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: 2016-04-27 14:34:49 +0200; commitID: e23c0ae-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: 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: 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 $\langle x|s\rangle$
- bug/inconsistency (up to version 'v4-'): I had used a tukey(0.1) window on SFT data (spanning $[10, 2000]$Hz) before inv-FFTing into time-domain, and then used data starting from 30Hz. At these low frequencies, however, the data was still affected by the Tukey-window (clearly visible in the previous PSD estimates). I've now reduced this to tukey(0.02), which is enough to avoid creating spurious time-domain noise, and even allowed me to push the low-frequency end down to 20 Hz.
- expanded used data-range from previously $[30, 1000]$ Hz to $[20, 1900]$ Hz (which avoids the lower and upper boundaries near 10Hz and 2000Hz affected by tukey(0.02)-windowing).
- apply tukey(0.1)-windowing to the $T = 8$ s time-series data before FFTing for whitening: this avoids red-noise artifacts, especially visible in L1 whitened spectra in previous versions (up to 'v4-'). Contrary to the frequency domain, none of the data near the start and end is actually used, so a wider window is used.
- PSD estimation: avoid using the central data-segment of $T = 8$ 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 estimate of Eq. (60) (rather than H-MP approximation of Eq. (64)) for amplitude parameters and resulting SNR. Essentially found to make no relevant difference.
- 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
- Finer τ resolution $d\tau = 0.2$ ms instead of 0.5 ms, but coarser f_0 resolution of $df_0 = 1$ Hz instead of 0.5 Hz, which is closer to actual parameter-space resolution as seen from posterior shapes.

All results shown in this version of the notes were produced using the ringdown pipeline in gitLab, version 7521757

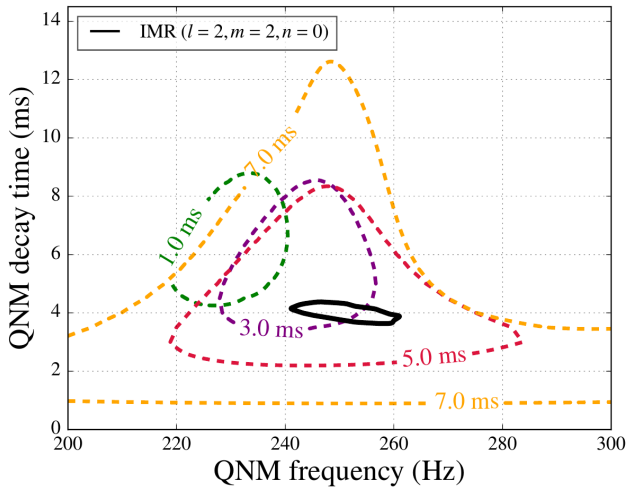
[a] Reinhard.Prix@ligo.org

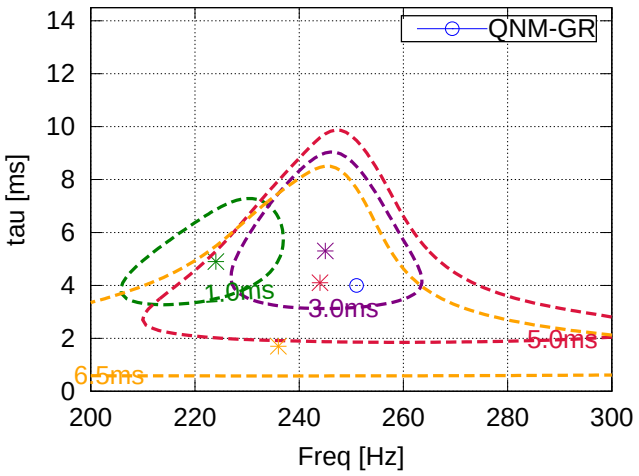
B. Proposed updated statements for PRL "Testing GR" paper

Qualitatively these results still agree with what has been shown in the first submitted PRL version. However the numbers and posterior contours are somewhat different, and this version is much more mature/converged, more fully tested and characterized, and a number of bugs and possible inconsistencies have been fixed since v3. We therefore propose to update the text and results when re-submitting the paper to PRL in the following way:

> We use 1800s of detector data containing GW150914 from H1 and L1, band-passed to $[20, 1900]$ Hz. We assume the signal arrived 7 ms earlier in L1 compared to H1, and that the amplitude has opposite sign in both detectors. [ref detection paper?]
> ...
> The 90% posterior contour overlaps the GR prediction from the IMR waveform starting at about $t_0 \gtrsim t_{\rm M} + 3\,{\rm ms}$, or $\sim 10\,M$ after merger. The corresponding Bayes factor at this point is $\log_{10} B \sim 14$ and the MAP waveform SNR is ~ 8. At $t_0 = t_{\rm M} + 5\,{\rm ms}$ the MAP parameters fall within the contour predicted in GR for the least-damped QNM, with $\log_{10} B \sim 6.5$ and SNR ~ 6.3. At $t_0 = t_{\rm M} + 6.5\,{\rm ms}$, or about $\sim 20\,M$ after merger, the Bayes factor is $\log_{10} B \sim 3.5$ with SNR ~ 4.8. The signal becomes undetectable after about $t_0 \gtrsim t_{\rm M} + 9\,{\rm ms}$, where $B \lesssim 1$.

New posterior contours [the official plot will again be produced by Walter] for selected time-steps (new result on the right, current plot on the left):





I. INTRODUCTION

Signal model: damped sinusoid starting at t_0, writing $\Delta t \equiv t - t_0$:

$$s(t;\, A, \phi_0,\, t_0, \tau, f_0) = \begin{cases} A\, \exp\left(-\frac{\Delta t}{\tau}\right)\, \cos\left(2\pi\, f_0\, \Delta t + \phi_0\right)\,, & \text{if } \Delta t \geq 0\,, \\ 0 & \text{if } \Delta t < 0\,. \end{cases} \tag{1}$$

We re-parametrize $\{A, \phi_0\}$ in terms of two amplitudes, namely

$$\mathcal{A}_\mathrm{s} = -A\, \sin\phi_0\,, \tag{2}$$
$$\mathcal{A}_\mathrm{c} = \;\; A\, \cos\phi_0\,, \tag{3}$$

and it is useful to distinguish "amplitude parameters" $\mathcal{A}$ and "evolution parameters" λ as

$$\mathcal{A} \equiv (\mathcal{A}_\mathrm{s}, \mathcal{A}_\mathrm{c})\,, \quad \lambda \equiv \{t_0, \tau, f_0\}\,. \tag{4}$$

Using this "JKS" factorization [1, 2], the signal model Eq. (1) now reads as:

$$s(t; \mathcal{A}, \lambda) = \mathcal{A}_\mathrm{s}\, h_\mathrm{s}(t;\lambda) + \mathcal{A}_\mathrm{c}\, h_\mathrm{c}(t;\lambda)\,, \tag{5}$$

with *basis functions*

$$h_\mathrm{s}(t;\lambda) \equiv \exp\left(-\Delta t/\tau\right)\, \sin(2\pi\, f_0\, \Delta t)\,, \tag{6}$$
$$h_\mathrm{c}(t;\lambda) \equiv \exp\left(-\Delta t/\tau\right)\, \cos(2\pi\, f_0\, \Delta t)\,. \tag{7}$$

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

$$P\left(x|\mathcal{H}_\mathrm{G}\right) = c\, \exp\left[-\frac{1}{2}\, \langle x|x\rangle\right]\,, \tag{8}$$

$$P\left(x|\mathcal{H}_\mathrm{S}, \mathcal{A}, \lambda\right) = c\, \exp\left[-\frac{1}{2}\, \langle x - s\left(\mathcal{A}, \lambda\right)|x - s\left(\mathcal{A}, \lambda\right)\rangle\right]\,, \tag{9}$$

with the multi-detector scalar product (over detectors indexed by X) defined as

$$\langle x|y\rangle \equiv \sum_X \left\langle x^X \middle| y^X \right\rangle = \sum_X 4\Re \int_0^\infty \frac{\widetilde{x}^X(f)\, \widetilde{y}^{*X}(f)}{S_X(f)}\, df\,, \tag{10}$$

where $S_X(f)$ is the single-sided noise PSD for detector X. The (marginal) likelihood for the signal model can be expressed as

$$P\left(x|\mathcal{H}_\mathrm{S}\right) = \int P\left(x|\mathcal{H}_\mathrm{S}, \mathcal{A}, \lambda\right)\, P\left(\mathcal{A}, \lambda|\mathcal{H}_\mathrm{S}\right)\, d\mathcal{A}\, d\lambda\,, \tag{11}$$

and the corresponding Bayes factor (or marginal likelihood ratio) is

$$B_\mathrm{S/G}(x) \equiv \frac{P\left(x|\mathcal{H}_\mathrm{S}\right)}{P\left(x|\mathcal{H}_\mathrm{G}\right)} = \int \mathcal{L}(x; \mathcal{A}, \lambda)\, P\left(\mathcal{A}, \lambda|\mathcal{H}_\mathrm{S}\right)\, d\mathcal{A}\, d\lambda\,, \tag{12}$$

with the likelihood-ratio *function* obtained from Eqs. (8) and (9) as

$$\mathcal{L}(x; \mathcal{A}, \lambda) \equiv \frac{P\left(x|\mathcal{H}_\mathrm{S}, \mathcal{A}, \lambda\right)}{P\left(x|\mathcal{H}_\mathrm{G}\right)} = \exp\left[\langle x|s\rangle - \frac{1}{2}\, \langle s|s\rangle\right]\,. \tag{13}$$

We further introduce the partially-marginalized (over amplitude parameters $\mathcal{A}$) Bayes factor $B_\mathrm{S/G}(x;\lambda)$ as

$$B_\mathrm{S/G}(x;\lambda) \equiv \int \mathcal{L}(x; \mathcal{A}, \lambda)\, P\left(\mathcal{A}|\lambda, \mathcal{H}_\mathrm{S}\right)\, d\mathcal{A}\,, \tag{14}$$

such that

$$B_{\mathrm{S/G}}(x) = \int \left[\int \mathcal{L}(x;\mathcal{A},\lambda)\, P\left(\mathcal{A}|\lambda,\mathcal{H}_{\mathrm{S}}\right)\, d\mathcal{A} \right] P\left(\lambda|\mathcal{H}_{\mathrm{S}}\right)\, d\lambda$$
$$= \int B_{\mathrm{S/G}}(x;\lambda)\, P\left(\lambda|\mathcal{H}_{\mathrm{S}}\right)\, d\lambda\,. \tag{15}$$

The posterior on the signal parameters is

$$P\left(\mathcal{A},\lambda|x,\mathcal{H}_{\mathrm{S}}\right) = P\left(x|\mathcal{H}_{\mathrm{S}},\mathcal{A},\lambda\right)\, \frac{P\left(\mathcal{A},\lambda|\mathcal{H}_{\mathrm{S}}\right)}{P\left(x|\mathcal{H}_{\mathrm{S}}\right)}$$
$$\propto \mathcal{L}(x;\mathcal{A},\lambda)\, P\left(\mathcal{A}|\lambda,\mathcal{H}_{\mathrm{S}}\right)\, P\left(\lambda|\mathcal{H}_{\mathrm{S}}\right)\,, \tag{16}$$

where we dropped all factors that independent of $\{\mathcal{A},\lambda\}$. The (marginal) posterior on the evolution parameters λ is therefore

$$P\left(\lambda|x,\mathcal{H}_{\mathrm{S}}\right) = \int P\left(\mathcal{A},\lambda|x,\mathcal{H}_{\mathrm{S}}\right)\, d\mathcal{A}$$
$$\propto B_{\mathrm{S/G}}(x;\lambda)\, P\left(\lambda|\mathcal{H}_{\mathrm{S}}\right)\,. \tag{17}$$

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

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

$$\langle s|s\rangle = \sum_X 2\int_{-\infty}^{\infty} \frac{|\widetilde{s}(f)|^2}{S_X(f)}\, df \tag{18}$$
$$= 2N_{\mathrm{det}} \int \frac{|\widetilde{s}(f)|^2}{\mathcal{S}(f)}\, df\,, \tag{19}$$

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

$$\mathcal{S}^{-1}(f) \equiv \frac{1}{N_{\mathrm{det}}} \sum_X S_X^{-1}(f)\,. \tag{20}$$

Using the factorization of Eq. (5), which in frequency domain yields

$$\widetilde{s}(f;\mathcal{A},\lambda) = \mathcal{A}_{\mathrm{s}}\, \widetilde{h}_{\mathrm{s}}(f;\lambda) + \mathcal{A}_{\mathrm{c}}\, \widetilde{h}_{\mathrm{c}}(f;\lambda)\,, \tag{21}$$

we can further write this as

$$\langle s|s\rangle = \mathcal{A}\cdot\mathcal{M}(\lambda)\cdot\mathcal{A}\,, \tag{22}$$

with

$$\mathcal{M}(\lambda) \equiv 2N_{\mathrm{det}} \begin{pmatrix} I_{\mathrm{s}} & I_{\mathrm{sc}} \\ I_{\mathrm{sc}} & I_{\mathrm{c}} \end{pmatrix}\,, \tag{23}$$

$$I_{\mathrm{s}}(\lambda) = 2\int_0^{\infty} \frac{|\widetilde{h}_{\mathrm{s}}(f)|^2}{\mathcal{S}(f)}\, df\,, \tag{24}$$

$$I_{\mathrm{c}}(\lambda) = 2\int_0^{\infty} \frac{|\widetilde{h}_{\mathrm{c}}(f)|^2}{\mathcal{S}(f)}\, df\,, \tag{25}$$

$$I_{\mathrm{sc}}(\lambda) = 2\int_0^{\infty} \frac{\Re[\widetilde{h}_{\mathrm{s}}(f)\, \widetilde{h}_{\mathrm{c}}^{\,*}(f)]}{\mathcal{S}(f)}\, df\,. \tag{26}$$

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

$$\widetilde{h}_{\rm s}(f;\lambda) = \tau\,\frac{2\pi f_0\,\tau}{1+i\,4\pi\,f\,\tau - 4\pi^2(f^2-f_0^2)\tau^2}\,e^{-i2\pi f t_0}\,, \tag{27}$$

$$\widetilde{h}_{\rm c}(f;\lambda) = \tau\,\frac{1+i\,2\pi\,f\,\tau}{1+i\,4\pi\,f\,\tau - 4\pi^2(f^2-f_0^2)\tau^2}\,e^{-i2\pi f t_0}\,, \tag{28}$$

and are shown for two parameter-space choices in Fig. 1.

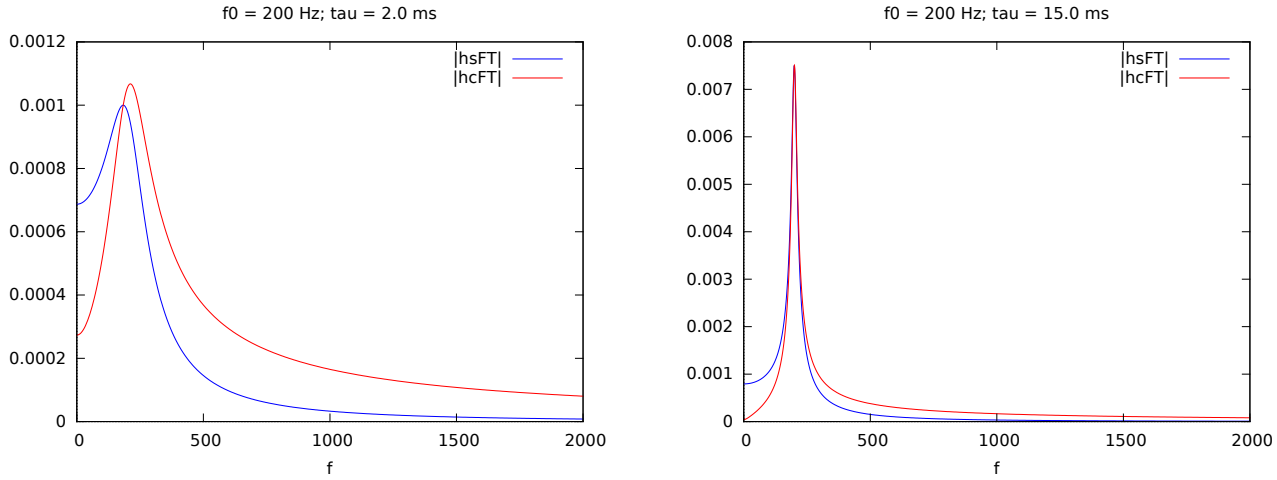


FIG. 1. Fourier power of ringdown components $|\widetilde{h}_{\rm s}|$ and $|\widetilde{h}_{\rm c}|$ for $f_0 = 200\,\text{Hz}$ and two different damping times, $\tau = 2\,\text{ms}$ (left) and $\tau = 15\,\text{ms}$ (right).

