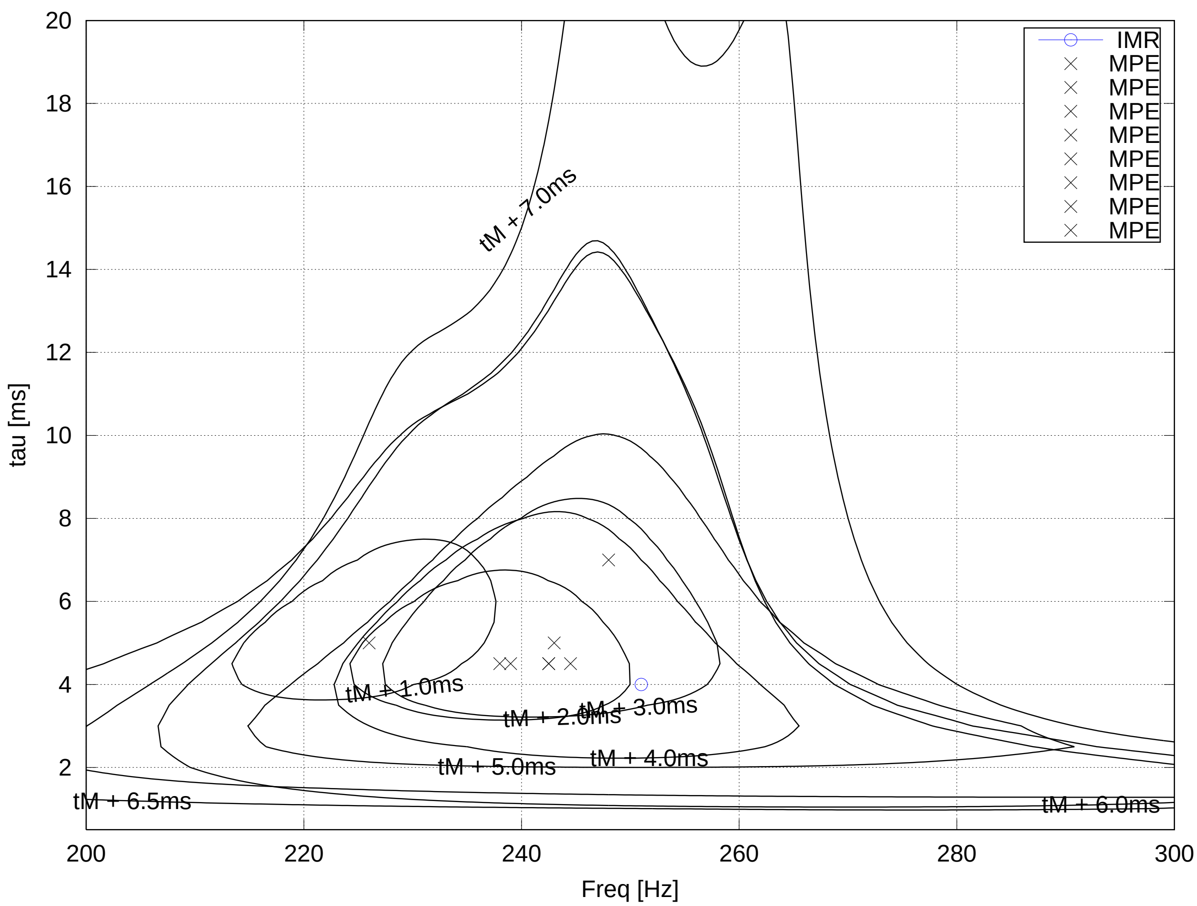
tau [ms]
tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 3.2
SNR0 = 4.9
tM + 6.5 ms
H1:1126259462.423000 s + tOffs [ms]



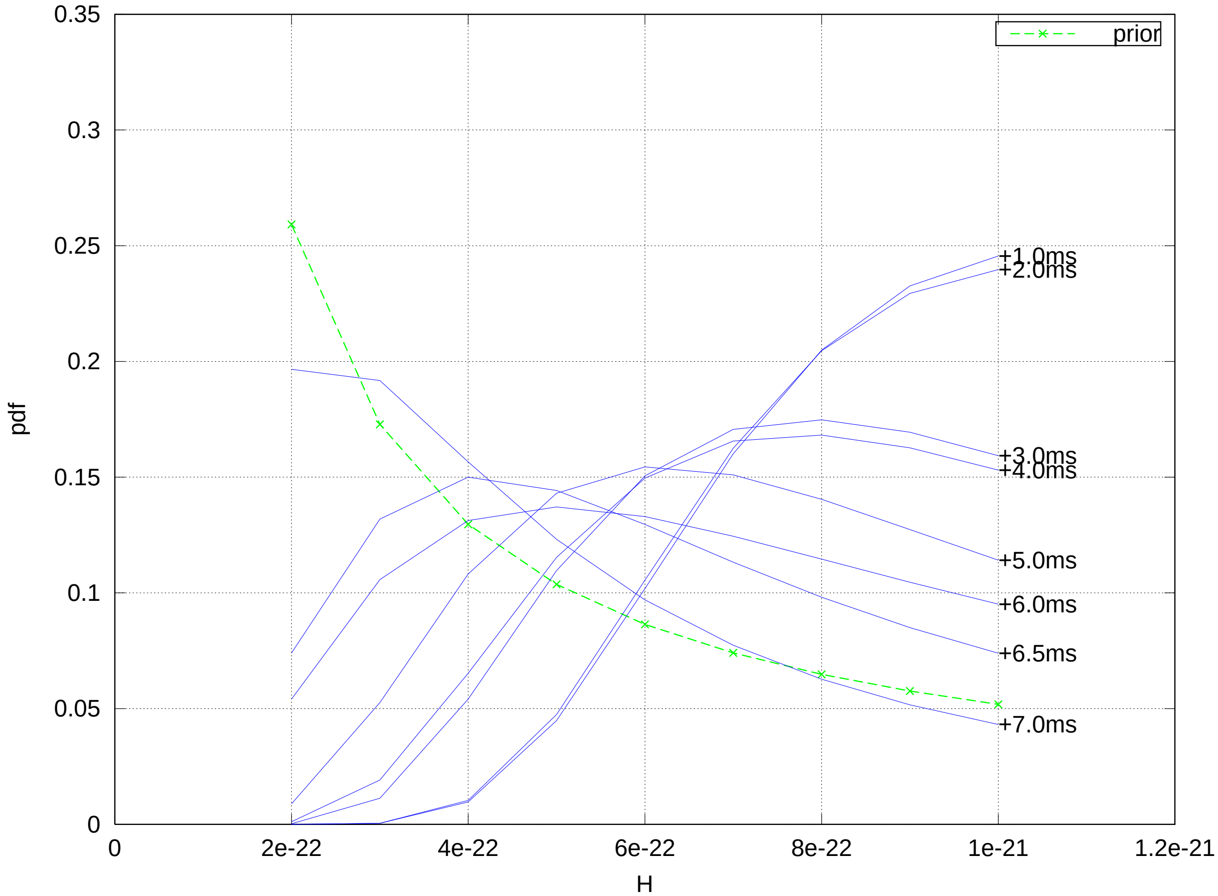
tau [ms]
tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 1.4
SNR0 = 3.9
tM + 7.0 ms
H1:1126259462.423000 s + tOffs [ms]