B. Expressing the "matched filter" $\langle x|s(\mathcal{A},\lambda)\rangle$

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

$$\widetilde{y}^X \equiv \frac{\widetilde{x}^X(f)}{S_X(f)}\,, \quad \widetilde{y} \equiv \sum_X \widetilde{y}^X\,. \tag{29}$$

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

$$\langle x|s\rangle = \sum_X 2\int_{-\infty}^{\infty} \widetilde{y}^X(f)\,\widetilde{s}^{*X}(f)\,df \tag{30}$$

$$= 2\int_{-\infty}^{\infty} \widetilde{y}(f)\,\widetilde{s}^{*}(f)\,df \tag{31}$$

$$= 2\int_{t_0}^{t_0+T} y(t)\,s(t)\,dt\,, \tag{32}$$

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. (5) we can further write

$$\langle x|s(\mathcal{A},\lambda)\rangle = \mathcal{A}\cdot\vec{x}(\lambda) = \mathcal{A}_{\rm s}\,x_{\rm s}(\lambda) + \mathcal{A}_{\rm c}\,x_{\rm c}(\lambda)\,, \tag{33}$$

with

$$x_{\mathrm{s}}(\lambda) \equiv 2 \int_{t_0}^{t_0+T} y(t)\, h_{\mathrm{s}}(t;\lambda)\, dt\,, \tag{34}$$

$$x_{\mathrm{c}}(\lambda) \equiv 2 \int_{t_0}^{t_0+T} y(t)\, h_{\mathrm{c}}(t;\lambda)\, dt\,. \tag{35}$$

Note that it will be convenient to write

$$h_{\mathrm{exp}} \equiv h_{\mathrm{c}}(t;\lambda) - i\, h_{\mathrm{s}}(t;\lambda) = e^{-\Delta t/\tau}\, e^{-i\, 2\pi\, f_0\, \Delta t} = e^{-\Delta t\, \varpi}\,, \tag{36}$$

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

$$\varpi \equiv \frac{1}{\tau} + i\, 2\pi f_0\,, \tag{37}$$

and so we obtain the complex matched-filter as

$$F \equiv x_{\mathrm{c}} - i\, x_{\mathrm{s}} = 2 \int_0^T y(t_0 + \Delta t)\, e^{-\Delta t\, \varpi}\, d\Delta t\,, \tag{38}$$

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. (13), we can write this as

$$\ln \mathcal{L}(x;\mathcal{A},\lambda) = \langle x|s\rangle - \frac{1}{2}\,\langle s|s\rangle \tag{39}$$

$$= -\frac{1}{2}\,\mathcal{A}\cdot\mathcal{M}\cdot\mathcal{A} + \mathcal{A}\cdot\vec{x}\,, \tag{40}$$

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. (15) for a suitable choice of prior $P\,(\mathcal{A}|\lambda,\mathcal{H}_{\mathrm{S}})$.

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. $P\,(\mathcal{A}|\lambda,\mathcal{H}_{\mathrm{S}}) = P\,(\mathcal{A}|\mathcal{H}_{\mathrm{S}})$

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

$$P\,(\mathcal{A}|\mathcal{H}_{\mathrm{S}},H) = \frac{1}{2\pi\, H^2}\, e^{-\frac{1}{2}\mathcal{A}\cdot\mathcal{A}/H^2}\,, \tag{41}$$

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

$$P\,(A|\mathcal{H}_{\mathrm{S}},H) = \frac{A}{H^2}\, e^{-\frac{A^2}{2H^2}}\,. \tag{42}$$

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

$$B_{\mathrm{S/G}}(x;\lambda,H) \equiv \frac{P\,(x|\mathcal{H}_{\mathrm{S}},\lambda,H)}{P\,(x|\mathcal{H}_{\mathrm{G}})} \tag{43}$$

$$= \frac{1}{2\pi H^2} \int e^{-\frac{1}{2}\,\mathcal{A}\cdot\gamma^{-1}\cdot\mathcal{A} + \mathcal{A}\cdot\vec{x}}\, d\mathcal{A} \tag{44}$$

$$= \frac{\sqrt{\det\gamma}}{H^2}\, e^{\frac{1}{2}\,\vec{x}\cdot\gamma\cdot\vec{x}} \tag{45}$$

with

$$\gamma^{-1}(\lambda) \equiv \mathcal{M} + H^{-2}\,\mathbb{I} = \begin{pmatrix} \mathcal{M}_{\mathrm{ss}} + H^{-2} & \mathcal{M}_{\mathrm{sc}} \\ \mathcal{M}_{\mathrm{sc}} & \mathcal{M}_{\mathrm{cc}} + H^{-2}\,. \end{pmatrix}, \tag{46}$$

and determinant

$$\det\gamma^{-1} = \left(\mathcal{M}_{\rm ss} + H^{-2}\right)\left(\mathcal{M}_{\rm cc} + H^{-2}\right) - \mathcal{M}_{\rm sc}^2 \tag{47}$$

$$= \det\mathcal{M} + H^{-2}\,\mathrm{tr}\mathcal{M} + H^{-4}\,, \tag{48}$$

inverse

$$\gamma(\lambda) = \frac{1}{\det\gamma^{-1}}\begin{pmatrix}\mathcal{M}_{\rm cc} + H^{-2} & -\mathcal{M}_{\rm sc}\\ -\mathcal{M}_{\rm sc} & \mathcal{M}_{\rm ss} + H^{-2}\end{pmatrix}\,, \tag{49}$$

and

$$\frac{\sqrt{\det\gamma}}{H^2} = \left[H^4\,\det\mathcal{M} + H^2\,\mathrm{tr}\mathcal{M} + 1\right]^{-1/2}\,. \tag{50}$$

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 $B_{\rm S/G} \to 1$. 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 $B_{\rm S/G} \to 0$ 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*

A 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

$$P\left(H_i|\mathcal{H}_{\rm S}\right) = c\,\frac{1}{H_i},, \quad \text{with} \quad c^{-1} = \sum_i H_i^{-1}\,. \tag{51}$$

The resulting amplitude prior is now obtained by marginalizing the fixed-H prior of Eq. (41) over H, which yields

$$P\left(\mathcal{A}|\mathcal{H}_{\rm S}\right) = \sum_i P\left(\mathcal{A}|\mathcal{H}_{\rm S}, H_i\right) P\left(H_i|\mathcal{H}_{\rm S}\right)\,. \tag{52}$$

We can further estimate the unknown H parameter from the data via Eq. (43), namely

$$P\left(H|x\right) = \int P\left(\lambda, H|x, \mathcal{H}_{\rm S}\right)\,d\lambda \tag{53}$$

$$\propto \left[\int P\left(x|\mathcal{H}_{\rm S}, \lambda, H\right)\,P\left(\lambda|\mathcal{H}_{\rm S}\right)\,d\lambda\right]\,P\left(H|\mathcal{H}_{\rm S}\right) \tag{54}$$

$$\propto B_{\rm S/G}(x; H)\,P\left(H|\mathcal{H}_{\rm S}\right) \tag{55}$$

$$= c\,\frac{1}{H_i}\,B_{\rm S/G}(x; H_i)\,. \tag{56}$$

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

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

$$= \int B_{\rm S/G}(x; \lambda, H) P\left(H|\mathcal{H}_{\rm S}\right)\,dH \tag{58}$$

$$= c\sum_i \frac{1}{H_i}\,B_{\rm S/G}(x; \lambda, H_i)\,. \tag{59}$$

III. AMPLITUDE PARAMETER ESTIMATION

The posterior for the amplitudes $\mathcal{A}$ at fixed $\lambda_{\rm MP}$ can directly be obtained from Eq. (16), Eq. (52) and Eq. (41) as

$$P\left(\mathcal{A}|x,\mathcal{H}_{\rm S}\right) \propto \mathcal{L}(x;\mathcal{A},\lambda)\, P\left(\mathcal{A}|\mathcal{H}_{\rm S}\right) \tag{60}$$

$$= \sum_i \mathcal{L}(x;\mathcal{A},\lambda)\, P\left(\mathcal{A}|\mathcal{H}_{\rm S},H_i\right)\, P\left(H_i|\mathcal{H}_{\rm S}\right) \tag{61}$$

$$= \sum_i \frac{1}{H_i^2}\, \exp\left[-\frac{1}{2}\mathcal{A}\cdot\gamma^{-1}(H_i)\cdot\mathcal{A} + \mathcal{A}\cdot\vec{x}\right]\, P\left(H_i|\mathcal{H}_{\rm S}\right)\,. \tag{62}$$

We can compute this expression and maximize over $P\left(A|x,\mathcal{H}_{\rm S}\right)$ to obtain $\mathcal{A}_{\rm MP}$. Alternatively, we could replace the unknown H by its maximum-posterior estimate $H_{\rm MP}$ of Eq. (53), and obtain the estimate

$$P\left(\mathcal{A}|x,\mathcal{H}_{\rm S},H_{\rm MP}\right) \propto \exp\left[-\frac{1}{2}\mathcal{A}\cdot\gamma^{-1}(H_{\rm MP})\cdot\mathcal{A} + \mathcal{A}\cdot\vec{x}\right]\,, \tag{63}$$

which can be maximized over $\mathcal{A}$ analytically to yield

$$\mathcal{A}'_{\rm MP} \approx \gamma(H_{\rm MP})\cdot\vec{x}\,. \tag{64}$$

From these estimates we can obtain $A = \sqrt{\mathcal{A}_{\rm s}^2 + \mathcal{A}_{\rm c}^2}$ and $\phi'_0 = -\tan^{-1}\left(\frac{\mathcal{A}_{\rm s}}{\mathcal{A}_{\rm c}}\right)$.

We can further estimate an "SNR" in the MPE template $\lambda_{\rm MP}$, by substituting the amplitude estimates $\mathcal{A}_{\rm MP}$ into the SNR expression of Eq. (22), i.e.

$$\rho_0^2 = \mathcal{A}_{\rm MP}\cdot\mathcal{M}\cdot\mathcal{A}_{\rm MP}\,. \tag{65}$$

IV. QNM SEARCH APPLIED TO GW150914

A. Prior choices

Isotropic 2D Gaussian amplitude prior (41) on $\{\mathcal{A}_{\rm s},\mathcal{A}_{\rm c}\}$ with characteristic amplitudes $H = [2:10]\times10^{-22}$, which corresponds to an isotropic prior in ϕ_0 and an A-prior of (42), as shown here:

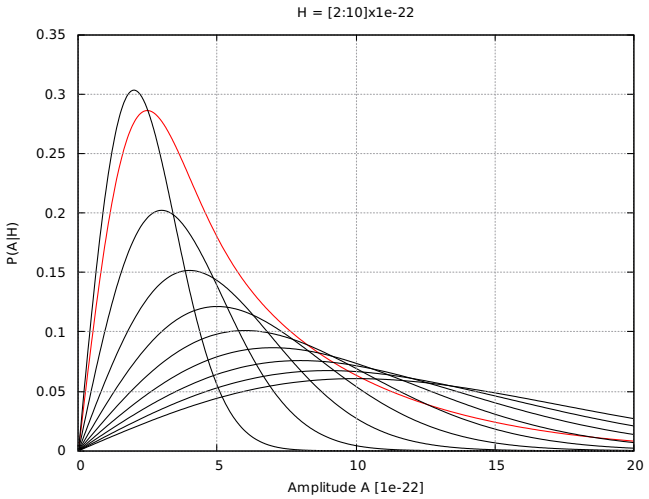


FIG. 2. Amplitude prior (red line) as a hyper-prior (weighted superposition) of several 2D-Gaussian distributions (black lines) with different scale parameters $H \in [2:10]\times10^{-22}$

B. Data preparation

We use the data from the respective 1800s-SFT (containing the frequency range $[10, 2000]$ Hz) from each detector covering GW150914. The data is inverse-FFTed into the time-domain after applying a tukey(0.02) windowing (to avoid boundary effects). The PSD is estimated on this 1800s time-series using `pwelch()` in octave, with a window-size of $T = 8$ s, but *excluding* the "on source" segment centered on GW150914 (or any injections) to avoid signals affecting the PSD estimate.

We then extract the "on source" $T = 8$ s segment centered on GW150914, FFT it back into frequency domain (after tukey(0.1) windowing to avoid boundary effects), and divide it by the PSD to over-whiten the data. We extract the frequency band of $[20, 1900]$ Hz, tukey(0.02)-window it, then inverse-FFT it back into the time-domain, resulting in the over-whitened timeseries y^X of Eq. (29). The L1 data is time-shifted by (delaying it by $+7$ ms) and multiplied it by (-1) to account for the inverse detector response.

The resulting PSD estimates and data spectra are shown in Fig. 3, and the whitened and over-whitened data spectra are shown in Fig. 4. Note that the 'data spectra' are showing the data from the $T = 8$ s "on-source" window.

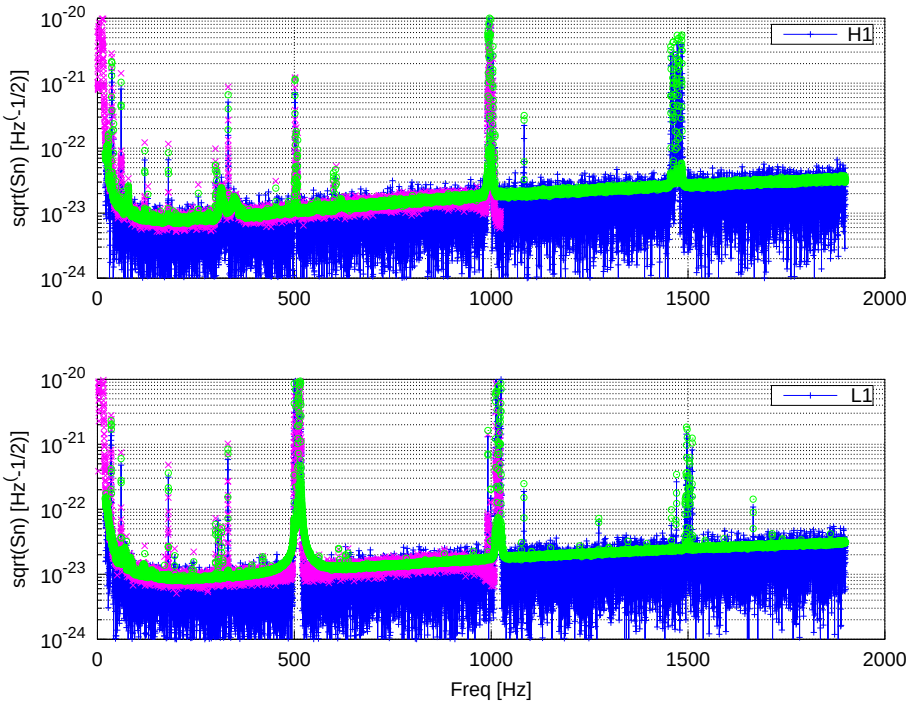


FIG. 3. Welch PSD estimate of $\sqrt{S_X(f)}$ (green 'o') over $f \in [20, 1900]$ Hz on 1800s of data (using an $T = 8$s window) for X = H1 (upper plot) and X = L1 (lower plot). LALInference PSD-estimate (magenta 'x') and normalized spectrum of $T = 8$ s of data used in the search, i.e. $|\widetilde{x}^X/\sqrt{T}|$ (blue line and '+').

Although the signal is quite wide-band in frequency domain (see Fig. 1), in Figs. 5 and 6 we show a zoom on the most relevant frequency range of $[200, 300]$ Hz.

C. Search results on GW150914

We search the $\{f_0, \tau\}$ range with uniform priors in $f_0 \in [200, 300]$ Hz and $\tau \in [0.5, 20]$ ms, in steps of $df_0 = 1$ Hz and $d\tau = 0.2$ 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 $t_{\rm M} = 1126259462.423$ (in H1 arrival time), as taken from Ian's wiki and rounded to ms accuracy. We compute snapshots for different QNM start-times $t_0 - t_{\rm M}$ (referring to H1 arrival times).