D. Summary plots





E. H-scale posteriors

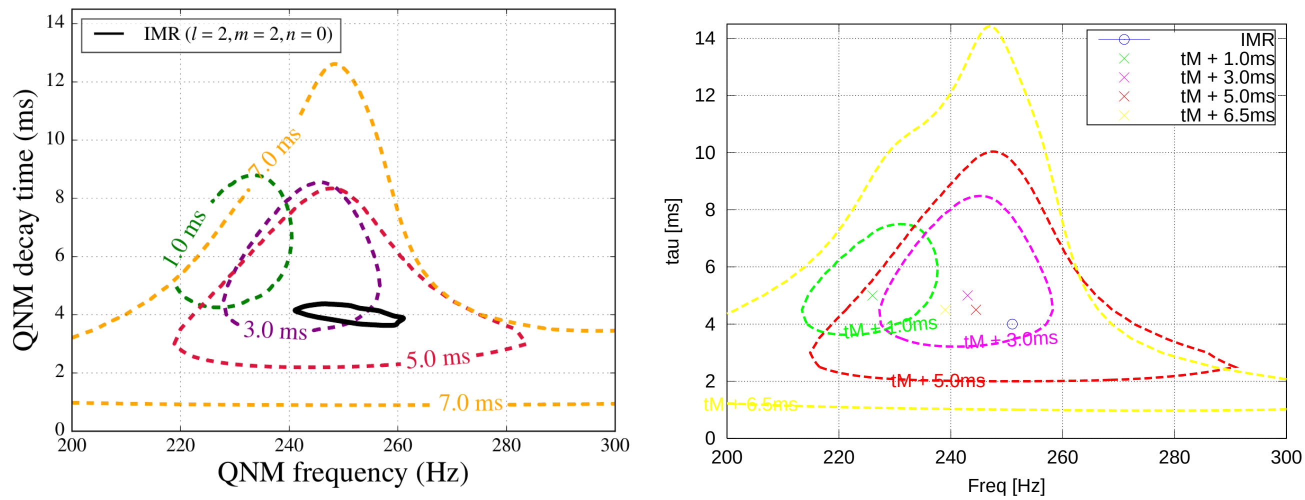


F. Proposed updated statements for PRL "Testing GR" paper

Qualitatively these results agree with what was stated and presented in the submitted PRL version. However the numbers *are* slightly different. Because we consider this version more "mature", more fully tested and characterized (and the previous results did contain a few minor bugs and inconsistencies that have been fixed), we propose to update the text and results when re-submitting the paper to PRL in the following way:

> The 90% posterior contour starts to overlap with GR prediction from the IMR waveform for $t_0 = t_\mathrm{M} + 3\,\mathrm{ms}$, or $\sim 10\,M$ after merger. The corresponding log Bayes factor at this point is $\log_{10} B \sim 17$ and the MAP waveform SNR is ~ 9. For $t_0 = t_\mathrm{M} + 5\,\mathrm{ms}$ the MAP parameters fall within the contour predicted in GR for the least-damped QNM, with $\log_{10} B \sim$ ~~9~~→8 and SNR ~ 7. At $t_0 = t_\mathrm{M} +$ ~~7~~→6.5 ms, or about $\sim 20\,M$ after merger, the posterior uncertainty becomes quite large, and the Bayes factor drops to $\log_{10} B \sim$ ~~2.6~~→3 with SNR $\sim$ ~~4.4~~→5. The signal becomes undetectable shortly thereafter, for $t0 \geq t_\mathrm{M} +$ ~~8~~→9 ms, where $B \sim 1$.

and new posterior-contour plot for these selected time-steps (new plot on the right, current plot on the left):



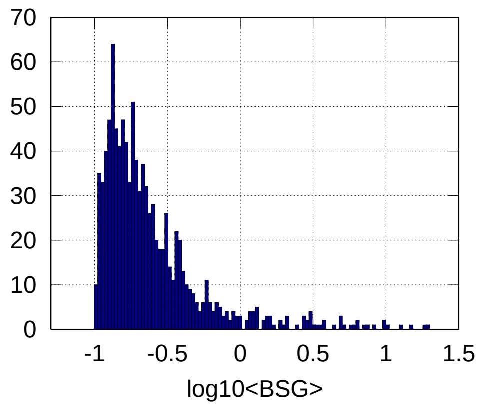
V. TESTING AND CHARACTERIZING PIPELINE PERFORMANCE

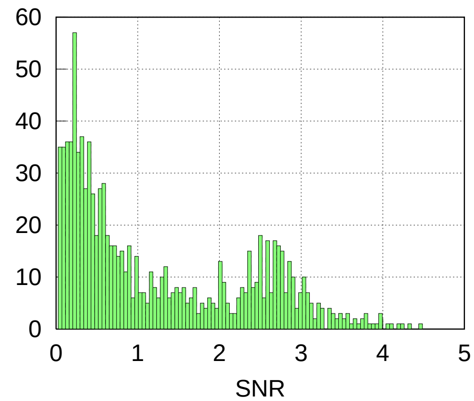
We perform two types of tests to characterize the performance of the search pipeline and parameter estimation (PE). First, we test the distribution of Bayes factors, SNR and PE-intervals on "off-source" pure noise data, in order to assess the significance of the "on source" values found for GW150914. Second, we perform searches on QNM-injections with known parameters and quantify the accuracy of parameter-estimation and of posterior coverage.

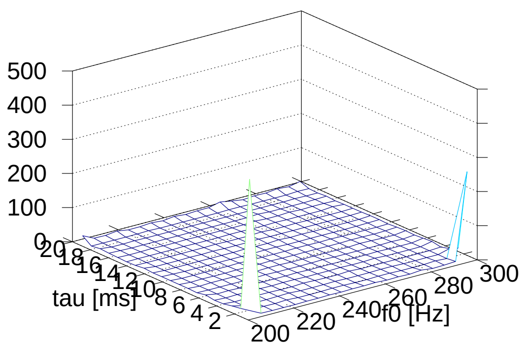
A. Off-source searches

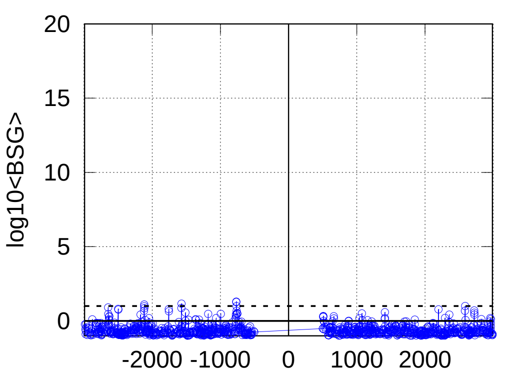
We test off-source search performance on Gaussian white noise, as well as on real detector data around GW150914, using 1000 random start times $t_0 \in ([-3, -0.5]\,\mathrm{s} \cup [0.5, 3]\,\mathrm{s}) + t_\mathrm{M}$.

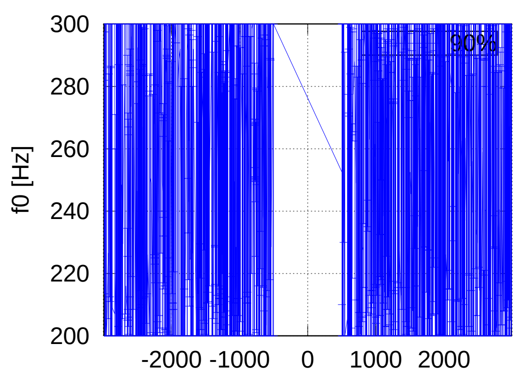
1. Gaussian white noise with PSD $\sqrt{S_X} = 8 \times 10^{-24}/\sqrt{\mathrm{Hz}}$

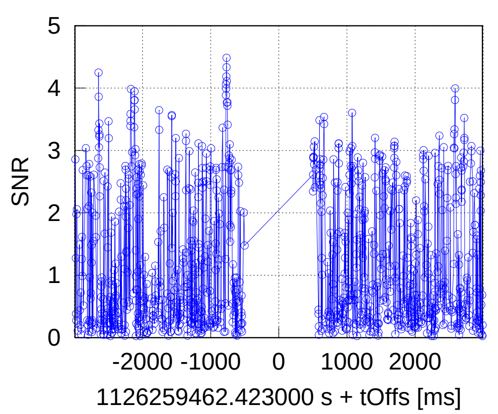


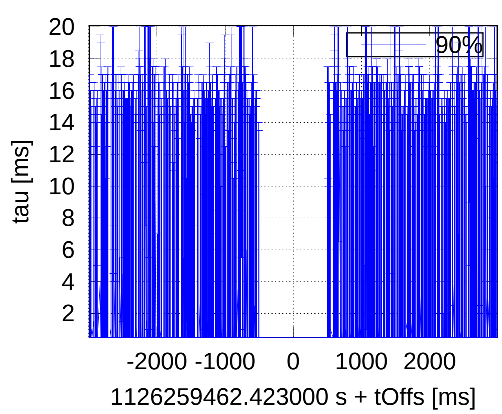


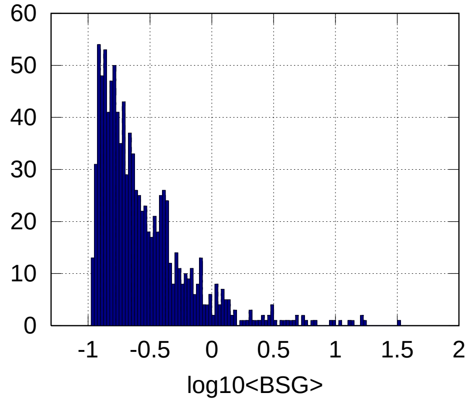


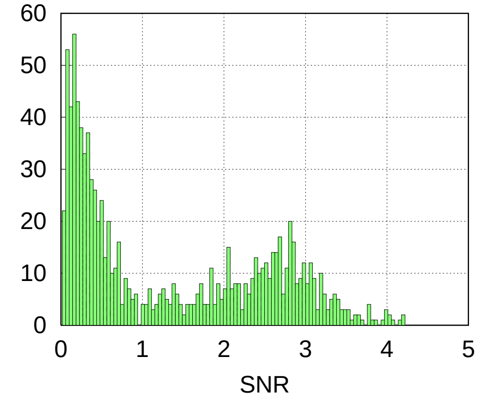


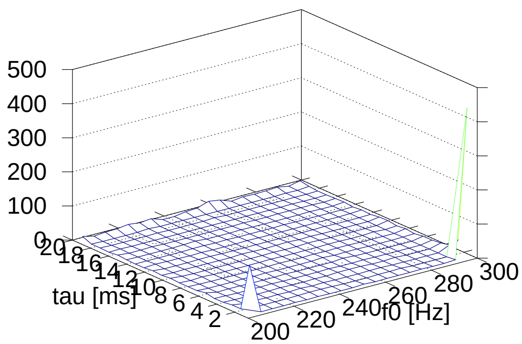


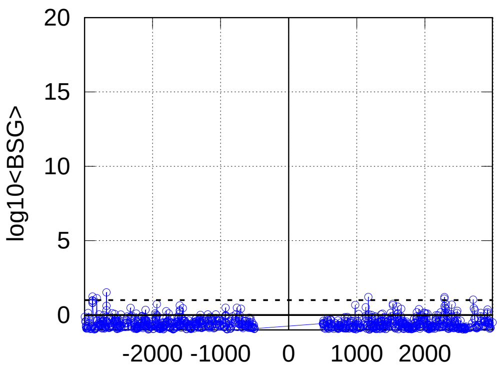


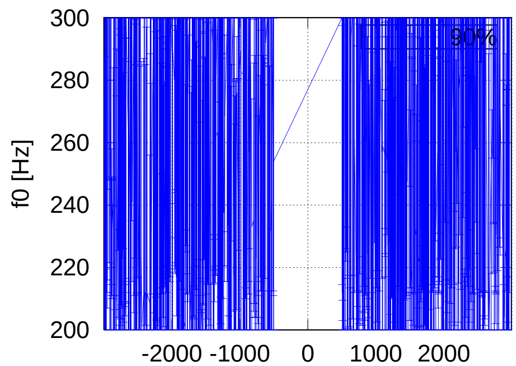


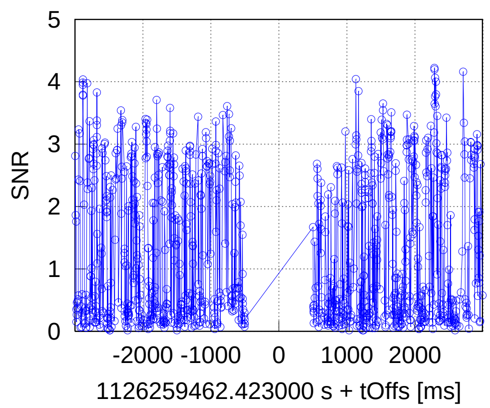

log10<BSG>

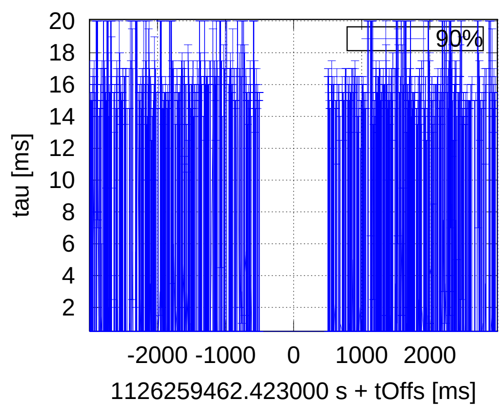

SNR


tau [ms]
f0 [Hz]


log10<BSG>


f0 [Hz]
90%


SNR
1126259462.423000 s + tOffs [ms]


tau [ms]
90%
1126259462.423000 s + tOffs [ms]