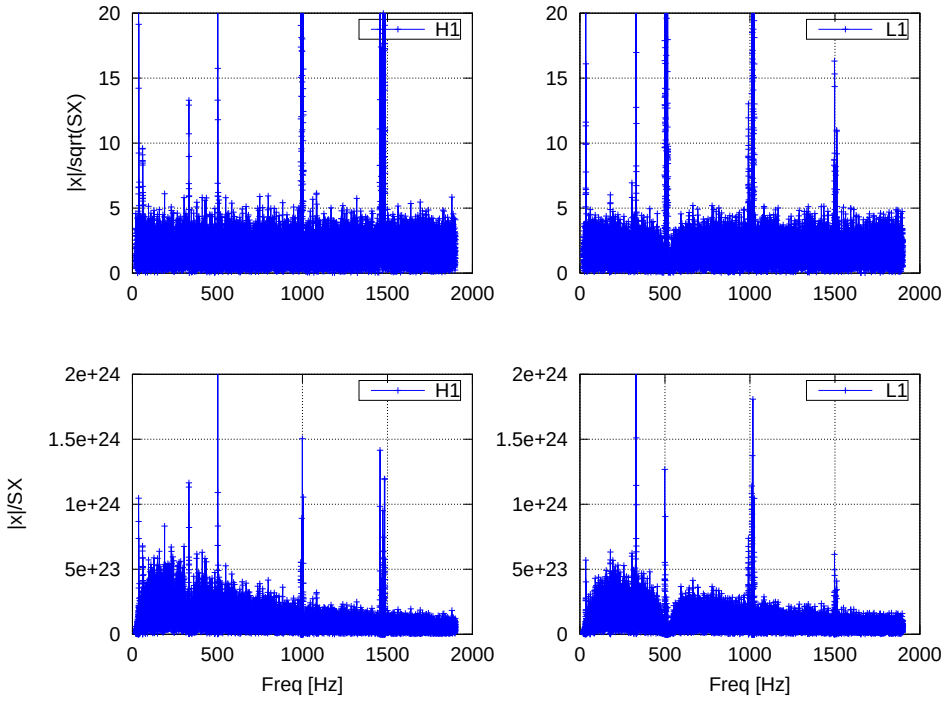


FIG. 4. Whitened $|\widetilde{x}^X/\sqrt{S_X}|$ (upper row) and over-whitened $|\widetilde{x}^X/S_X|$ (lower row) spectra of $T = 8\,\mathrm{s}$ of data used in the search, for X = H1 (left column) and L1 (right column).

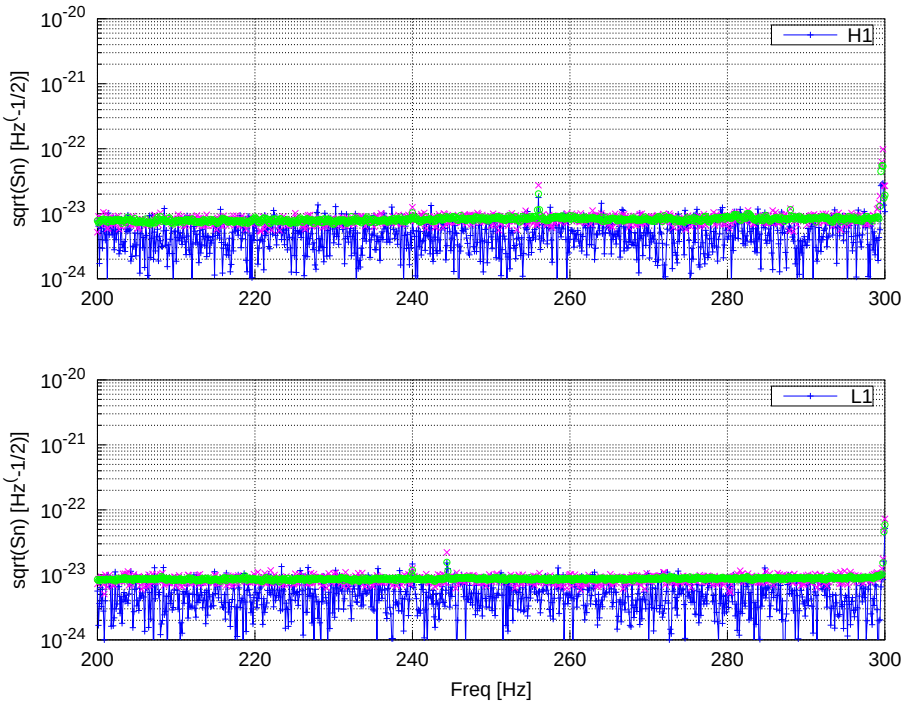


FIG. 5. Same as Fig. 3 but zoomed on frequency range $f \in [200, 300]\,\mathrm{Hz}$.

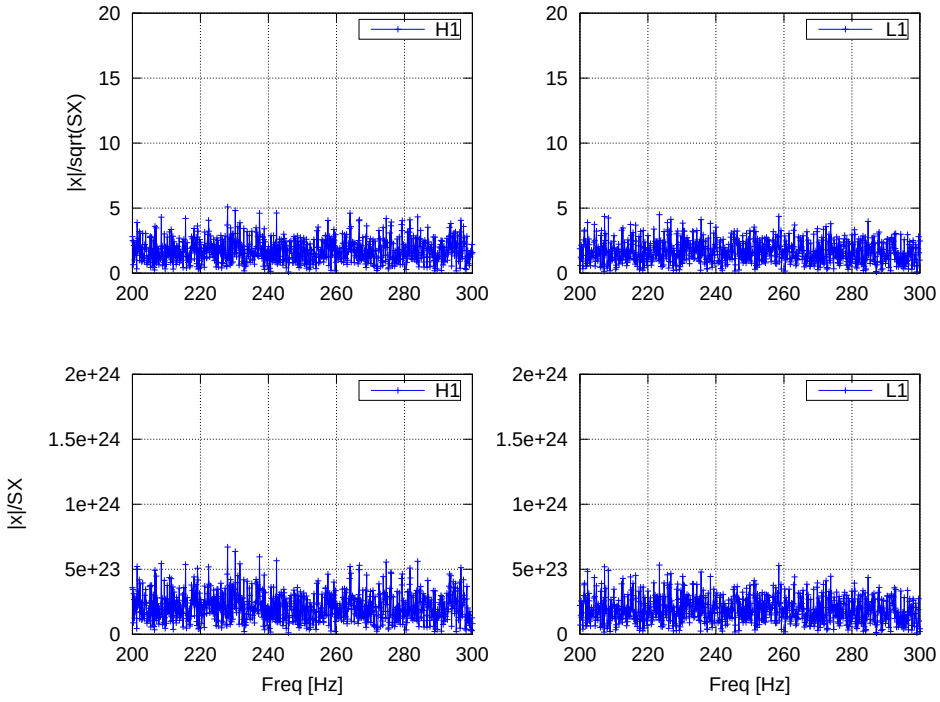
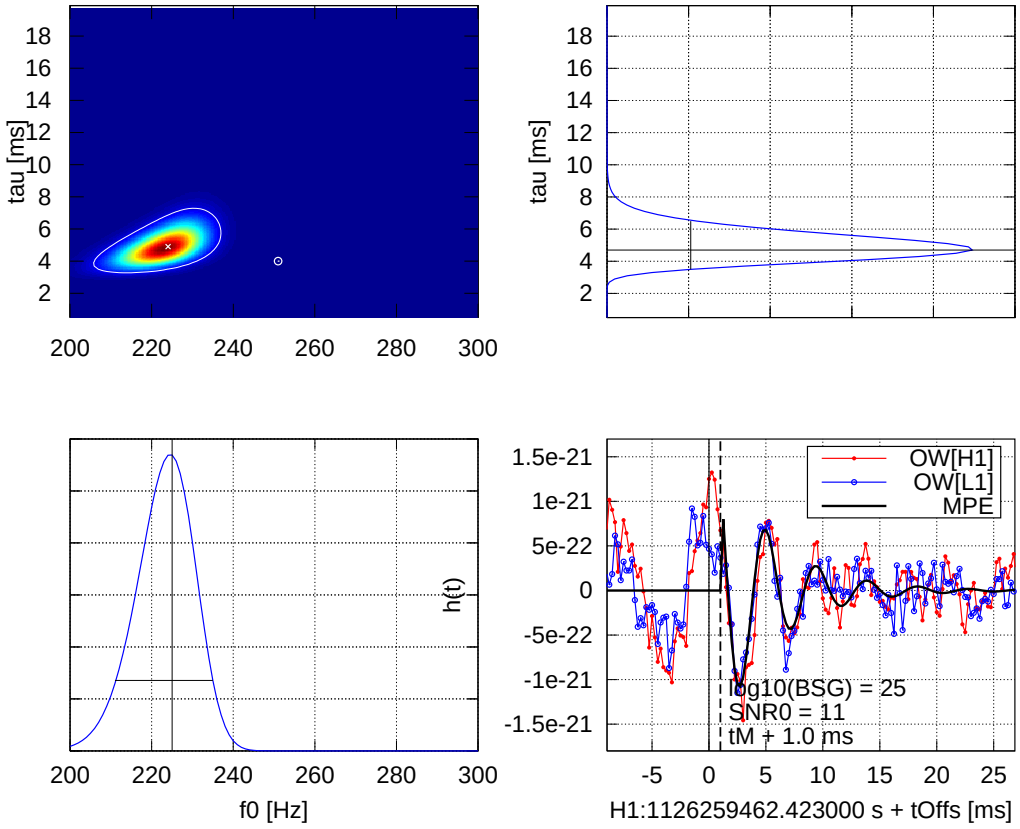
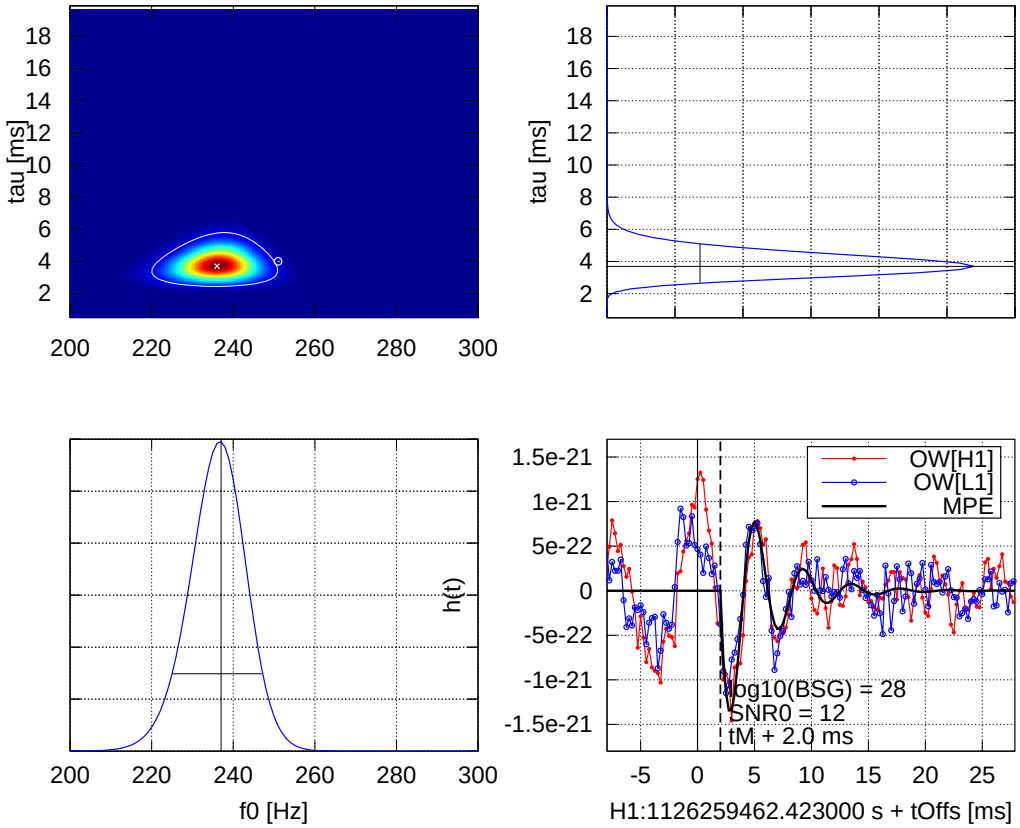
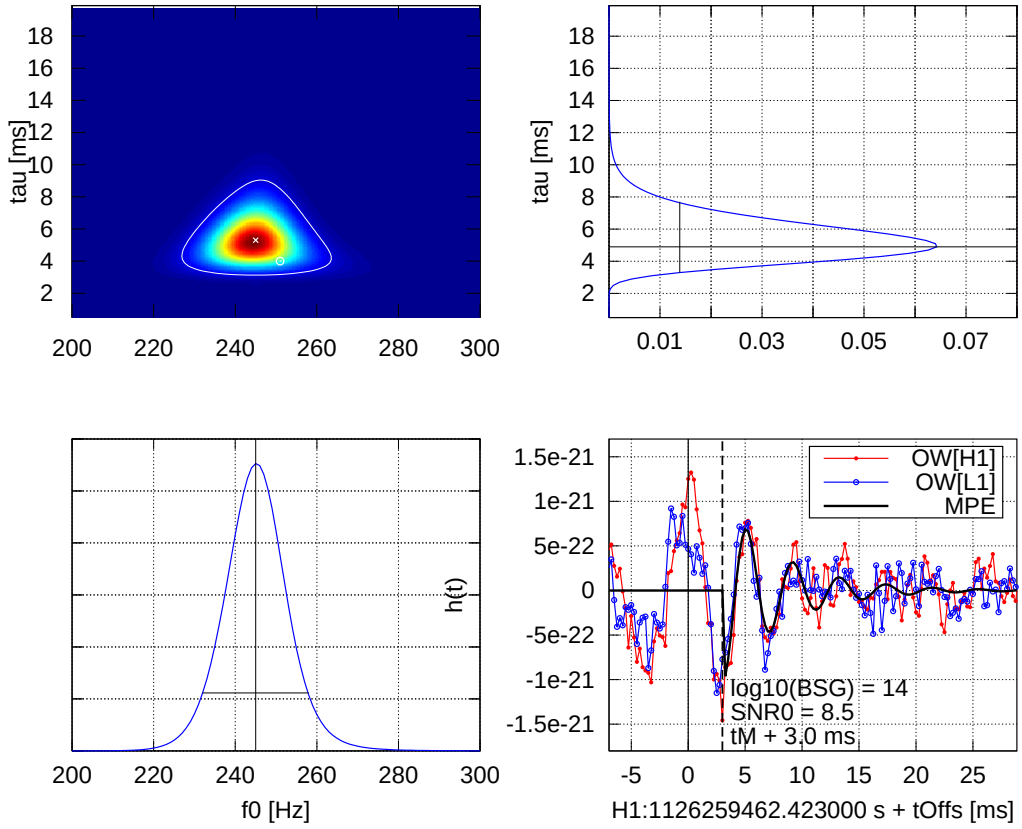
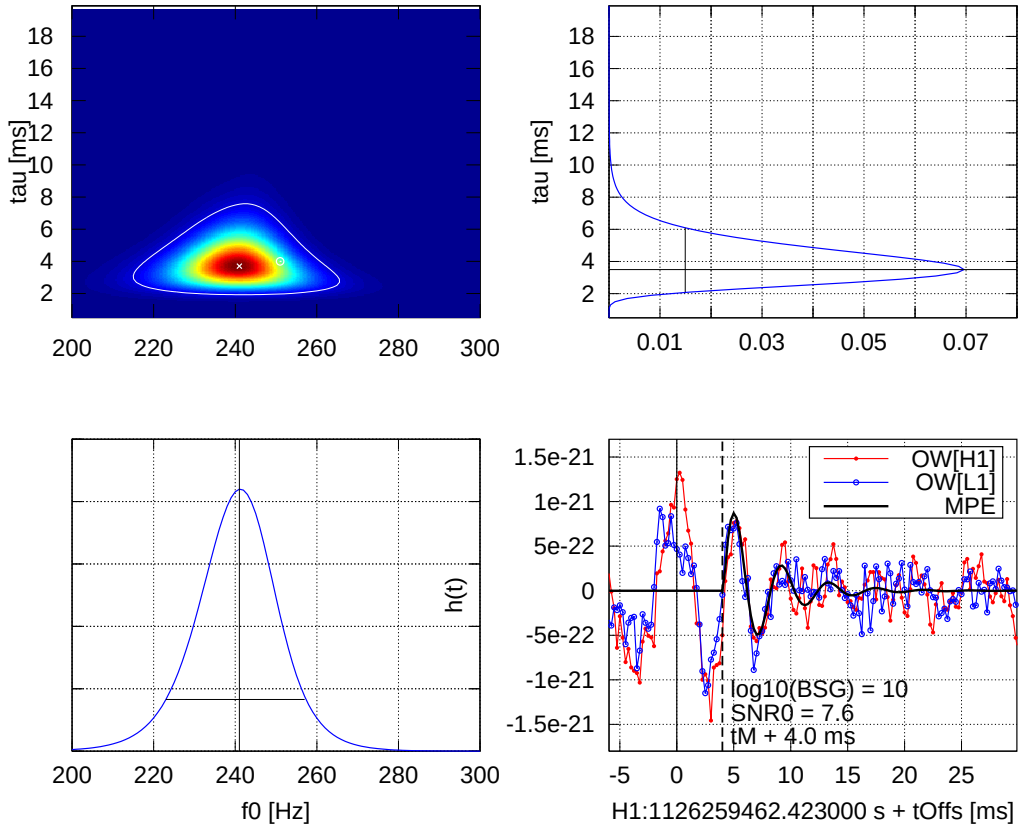


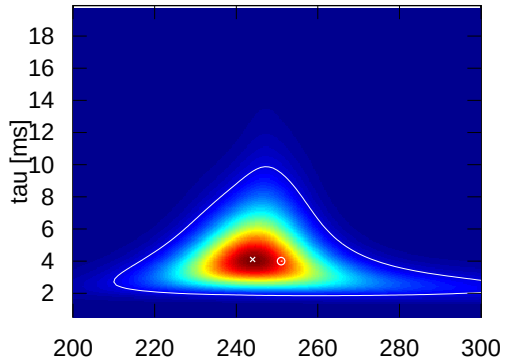
FIG. 6. Same as Fig. 4 but zoomed on frequency range $f \in [200, 300]$ Hz.