B. Testing parameter-estimation accuracy on injections

We test off-source search and PE performance on injected QNM signals with perfectly-matched start-time t_0, using 1000 random start times $t_0 \in ([-3, -0.5]\,\mathrm{s} \cup [0.5, 3]\,\mathrm{s}) + t_\mathrm{M}$. The QNM parameters are drawn from the following *uniform* distributions

$$A \in [3, 8] \times 10^{-22}\,, \tag{60}$$
$$\phi_0 \in [0, 2\pi]\,, \tag{61}$$
$$f_0 \in [200, 300]\,\mathrm{Hz}\,, \tag{62}$$
$$\tau \in [0.5, 20]\,\mathrm{ms}\,, \tag{63}$$

which correspond to the assumed prior distributions for these parameters, except for the amplitude A, which has a more complicated (non-flat) prior, as shown in Fig. 1. In order to avoid injecting a discontinuity at t_0 (which would create problems when band-passing and FFTing the data), we include a smooth "ringup" with characteristic timescale $-\tau/10$ in the injections, i.e. $s(t<t_0) = A\, e^{10\frac{t-t_0}{\tau}} \cos\left(2\pi f_0 (t-t_0) + \phi_0\right)$ instead of $s(t<t_0) = 0$ as used in the template.

–¿ plot figure

We consider 4 injection scenarios:

1. Pure signals without noise, assuming a white noise-floor of $\sqrt{S_X} = 8 \times 10^{-24}/\sqrt{\mathrm{Hz}}$.

   These "signal-only" injections focus on the accuracy of the MPE estimate versus the true injected parameters, which should essentially be perfect at sufficiently large SNR (wrt the assumed noise-floor). For "lower SNR" signals, the amplitude prior starts to affect parameter estimation and leads to deviations from the injected signal parameters, which is perfectly normal and expected.

2. Signals injected in Gaussian white noise of $\sqrt{S_X} = 8 \times 10^{-24}/\sqrt{\mathrm{Hz}}$, assuming this exact noise PSD.

   The injections into *known* Gaussian noise exactly realize all of the assumptions, and as we're drawing signals essentially from the priors, we expect the parameter posteriors to accurately predict their frequency of coverage [e.g. see 4].

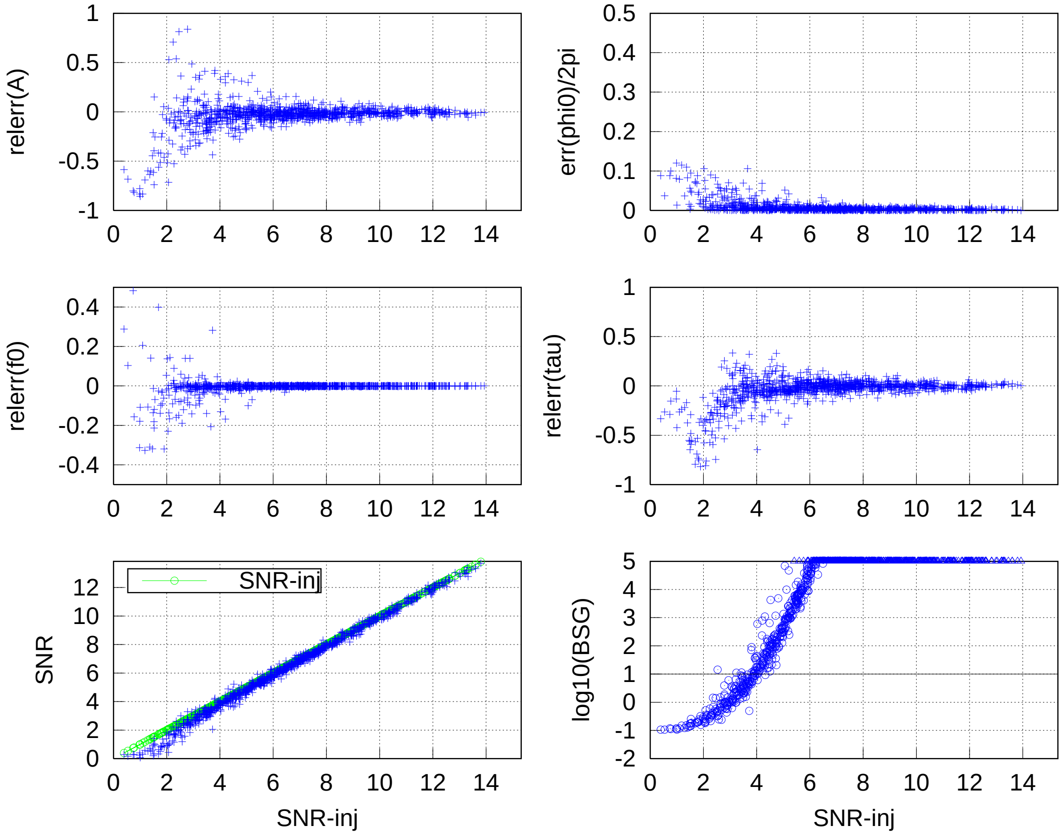
3. Signals injected in Gaussian white noise of $\sqrt{S_X} = 8 \times 10^{-24}/\sqrt{\mathrm{Hz}}$, estimating PSD from the data.

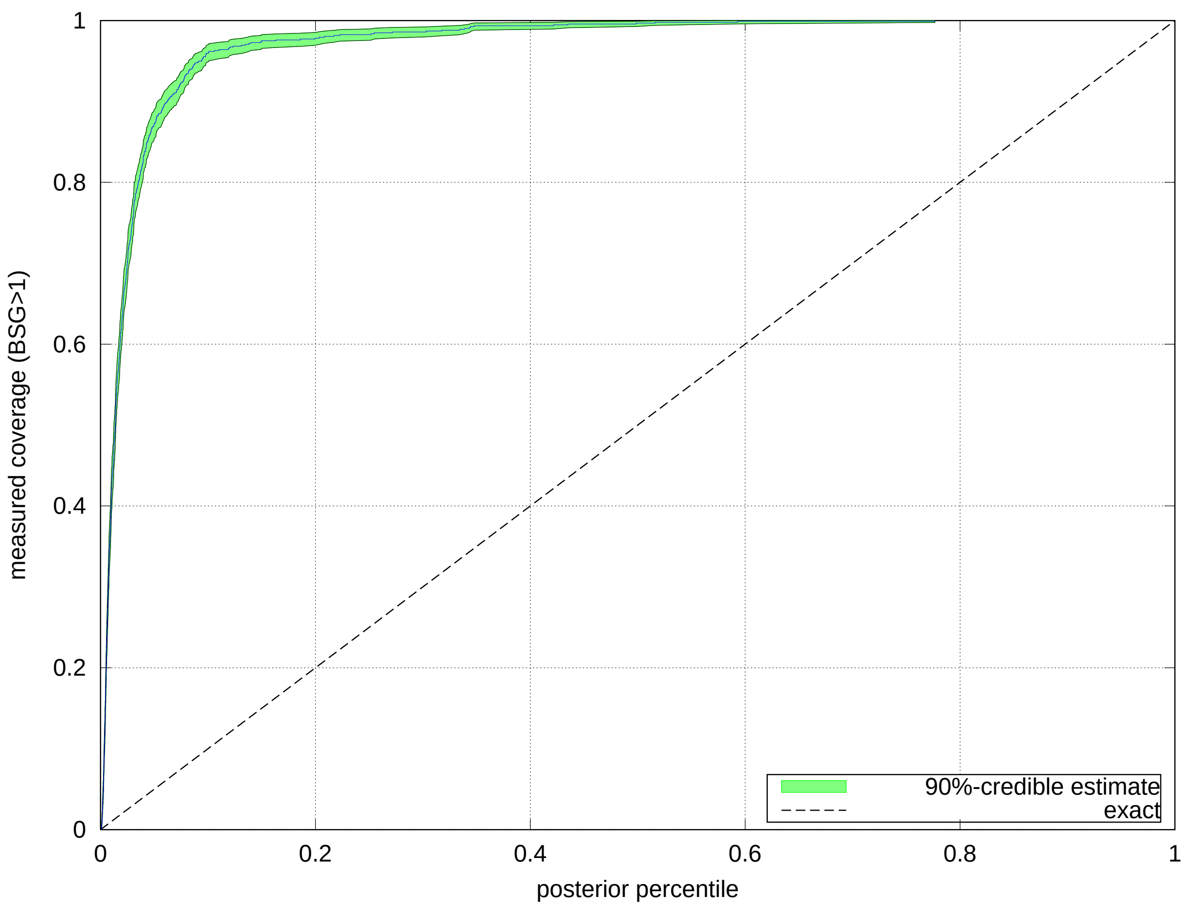
   Compared to the previous case this basically just tests the reliability of noise-estimation, and we're still expecting posteriors to predict their coverage.

4. Signals injected into real off-source detector data.

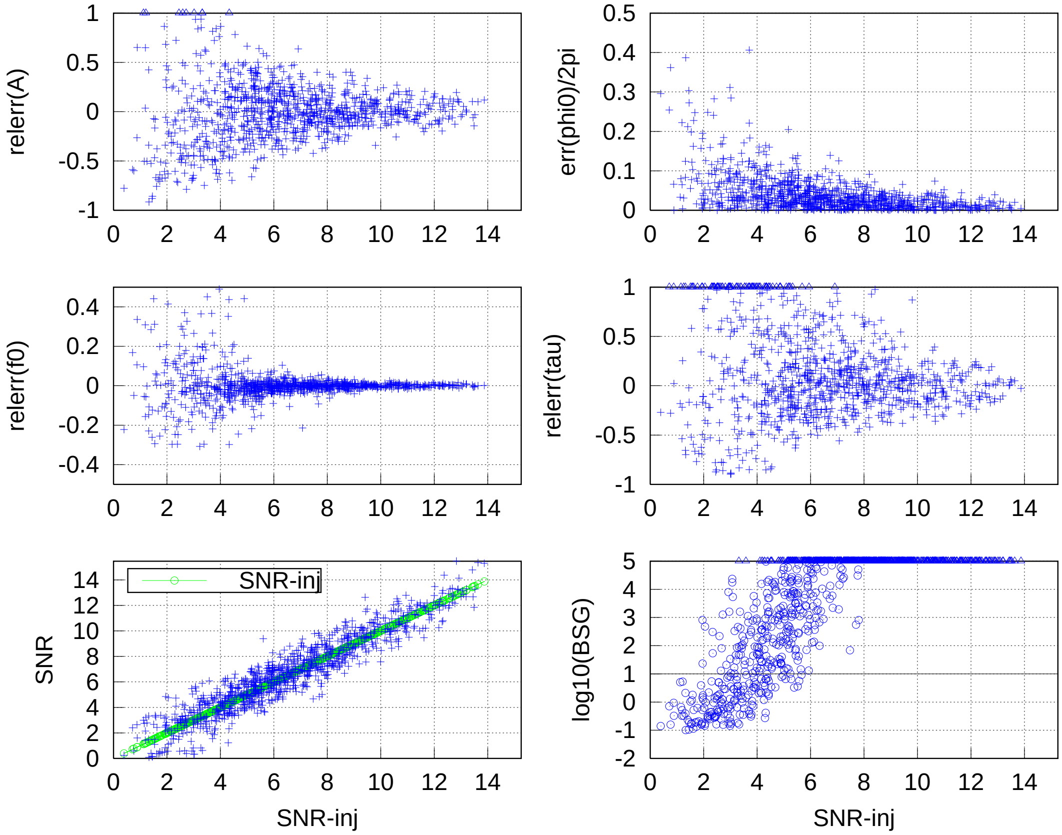
   Here we can quantify how much of an effect "real noise" instead of Gaussian noise has on the coverage of the posteriors, which can lead to over- or under-coverage.

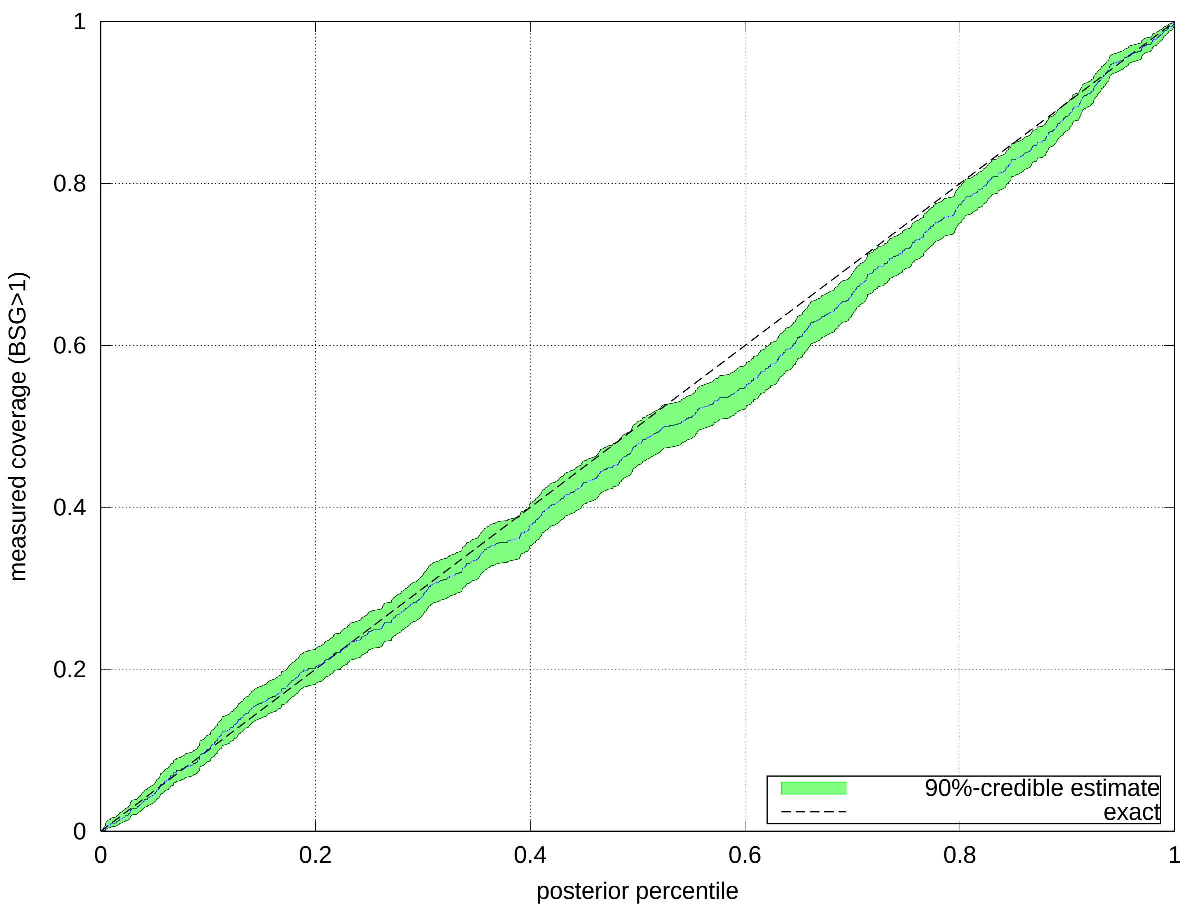
Cases 1-3 essentially serve to test self-consistency of the method and implementation, given that we know exactly what to expect, i.e. in order to catch mistakes or bugs. Case 4 serves to test the validity of our assumptions (mostly about the noise) in real detector data.