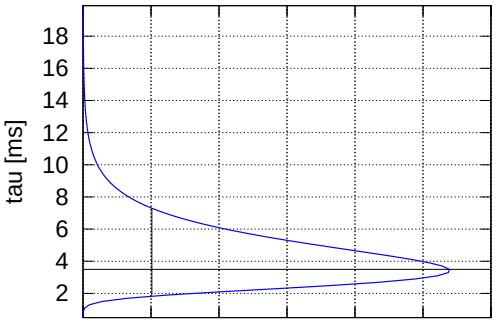

tau [ms]
18
16
14
12
10
8
6
4
2
200
220
240
260
280
300
h(t)
f0 [Hz]
1.5e-21
1e-21
5e-22
0
-5e-22
-1e-21
-1.5e-21
OW[H1]
OW[L1]
MPE
log10(BSG) = 25
SNR0 = 11
tM + 1.0 ms
-5
0
5
10
15
20
25
H1:1126259462.423000 s + tOffs [ms]

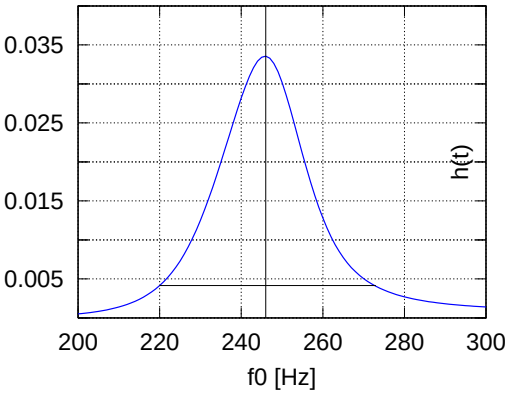

tau [ms]
18
16
14
12
10
8
6
4
2
200
220
240
260
280
300
h(t)
f0 [Hz]
1.5e-21
1e-21
5e-22
0
-5e-22
-1e-21
-1.5e-21
OW[H1]
OW[L1]
MPE
log10(BSG) = 28
SNR0 = 12
tM + 2.0 ms
-5
0
5
10
15
20
25
H1:1126259462.423000 s + tOffs [ms]

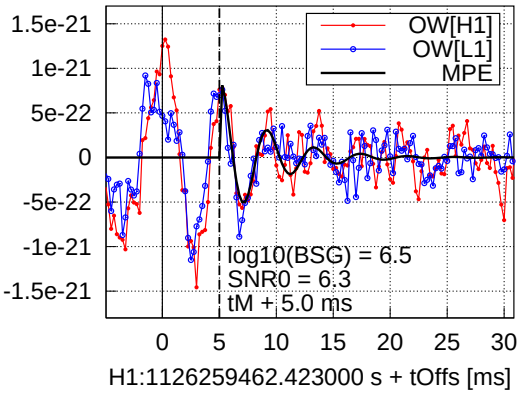

tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 14
SNR0 = 8.5
tM + 3.0 ms
H1:1126259462.423000 s + tOffs [ms]

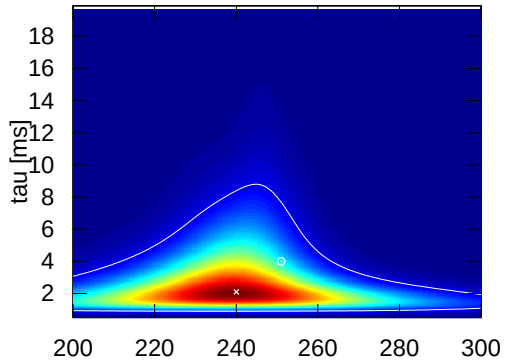

tau [ms]
f0 [Hz]
h(t)
OW[H1]
OW[L1]
MPE
log10(BSG) = 10
SNR0 = 7.6
tM + 4.0 ms
H1:1126259462.423000 s + tOffs [ms]

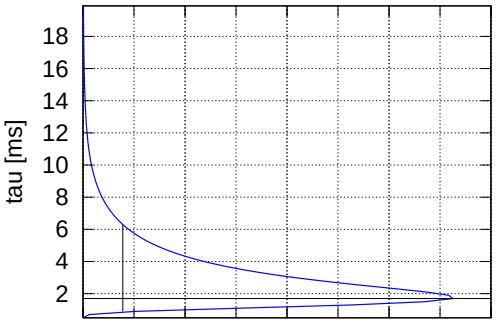

tau [ms]
18
16
14
12
10
8
6
4
2
200
220
240
260
280
300

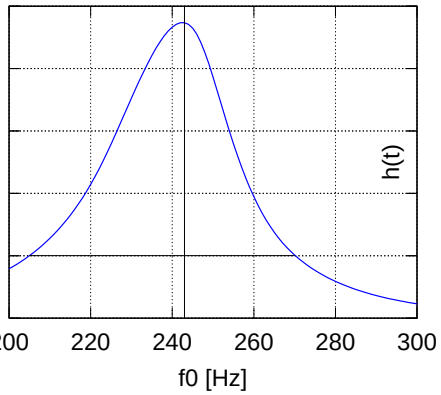

tau [ms]
18
16
14
12
10
8
6
4
2

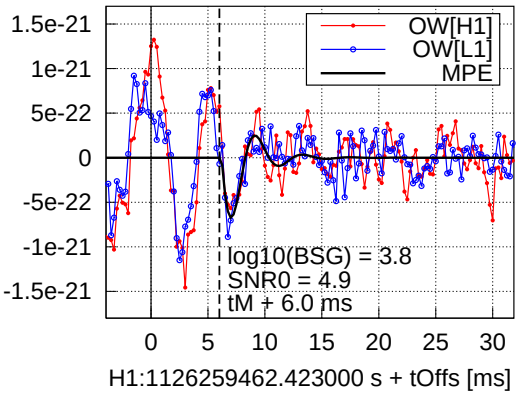

0.035
0.025
0.015
0.005
h(t)
200
220
240
260
280
300
f0 [Hz]

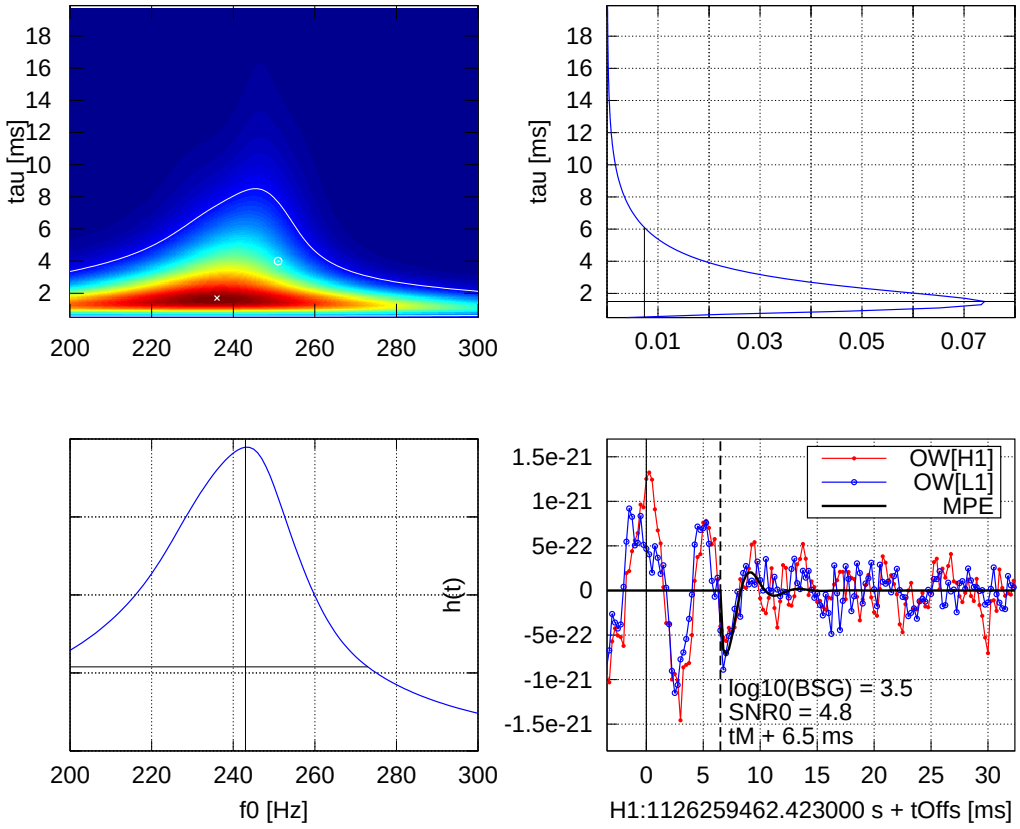

1.5e-21
1e-21
5e-22
0
-5e-22
-1e-21
-1.5e-21
OW[H1]
OW[L1]
MPE
log10(BSG) = 6.5
SNR0 = 6.3
tM + 5.0 ms
0
5
10
15
20
25
30
H1:1126259462.423000 s + tOffs [ms]

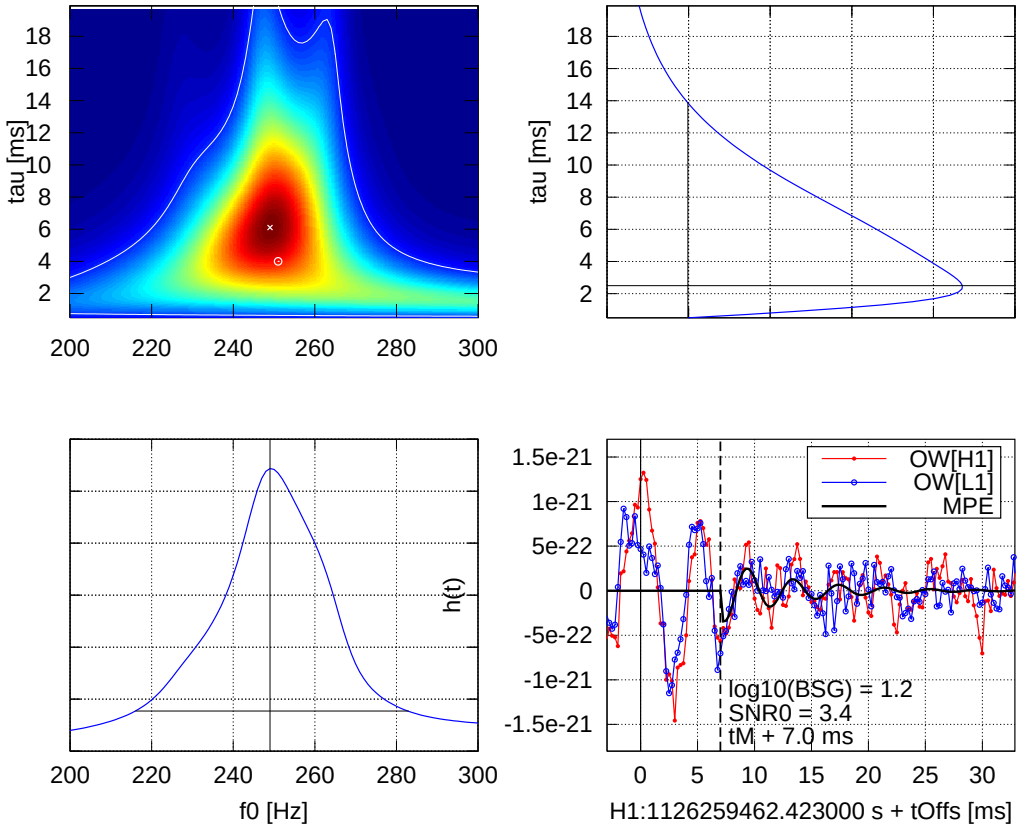

tau [ms]
18
16
14
12
10
8
6
4
2
200
220
240
260
280
300


tau [ms]
18
16
14
12
10
8
6
4
2
0.01
0.03
0.05
0.07

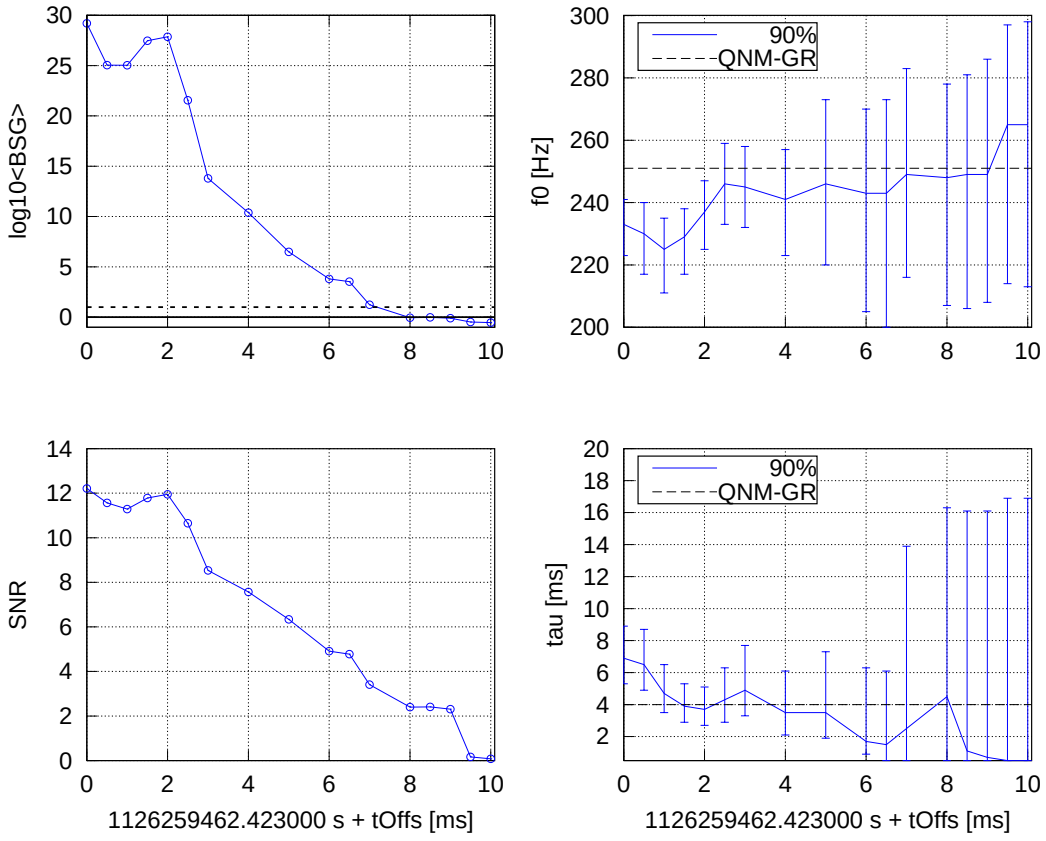

h(t)
200
220
240
260
280
300
f0 [Hz]