relerr(A)
err(phi0)/2pi
relerr(f0)
relerr(tau)
SNR
SNR-inj
log10(BSG)
SNR-inj
SNR-inj

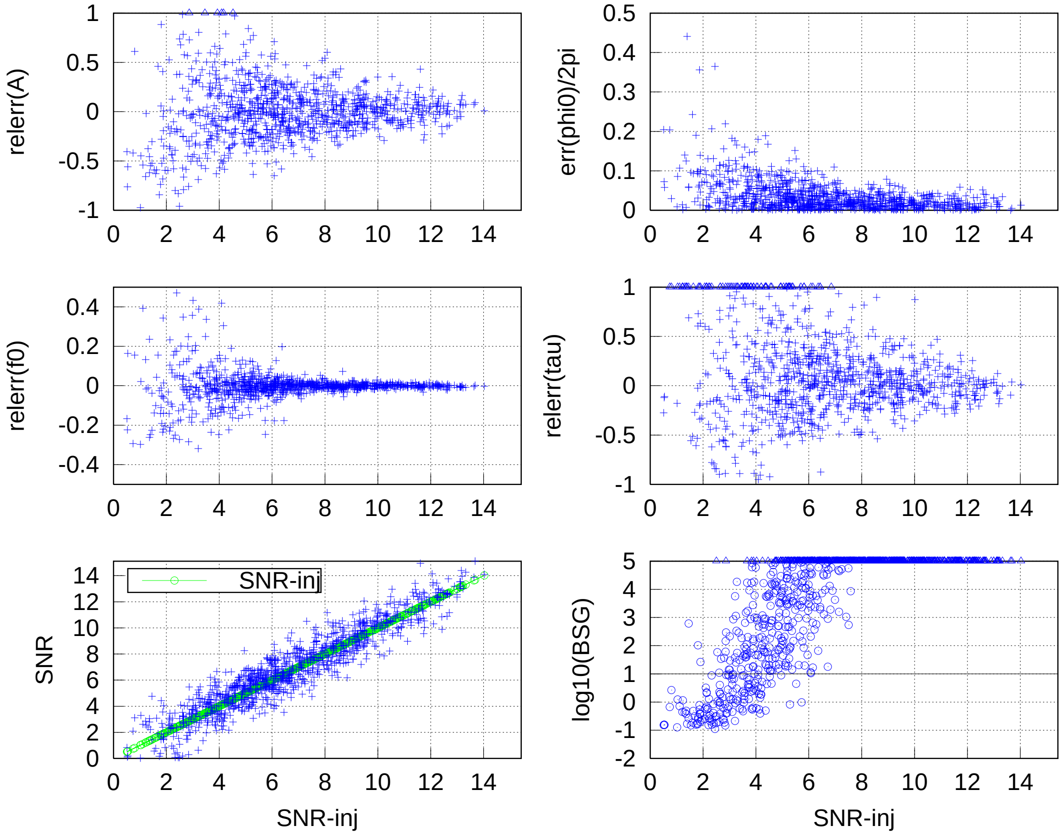

measured coverage (BSG>1)
posterior percentile
90%-credible estimate
exact

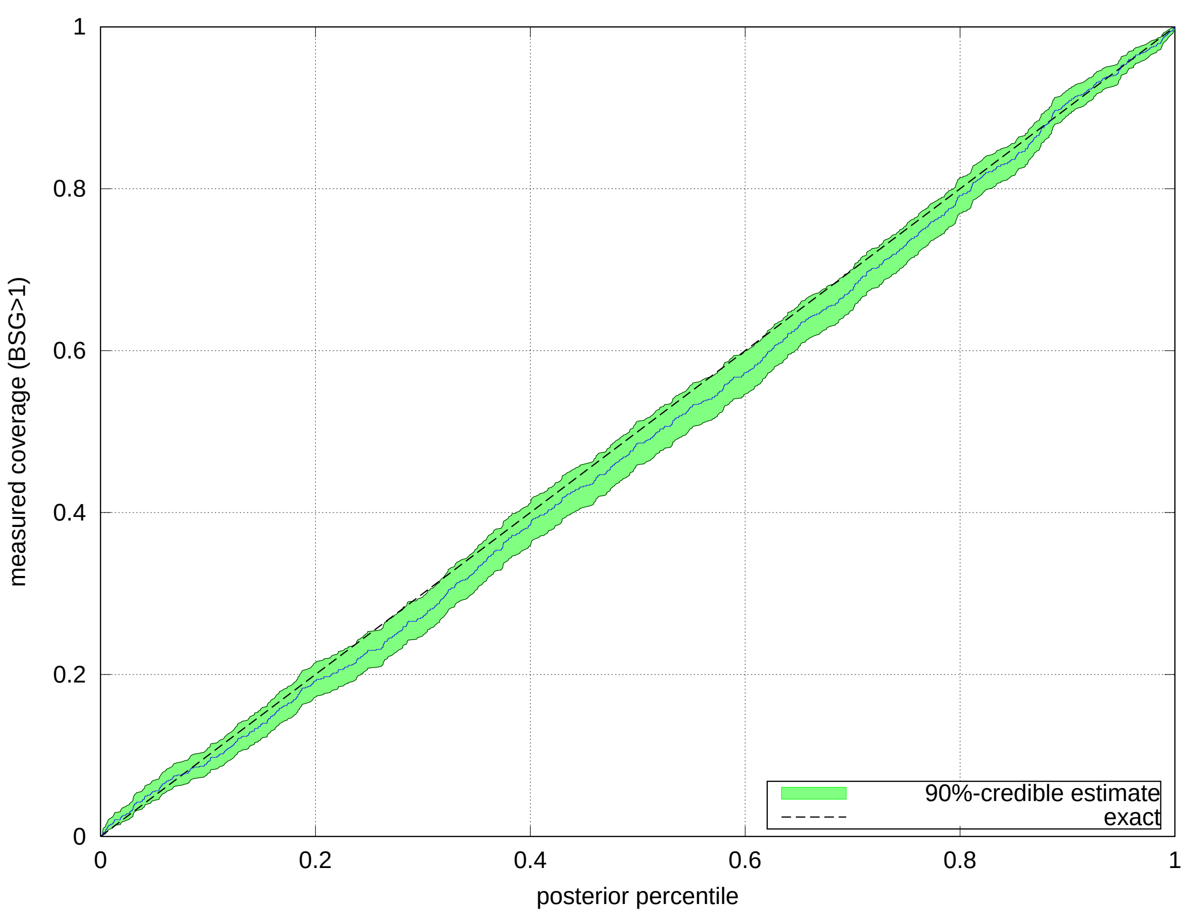
2. *Case 2: Injections into **known** Gaussian noise of* $\sqrt{S_X} = 8 \times 10^{-24}/\sqrt{\mathrm{Hz}}$