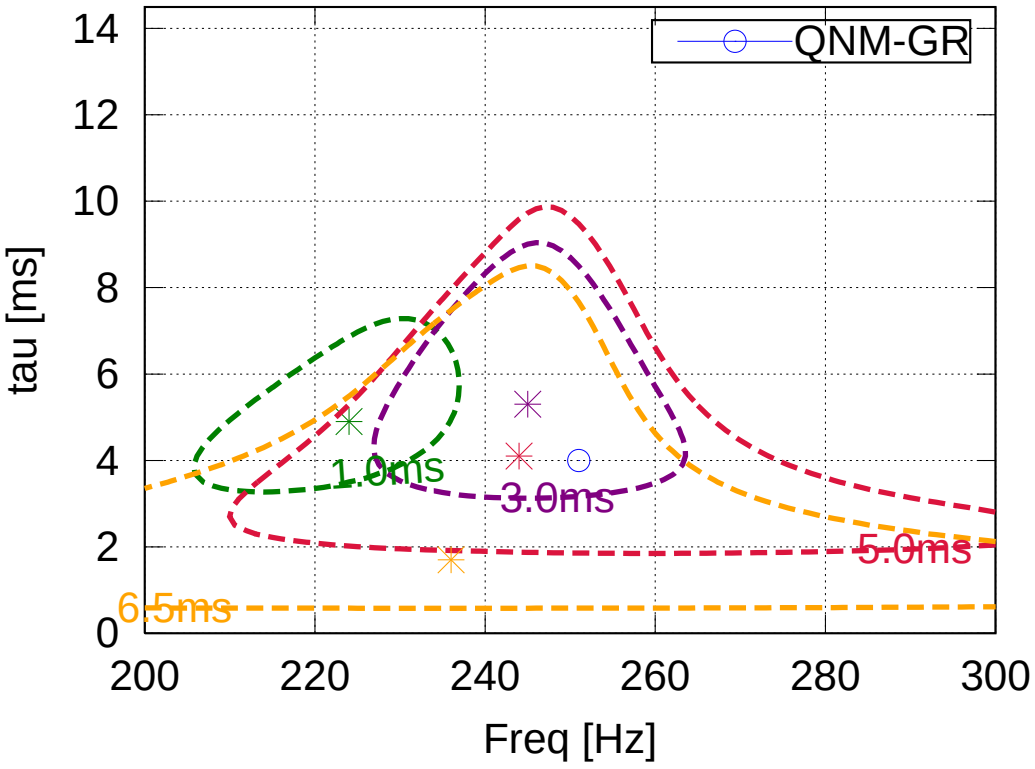

1.5e-21
1e-21
5e-22
0
-5e-22
-1e-21
-1.5e-21
OW[H1]
OW[L1]
MPE
log10(BSG) = 3.8
SNR0 = 4.9
tM + 6.0 ms
0
5
10
15
20
25
30
H1:1126259462.423000 s + tOffs [ms]


tau [ms]
18
16
14
12
10
8
6
4
2
200
220
240
260
280
300
0.01
0.03
0.05
0.07
h(t)
f0 [Hz]
1.5e-21
1e-21
5e-22
0
-5e-22
-1e-21
-1.5e-21
OW[H1]
OW[L1]
MPE
log10(BSG) = 3.5
SNR0 = 4.8
tM + 6.5 ms
0
5
10
15
20
25
30
H1:1126259462.423000 s + tOffs [ms]


tau [ms]
18
16
14
12
10
8
6
4
2
200
220
240
260
280
300
h(t)
f0 [Hz]
1.5e-21
1e-21
5e-22
0
-5e-22
-1e-21
-1.5e-21
OW[H1]
OW[L1]
MPE
log10(BSG) = 1.2
SNR0 = 3.4
tM + 7.0 ms
0
5
10
15
20
25
30
H1:1126259462.423000 s + tOffs [ms]

D. Dependence on $t_0 - t_M$



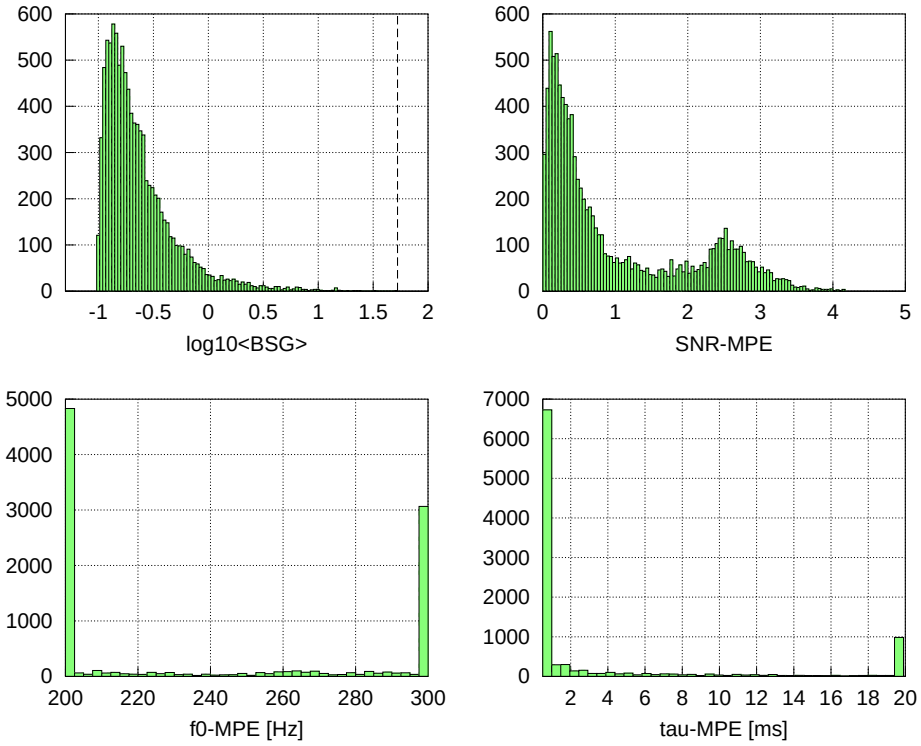


V. TESTING AND CHARACTERIZING PIPELINE PERFORMANCE

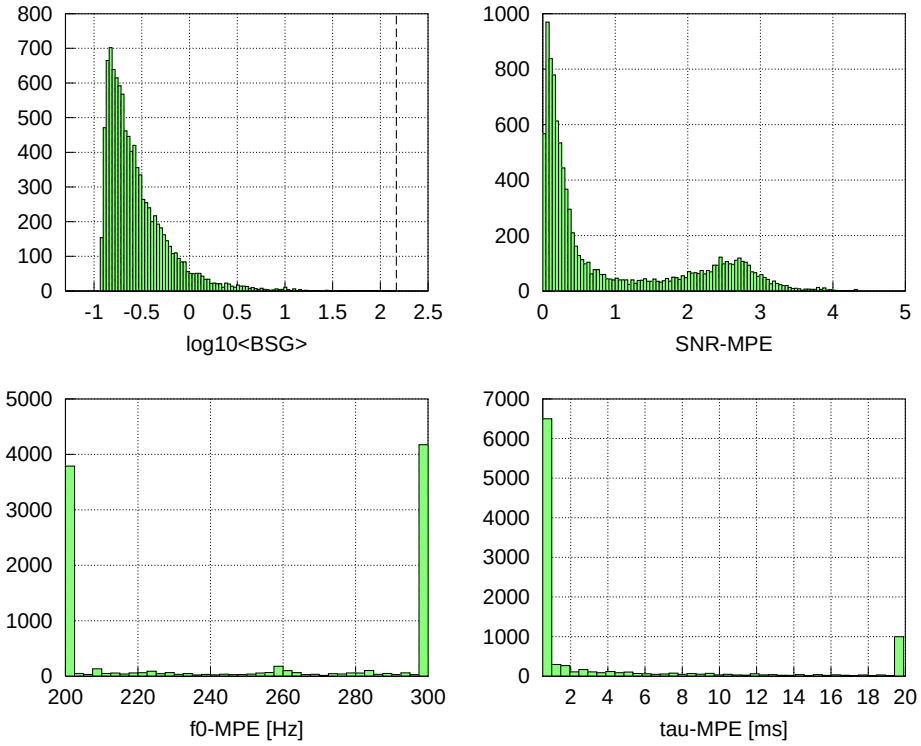
A. Off-source searches

We test off-source search performance on Gaussian white noise, as well as on real detector data around GW150914, using $10,000$ random start times $t_0 \in ([-3, -0.5]\,\mathrm{s} \cup [0.5, 3]\,\mathrm{s}) + t_\mathrm{M}$.

Searches in Gaussian noise:



Searches in real (off-source) data:



B. Testing parameter-estimation accuracy on injections

We test PE performance on injected QNM signals with perfectly-matched start-time t_0, using $10,000$ random start times $t_0 \in ([-3,-0.5]\,\mathrm{s} \cup [0.5,3]\,\mathrm{s}) + t_\mathrm{M}$.

In order to avoid injecting a discontinuity at t_0 (which would create problems when band-passing and FFTing the data), we preface the ringdown with a short "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 for the template, as shown in Fig. 7.

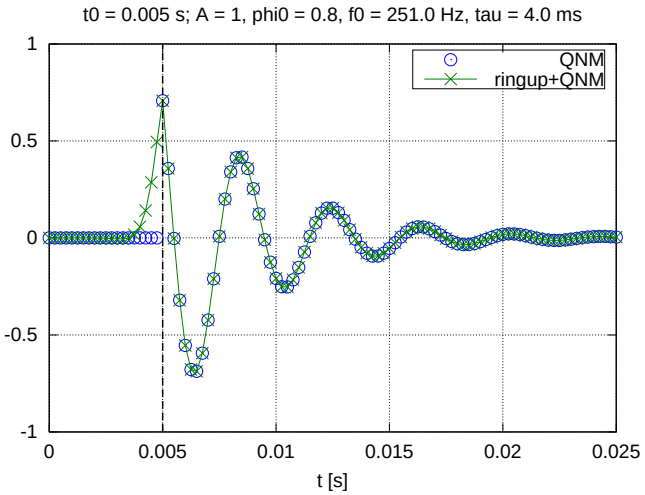


FIG. 7. Example QNM template (labelled 'QNM') used for matching, and ringup+QNM used for injections

We consider 5 different injection scenarios:

1. all QNM parameters drawn drawn from the priors: A via Eq. (52) and shown in Fig. 2, $\phi_0 \in [0,2\pi]$, $f_0 \in [200,300]\,\mathrm{Hz}$, and $\tau \in [0.5,20]\,\mathrm{ms}$, and

   (a) Pure signals without noise, assuming a white noise-floor of $\sqrt{S_X} = 8\times10^{-24}/\sqrt{\mathrm{Hz}}$.

   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 expected.

   (b) Signals injected in Gaussian white noise of $\sqrt{S_X} = 8\times10^{-24}/\sqrt{\mathrm{Hz}}$, assuming perfect knowledge of this noise PSD (instead of estimating it).

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

   (c) Signals injected in Gaussian white noise of $\sqrt{S_X} = 8\times10^{-24}/\sqrt{\mathrm{Hz}}$, with estimating the 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.

   (d) 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, i.e. whether it leads to over- or under-coverage.

2. QNM parameters with *fixed* "GR values" $f_0 = 251\,\mathrm{Hz}$, $\tau = 4\,\mathrm{ms}$, with A drawn uniformly $A \in [3,8]\times10^{-22}$, and $\phi_0 \in [0,2\pi]$, using real off-source detector data.

   This case only serves to illustrate how we expect the posterior coverage to perform if the signal actually was conforming to our expectation from GR, as suggested by the "on-source" results.

1. Case 1(a): QNM parameter drawn from priors, signal-only with assumed white noise

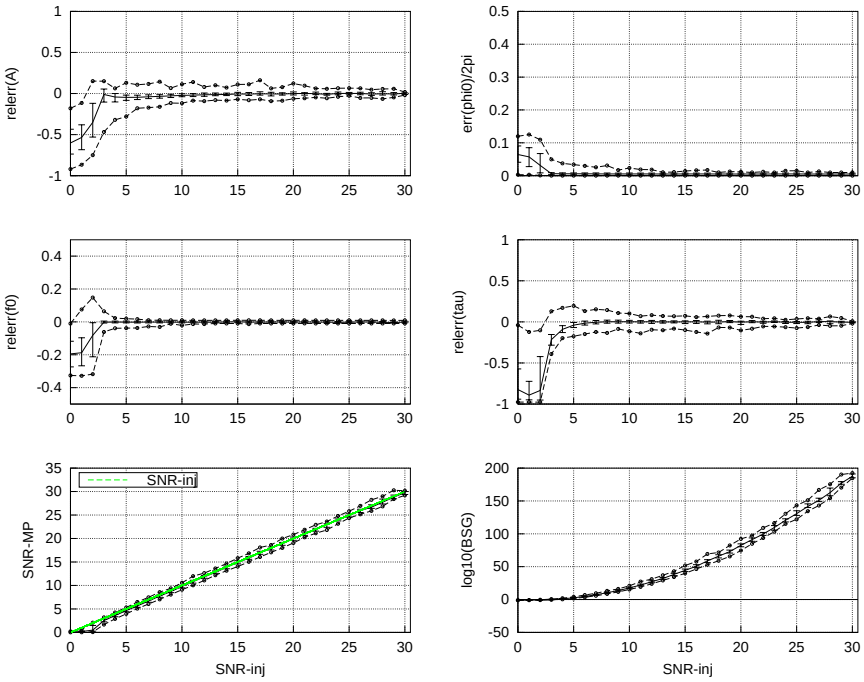


FIG. 8. MP estimation errors (top 2 rows), recovered SNR (bottom left) and Bayes-factor (bottom right) as functions of injected SNR. Solid line denotes the median, error-bars denote ±25%-ile and dashed lines are 2.5%- and 97.5%-iles, respectively.

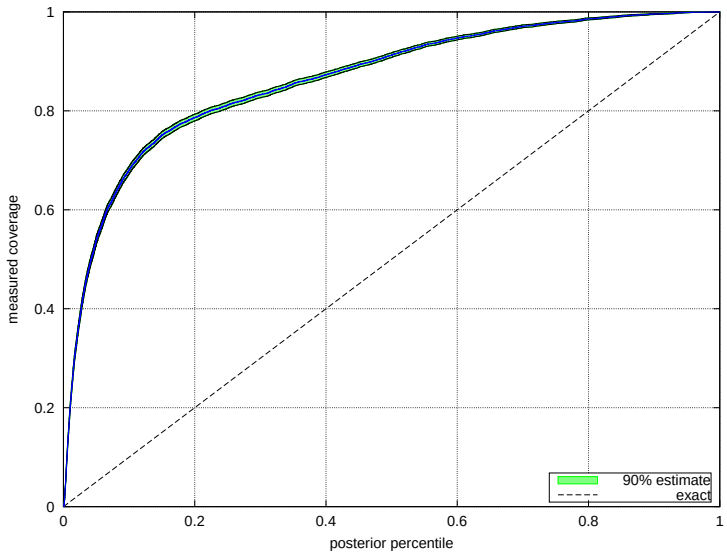
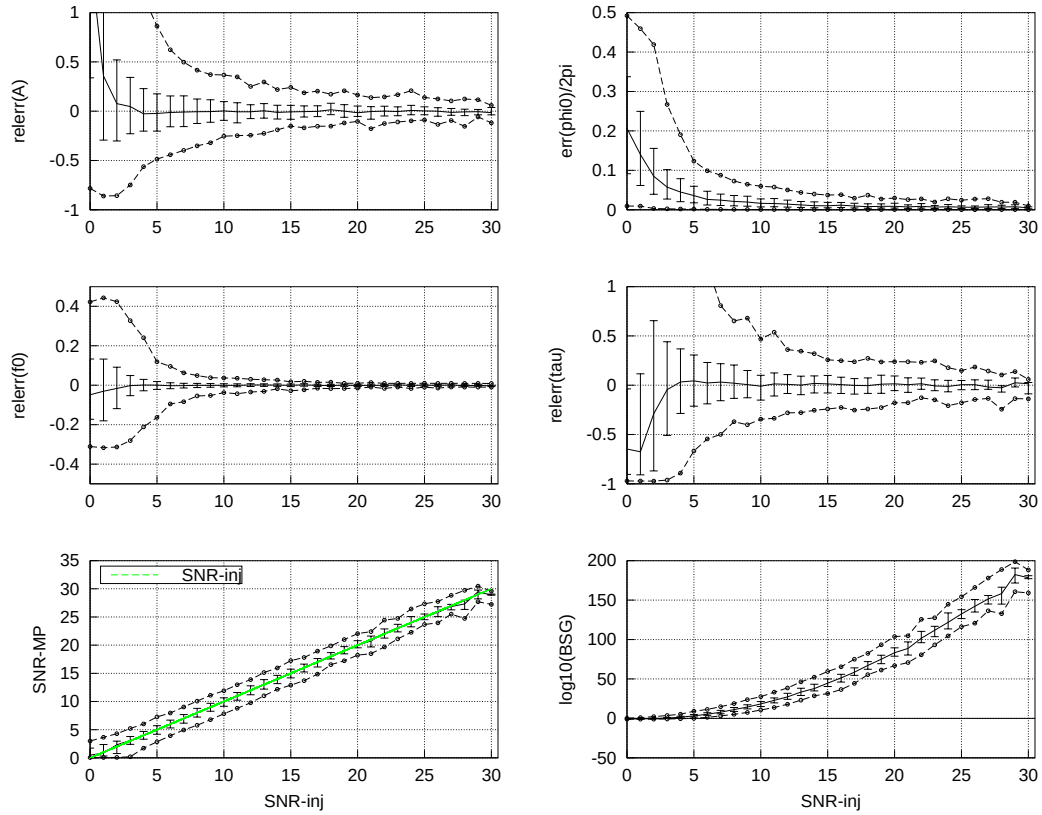
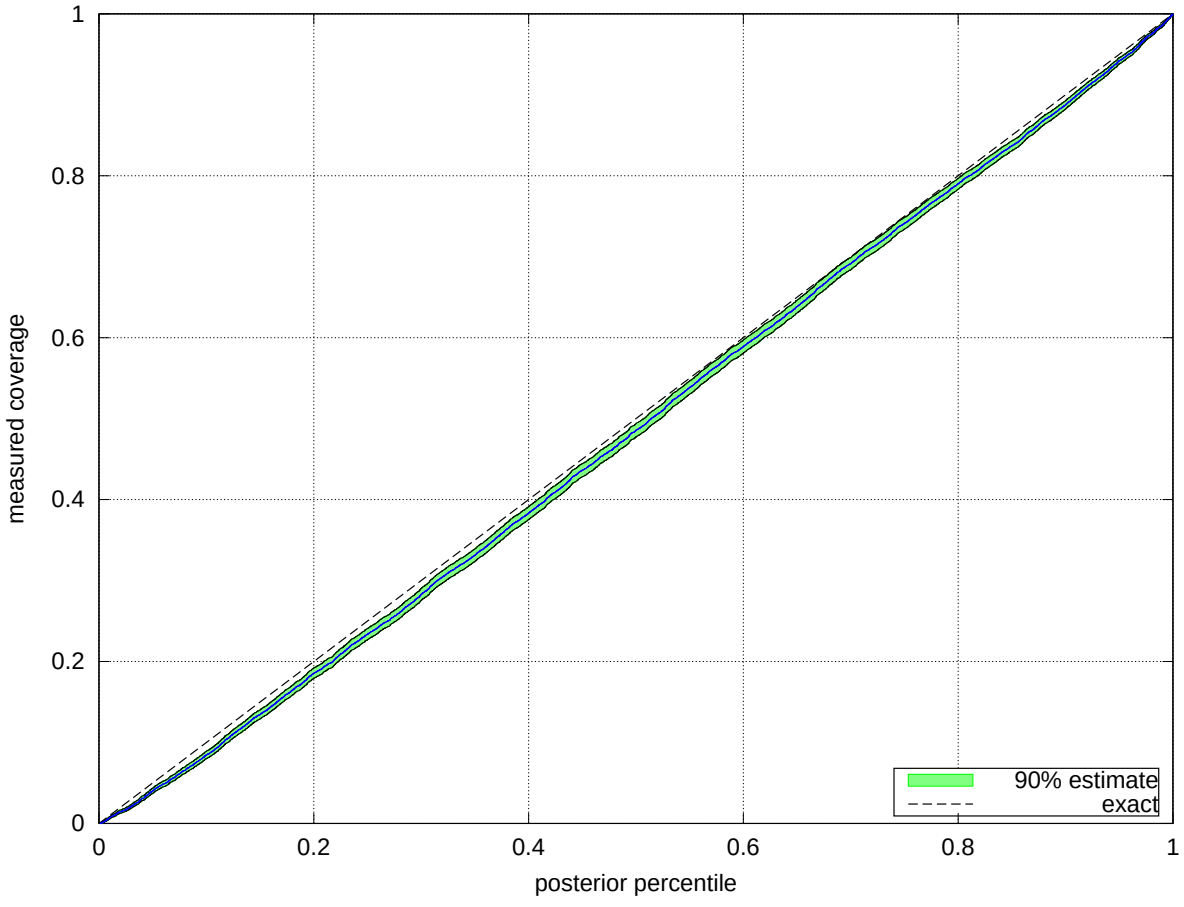


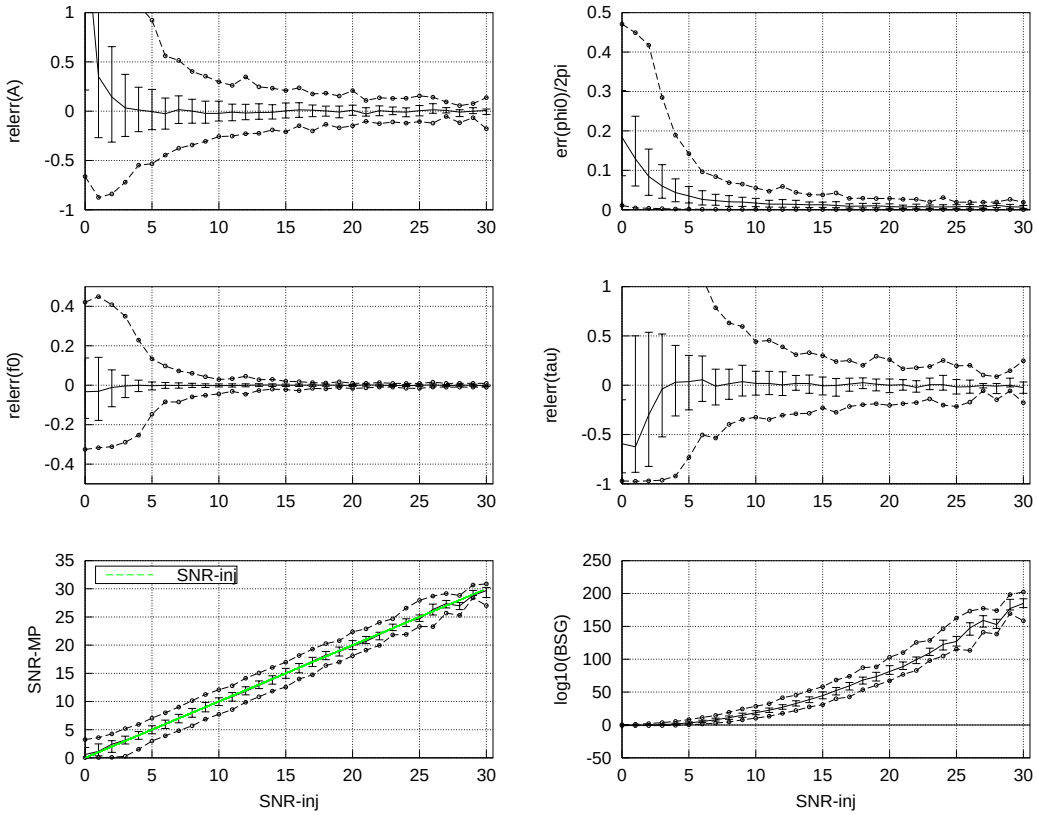
FIG. 9. Posterior coverage versus percentile ("pp-plot"), showing mean (solid line) and 90%-credible interval (green).

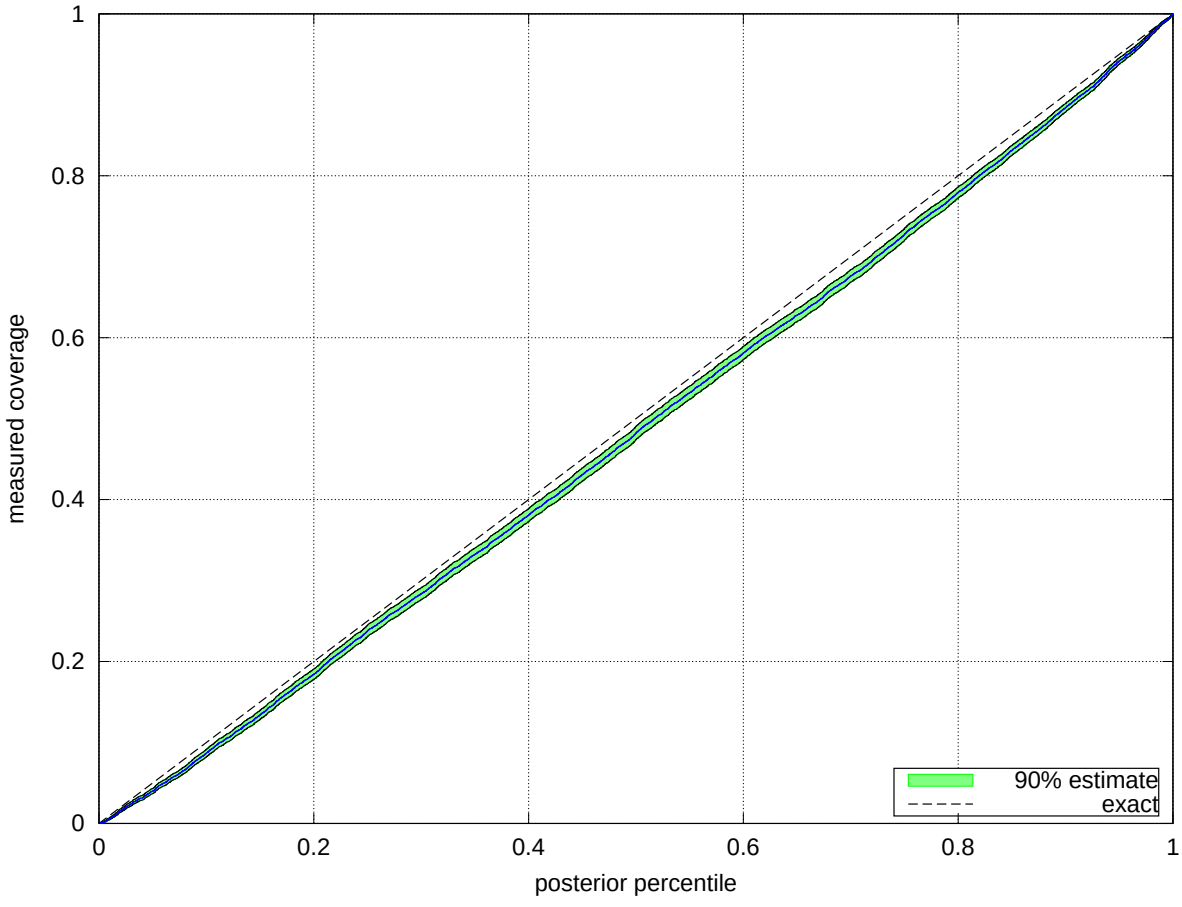
*2. Case 1(b): QNM parameter drawn from priors, **known** Gaussian white noise*



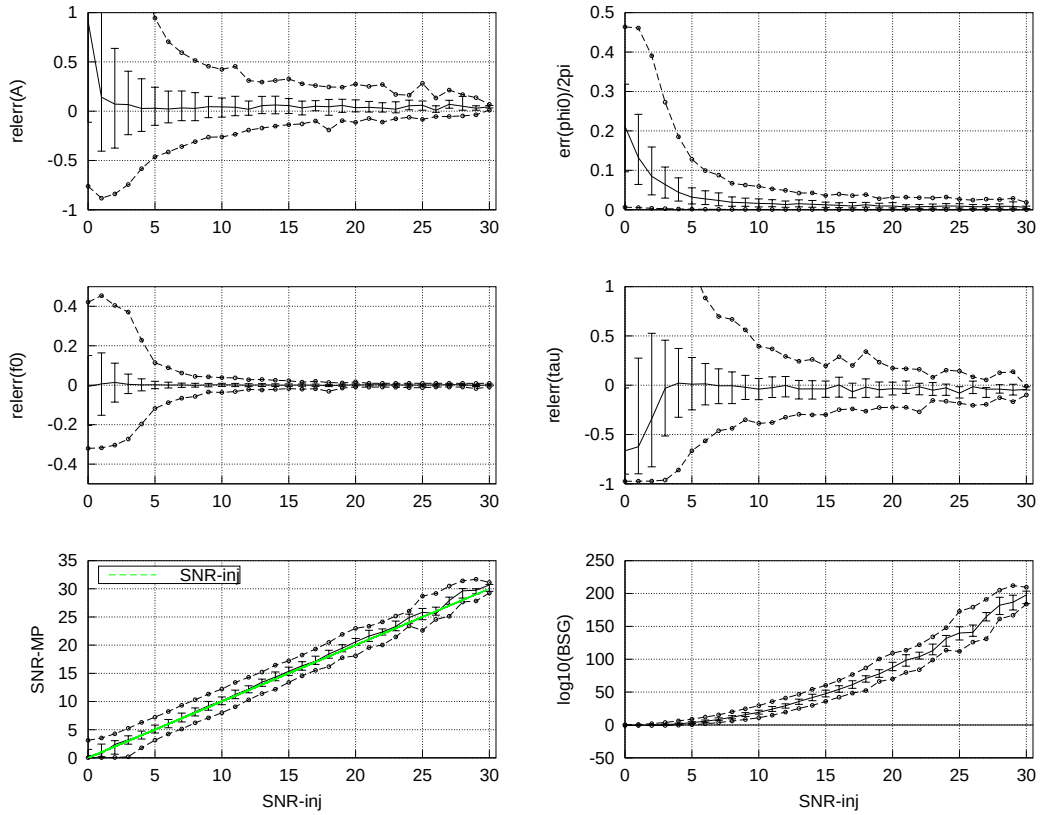


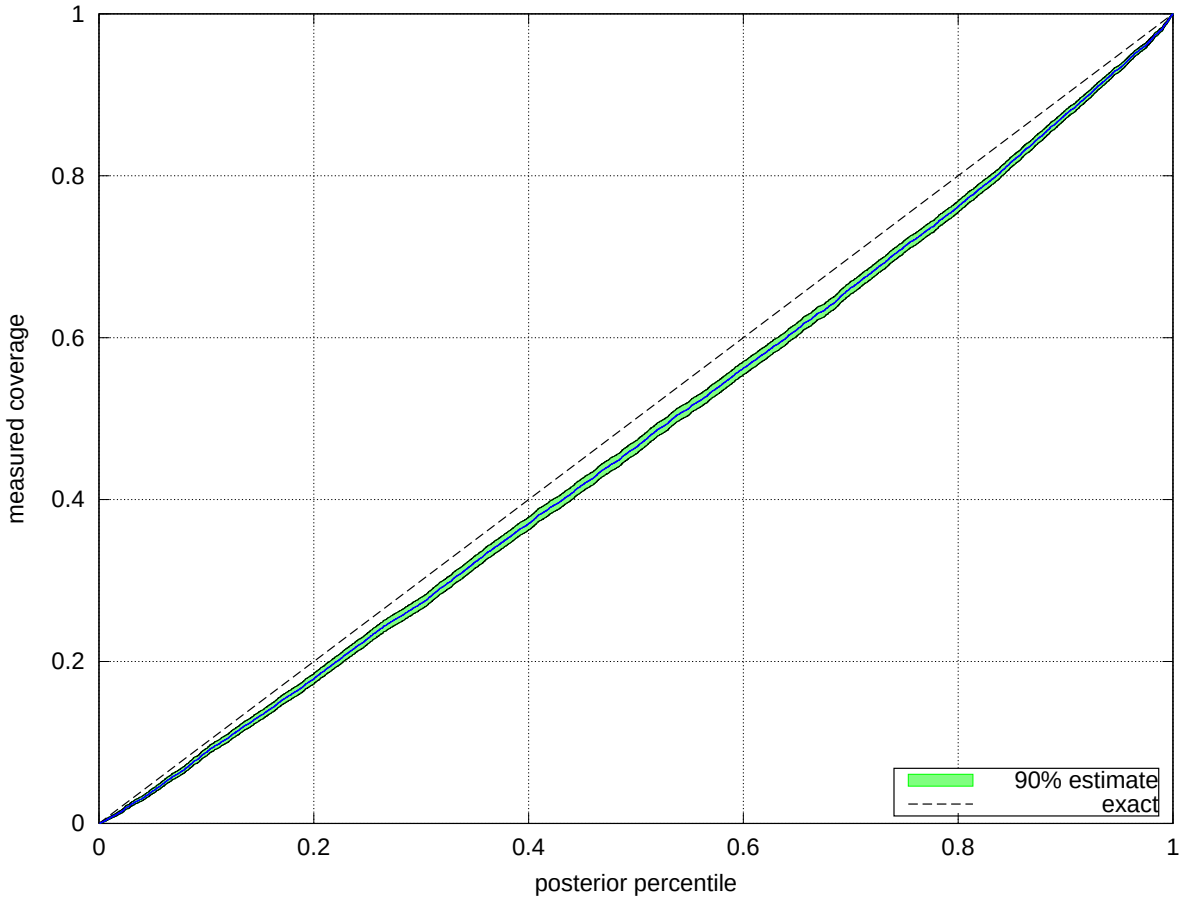
3. *Case 1(c): QNM parameter drawn from priors,* ***unknown*** *Gaussian white noise*