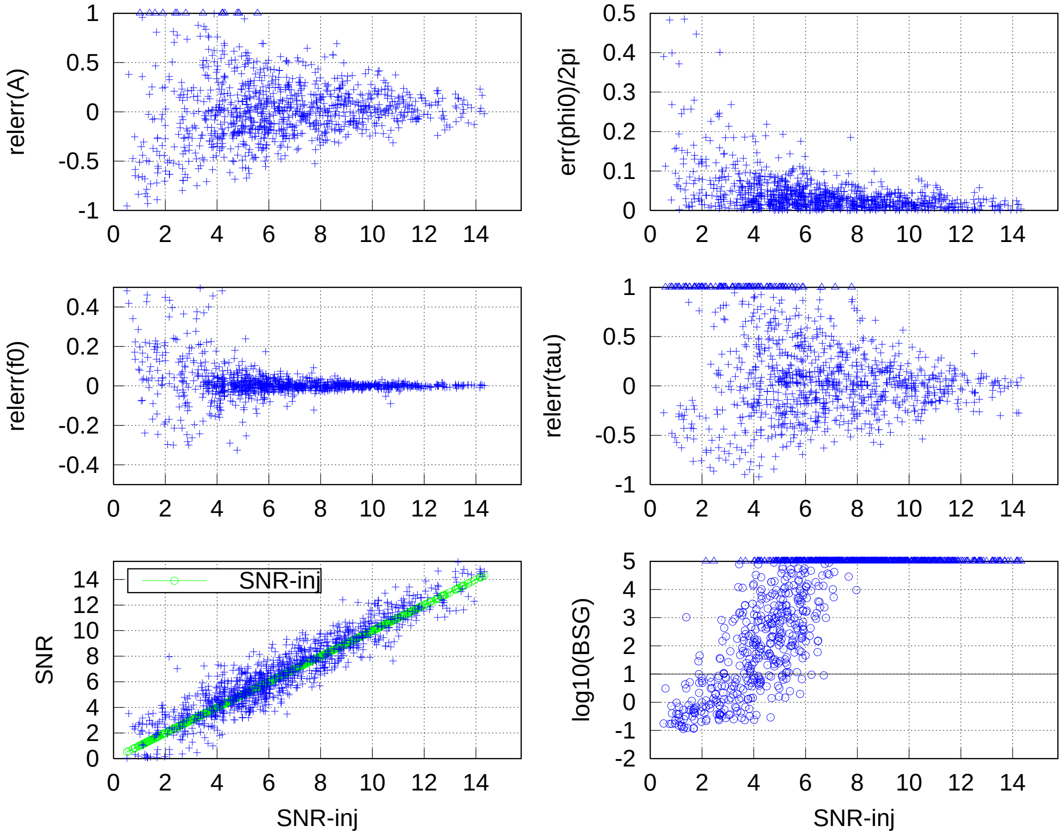


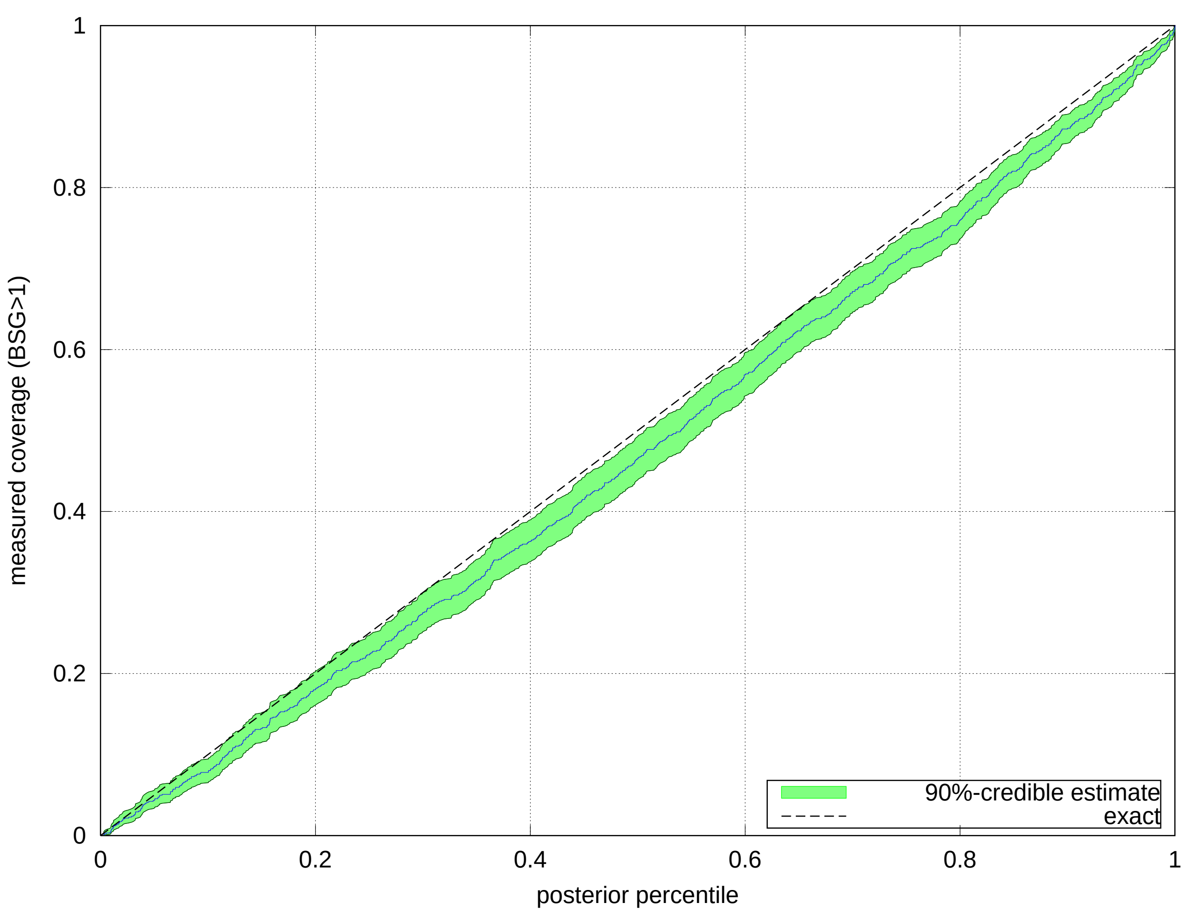


*3. Case 3: Injections into **unknown** Gaussian noise of* $\sqrt{S_X} = 8 \times 10^{-24}/\sqrt{\mathrm{Hz}}$






relerr(A)
err(phi0)/2pi
relerr(f0)
relerr(tau)
SNR
SNR-inj
log10(BSG)
SNR-inj
SNR-inj


measured coverage (BSG>1)
posterior percentile
90%-credible estimate
exact

Appendix A: Deprecated old way to compute $\langle s|s\rangle$: const noise floor + time-domain integration