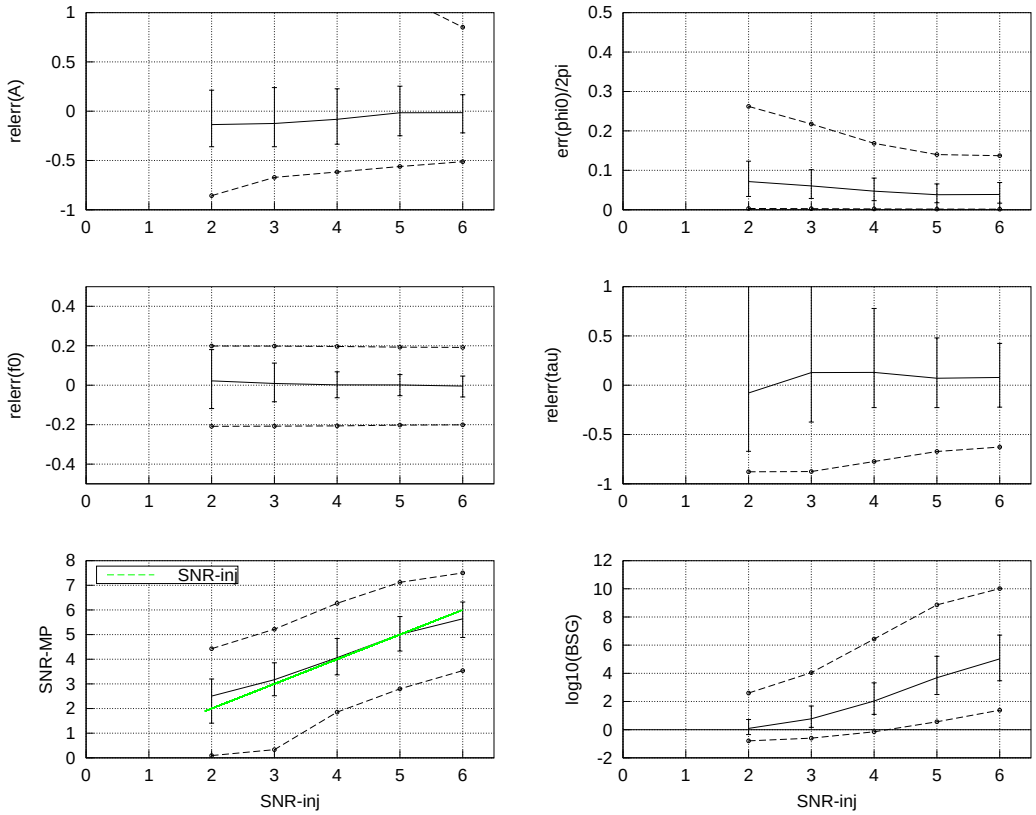


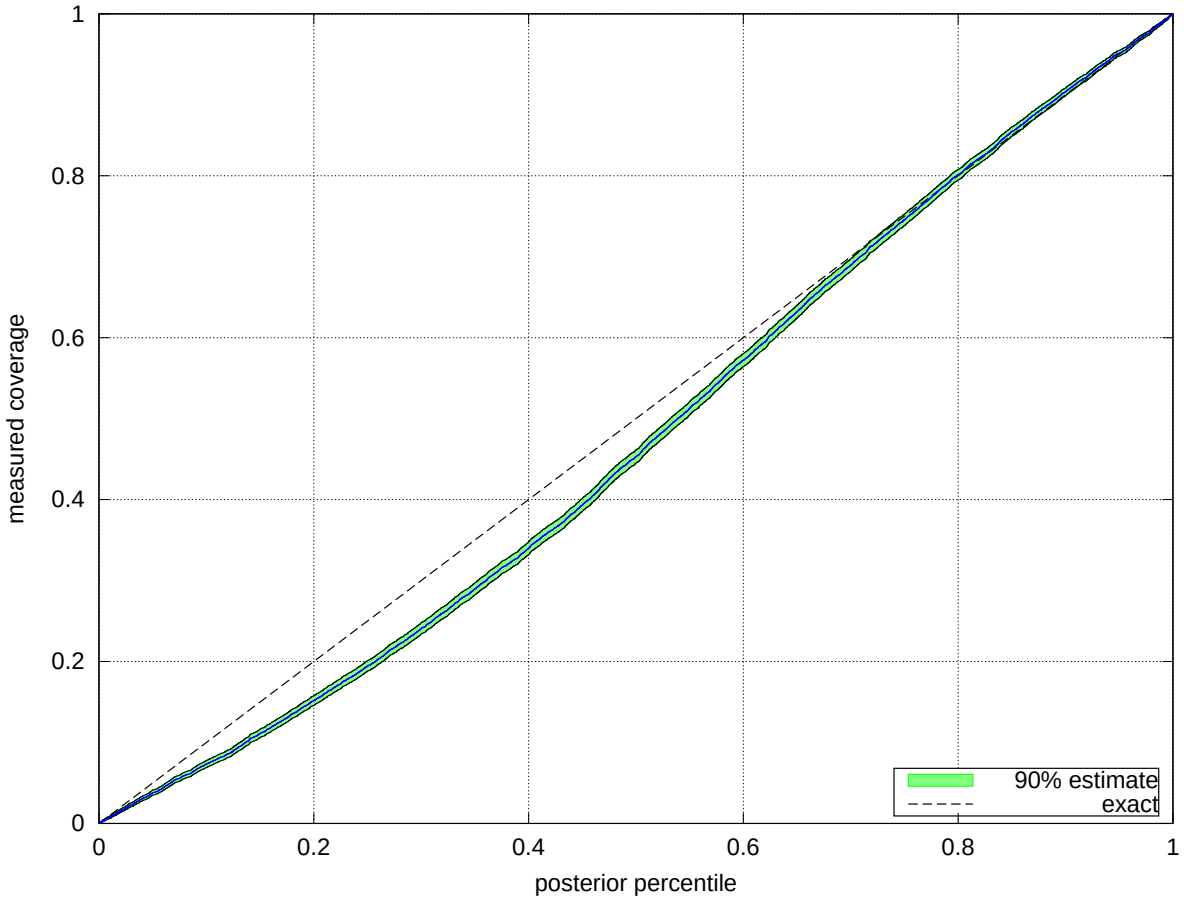
4. Case 1(d): QNM parameter drawn from priors, off-source detector data





5. *Case 2: fixed "GR injections"* $\{A \in [3,8] \times 10^{-22}, f_0 = 251\text{Hz}, \tau = 4\text{ms}\}$*, off-source detector data*





C. Varying noise-realization at fixed-$\{f_0 = 251\text{Hz}, \tau = 4\text{ms}, A = 2.5 \times 10^{21}, \phi_0 = 0\}$ injection

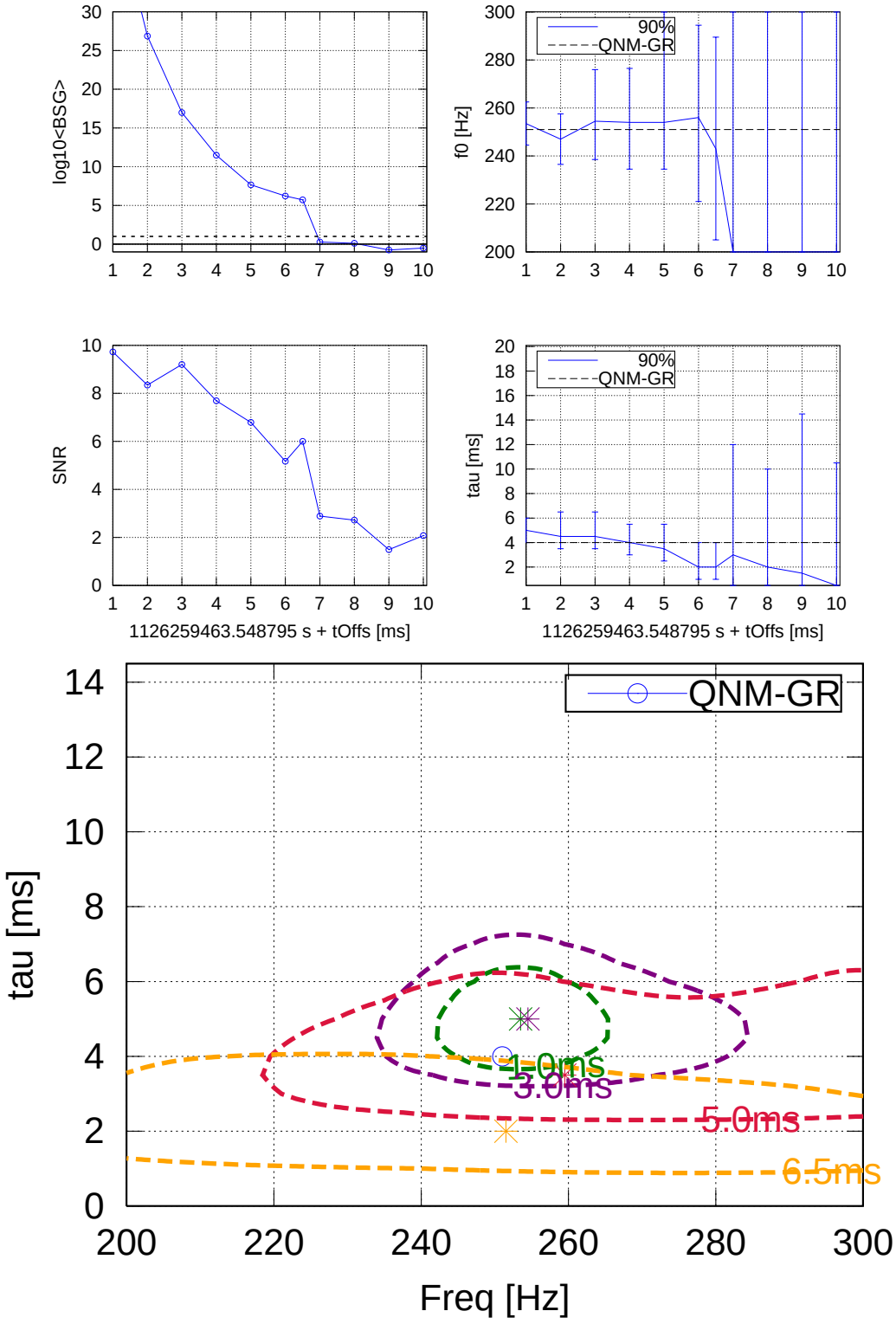
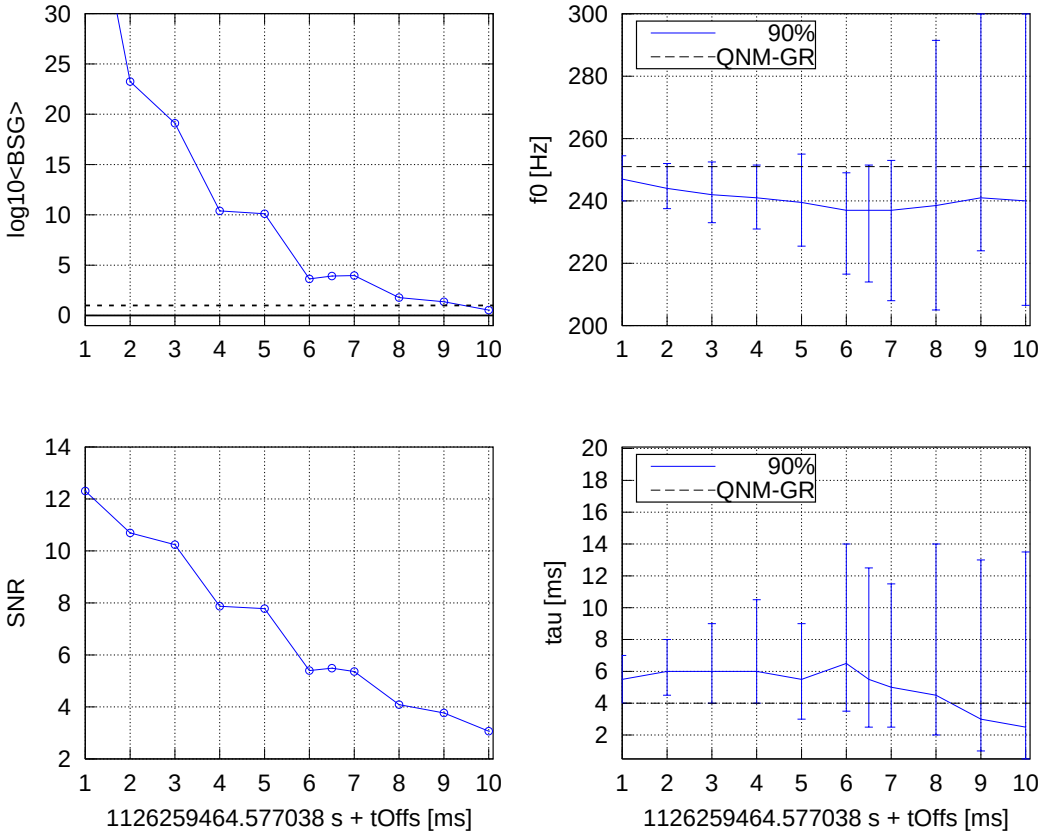


FIG. 10. Example 1 in Gaussian noise



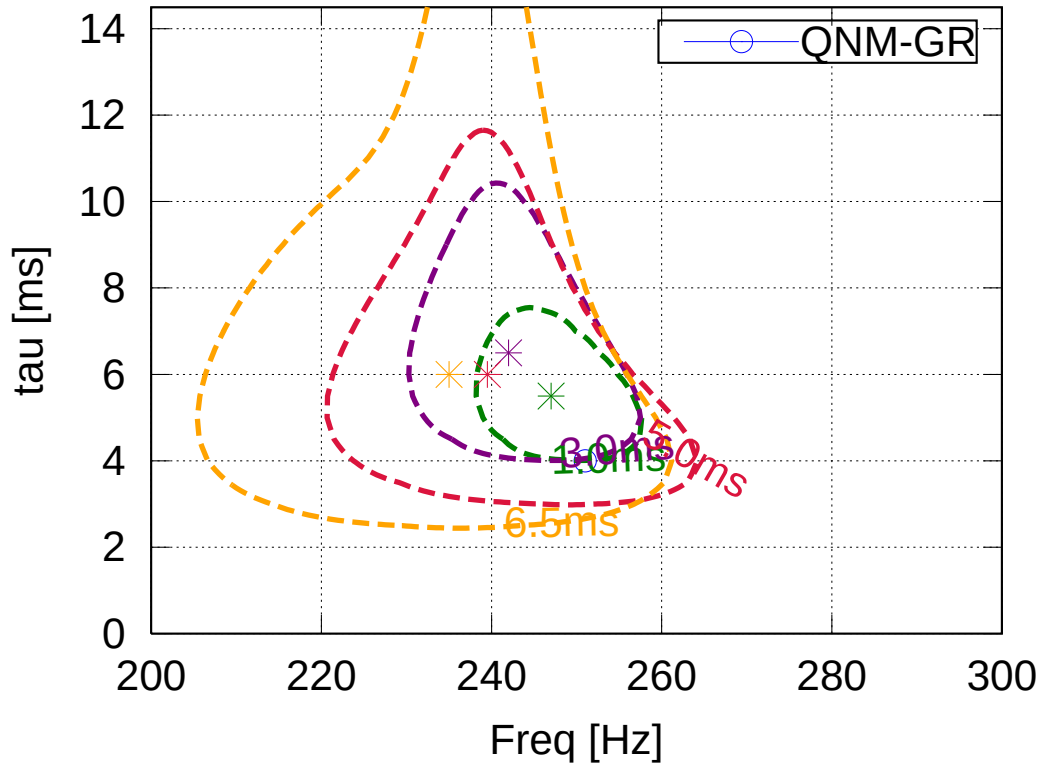


FIG. 11. Example 2 in Gaussian noise

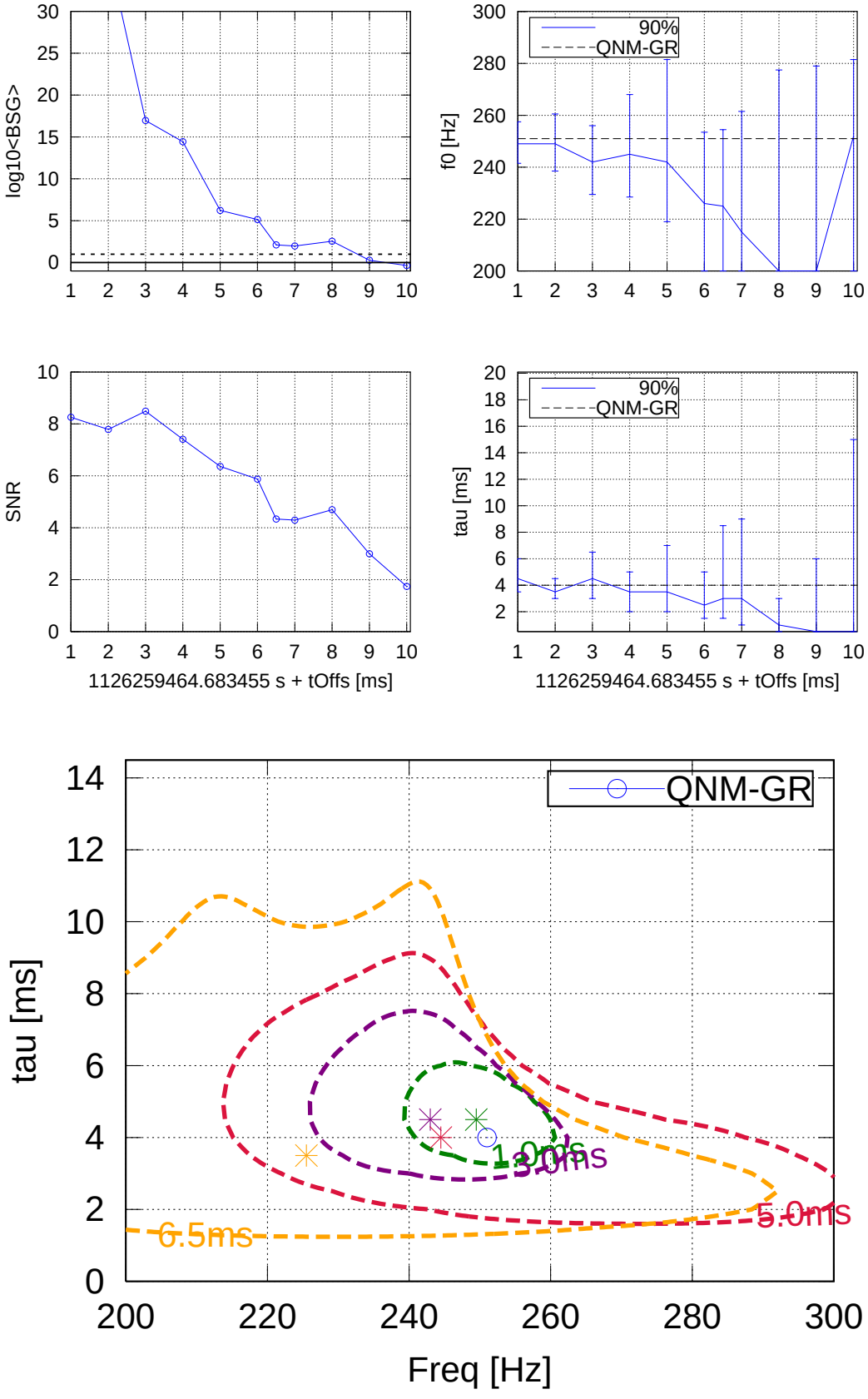


FIG. 12. Example 3 in Gaussian noise

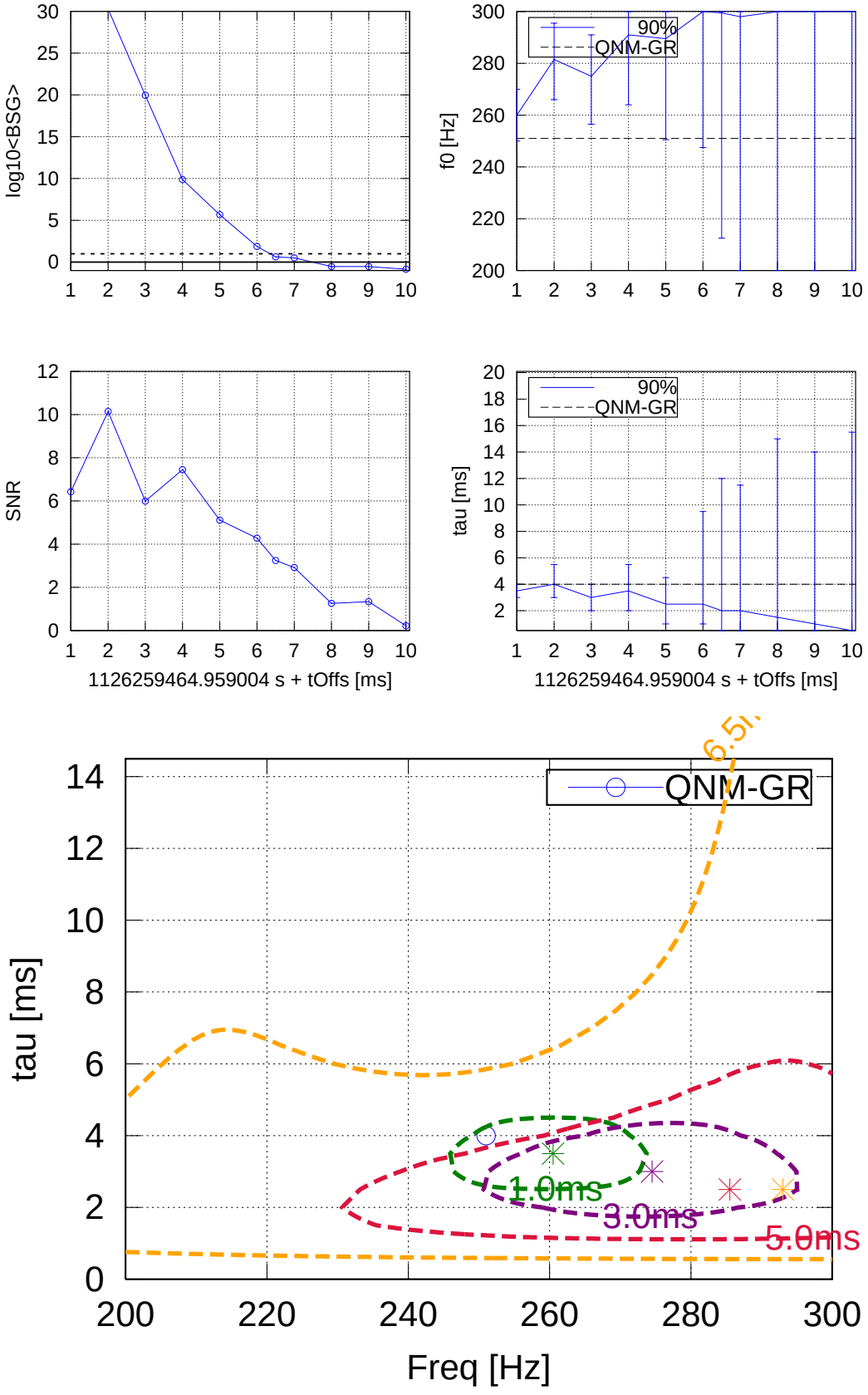


FIG. 13. Example 1 in real off-source data

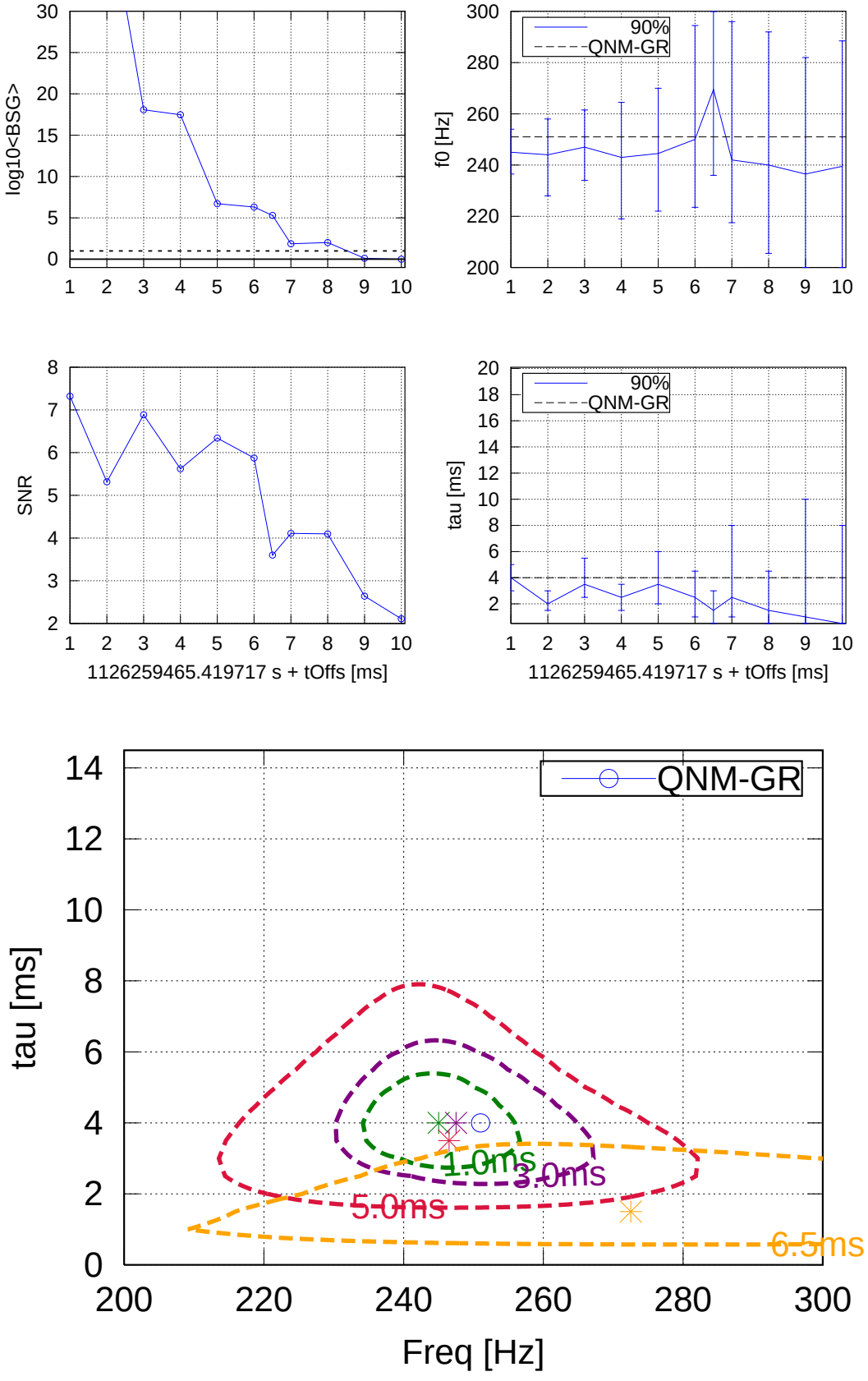


FIG. 14. Example 2 in real off-source data

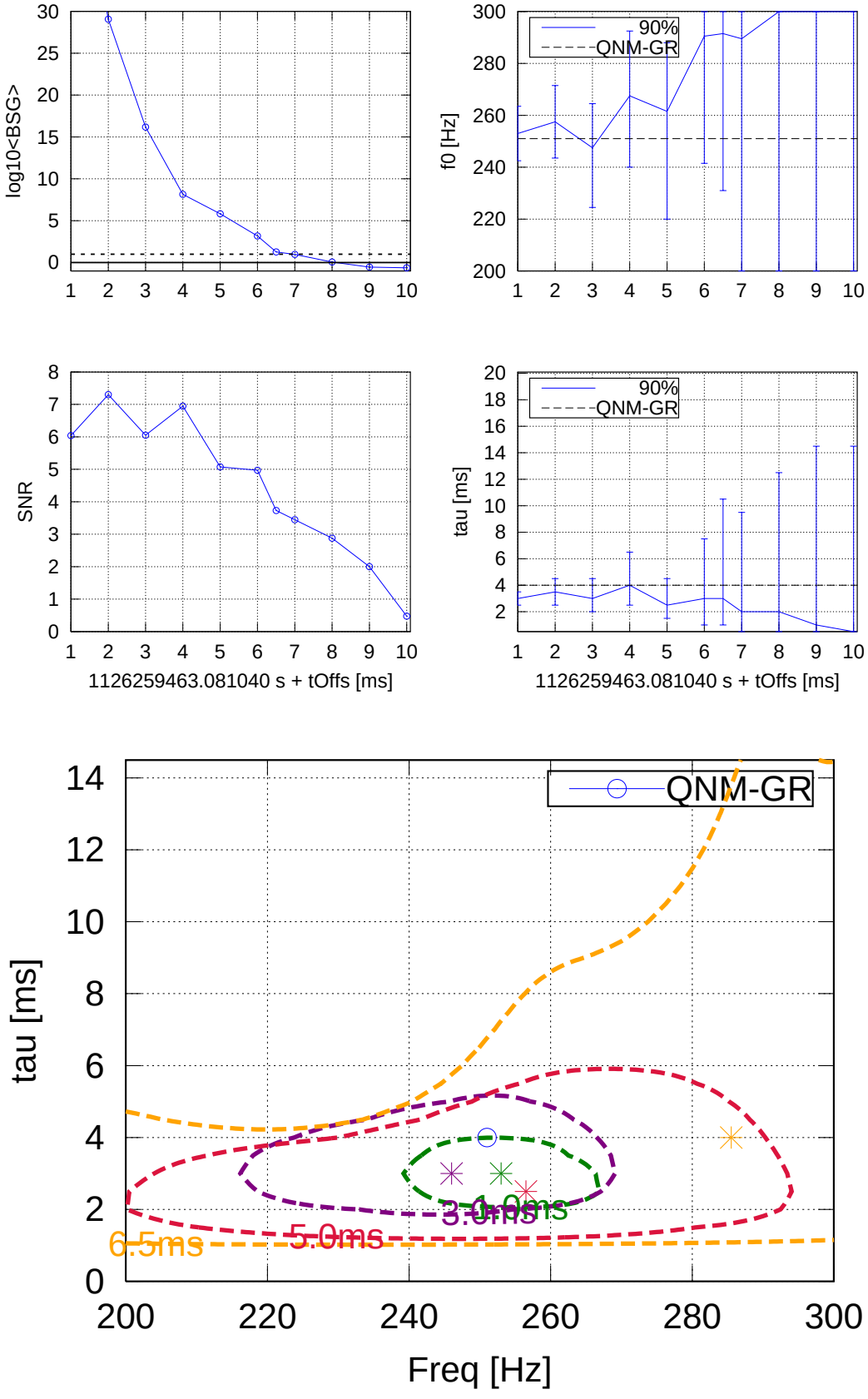


FIG. 15. Example 3 in real off-source data

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

$$\langle s|s\rangle = \sum_X 2\int_{-\infty}^{\infty} \frac{\widetilde{s}^X(f)\,\widetilde{s}^{*X}(f)}{S_X(f)}\,df \tag{A1}$$

$$\sim \sum_X \frac{2}{S_X(f')}\int_0^{\infty} s^2(t;\mathcal{A},\lambda)\,dt \tag{A2}$$

$$= \frac{2\,N_{\rm det}}{\mathcal{S}(f')}\int \left(\mathcal{A}_{\rm s}^2\,h_{\rm s}^2(t) + 2\mathcal{A}_{\rm s}\mathcal{A}_{\rm c}\,h_{\rm s}h_{\rm c} + \mathcal{A}_{\rm c}^2\,h_{\rm c}^2\right)\,dt \tag{A3}$$

$$= \mathcal{A}\cdot\mathcal{M}\cdot\mathcal{A}\,, \tag{A4}$$

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

$$I_{\rm s} \equiv \frac{1}{\mathcal{S}(f')}\int_0^{\infty} e^{-\frac{2t}{\tau}}\,\sin^2(2\pi f t)\,dt = \frac{1}{2\pi f}\int e^{-\frac{\varphi}{Q}}\sin^2\varphi\,d\varphi \tag{A6}$$

$$I_{\rm c} \equiv \int_0^{\infty} e^{-\frac{2t}{\tau}}\,\cos^2(2\pi f t)\,dt = \frac{1}{2\pi f}\int e^{-\frac{\varphi}{Q}}\cos^2\varphi\,d\varphi \tag{A7}$$

$$I_{\rm sc} \equiv \int_0^{\infty} e^{-\frac{2t}{\tau}}\,\sin(2\pi f t)\cos(2\pi f t)\,dt = \frac{1}{4\pi f}\int e^{-\frac{\varphi}{Q}}\sin 2\varphi\,d\varphi\,, \tag{A8}$$

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

$$\varphi \equiv 2\pi f\Delta t\,, \tag{A10}$$

$$Q \equiv \pi f\tau\,. \tag{A11}$$

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

$$I'_{\rm s} = \frac{-1}{2\pi f}\frac{Q^2}{1+4Q^2}e^{-\frac{\varphi}{Q}}\left[\sin 2\varphi + 2Q + \frac{\sin^2\varphi}{Q}\right]\Big|_0^{\infty} = \frac{2Q}{2\pi f}\,\frac{Q^2}{1+4Q^2} = \frac{\tau}{4+Q^{-2}} \tag{A12}$$

$$\overset{Q\gg 1}{\approx} \frac{\tau}{4}\,, \tag{A13}$$

$$I'_{\rm c} = \frac{-1}{2\pi f}\frac{Q^2}{1+4Q^2}e^{-\frac{\varphi}{Q}}\left[-\sin 2\varphi + 2Q + \frac{\cos^2\varphi}{Q}\right]\Big|_0^{\infty} = \frac{2Q+\frac{1}{Q}}{2\pi f}\,\frac{Q^2}{1+4Q^2} = \frac{\tau}{4}\left(\frac{2+Q^{-2}}{2+Q^{-2}/2}\right) \tag{A14}$$

$$\overset{Q\gg 1}{\approx} \frac{\tau}{4}\,, \tag{A15}$$

$$I'_{\rm sc} = \frac{-1}{2\pi f}\frac{Q^2}{1+4Q^2}e^{-\frac{\varphi}{Q}}\left[2\cos^2\varphi - 1 + \frac{\sin 2\varphi}{2Q}\right]\Big|_0^{\infty} = \frac{1}{2\pi f}\,\frac{Q^2}{1+4Q^2} = \frac{\tau}{8Q\,(1+Q^{-2}/4)} \tag{A16}$$

$$\overset{Q\gg 1}{\approx} \frac{1}{2Q}\,I_{\rm s} \ll I_{\rm s} \approx 0\,. \tag{A17}$$

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.