Given that aLIGO noise-curve is relatively "white" over a broad-band in the "bucket", and the signal $s(t)$ of Eq. (1) can still be considered relatively "narrow band" ($\sim \pm 100$Hz) with respect to this noise curve, we can approximate the signal-normalization integral as

$$
\begin{aligned}
\langle s|s\rangle &= \sum_X 2\int_{-\infty}^{\infty} \frac{\widetilde{s}^X(f)\,\widetilde{s}^{*X}(f)}{S_X(f)}\,df && \text{(A1)}\\
&\sim \sum_X \frac{2}{S_X(f')}\int_0^{\infty} s^2(t;\theta)\,dt && \text{(A2)}\\
&= \frac{2\,N_{\rm det}}{\mathcal{S}(f')}\int \left(\mathcal{A}_{\rm s}^2\,h_{\rm s}^2(t) + 2\mathcal{A}_{\rm s}\mathcal{A}_{\rm c}\,h_{\rm s}h_{\rm c} + \mathcal{A}_{\rm c}^2\,h_{\rm c}^2\right)\,dt && \text{(A3)}\\
&= \vec{\mathcal{A}}\cdot\mathcal{M}\cdot\vec{\mathcal{A}}, && \text{(A4)}
\end{aligned}
$$

with

$$
\mathcal{M} \equiv 2N_{\rm det}\begin{pmatrix} I_{\rm s} & I_{\rm sc}\\ I_{\rm sc} & I_{\rm c}\end{pmatrix} \tag{A5}
$$

with

$$
\begin{aligned}
I_{\rm s} &\equiv \frac{1}{\mathcal{S}(f')}\int_0^{\infty} e^{-\frac{2t}{\tau}}\,\sin^2(2\pi f t)\,dt = \frac{1}{2\pi f}\int e^{-\frac{\varphi}{Q}}\sin^2\varphi\,d\varphi && \text{(A6)}\\
I_{\rm c} &\equiv \int_0^{\infty} e^{-\frac{2t}{\tau}}\,\cos^2(2\pi f t)\,dt = \frac{1}{2\pi f}\int e^{-\frac{\varphi}{Q}}\cos^2\varphi\,d\varphi && \text{(A7)}\\
I_{\rm sc} &\equiv \int_0^{\infty} e^{-\frac{2t}{\tau}}\,\sin(2\pi f t)\cos(2\pi f t)\,dt = \frac{1}{4\pi f}\int e^{-\frac{\varphi}{Q}}\sin 2\varphi\,d\varphi\,, && \text{(A8)}
\end{aligned}
$$

where f' is some (unknown) frequency within the effective frequency band around the central signal frequency f (using mean-value theorem), and we have used the assumption of identical signal model in both detectors (after time-shifting the data and correcting for antenna-pattern differences).

The respective integrals to compute are

$$
\tag{A9}
$$

using the definitions

$$
\begin{aligned}
\varphi &\equiv 2\pi f\Delta t\,, && \text{(A10)}\\
Q &\equiv \pi f\tau\,. && \text{(A11)}
\end{aligned}
$$

Assuming only non-critically damped signals, i.e. $Q \gtrsim \mathcal{O}(\pi)$, these integrals can be approximated computed analytically as

$$
\begin{aligned}
I_{\rm s}' &= \frac{-1}{2\pi f}\frac{Q^2}{1+4Q^2}e^{-\frac{\varphi}{Q}}\left[\sin 2\varphi + 2Q + \frac{\sin^2\varphi}{Q}\right]\Bigg|_0^{\infty} = \frac{2Q}{2\pi f}\,\frac{Q^2}{1+4Q^2} = \frac{\tau}{4+Q^{-2}} && \text{(A12)}\\
&\overset{Q\gg 1}{\approx} \frac{\tau}{4}\,, && \text{(A13)}\\
I_{\rm c}' &= \frac{-1}{2\pi f}\frac{Q^2}{1+4Q^2}e^{-\frac{\varphi}{Q}}\left[-\sin 2\varphi + 2Q + \frac{\cos^2\varphi}{Q}\right]\Bigg|_0^{\infty} = \frac{2Q+\frac{1}{Q}}{2\pi f}\,\frac{Q^2}{1+4Q^2} = \frac{\tau}{4}\left(\frac{2+Q^{-2}}{2+Q^{-2}/2}\right) && \text{(A14)}\\
&\overset{Q\gg 1}{\approx} \frac{\tau}{4}\,, && \text{(A15)}\\
I_{\rm sc}' &= \frac{-1}{2\pi f}\frac{Q^2}{1+4Q^2}e^{-\frac{\varphi}{Q}}\left[2\cos^2\varphi - 1 + \frac{\sin 2\varphi}{2Q}\right]\Bigg|_0^{\infty} = \frac{1}{2\pi f}\,\frac{Q^2}{1+4Q^2} = \frac{\tau}{8Q\,(1+Q^{-2}/4)} && \text{(A16)}\\
&\overset{Q\gg 1}{\approx} \frac{1}{2Q}\,I_{\rm s} \ll I_{\rm s} \approx 0\,. && \text{(A17)}
\end{aligned}
$$

So $I_\mathrm{s} \approx I_\mathrm{c} \approx \frac{N_\mathrm{det}\tau}{2\mathcal{S}(f')}$ and $I_\mathrm{sc} \approx 0$, and we obtain the approximate $\mathcal{M}$-matrix as

$$\mathcal{M} \approx \frac{N_\mathrm{det}\,\tau}{2\mathcal{S}(f')}\,\mathbb{I} = \begin{pmatrix} I_0 & 0 \\ 0 & I_0 \end{pmatrix}\,. \tag{A18}$$

Note: in the QNM search we'll approximate $\mathcal{S}(f')$ in this expression by the arithmetic mean $\langle \mathcal{S}(f) \rangle_{f \pm \Delta f}$ around each template frequency f. This fixed-SN high-Q limit was originally used in the v1 of this search and document, which was originally circulated.

[1] G. L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation*, Lecture notes in statistics (Springer-Verlag, New York, 1988).

[2] P. Jaranowski, A. Królak, and B. F. Schutz, Phys. Rev. D **58**, 063001 (1998).

[3] E. T. Jaynes, *Probability Theory. The Logic of Science* (Cambridge University Press, 2003).

[4] A. C. Searle, ArXiv e-prints (2008), 0804.1161.