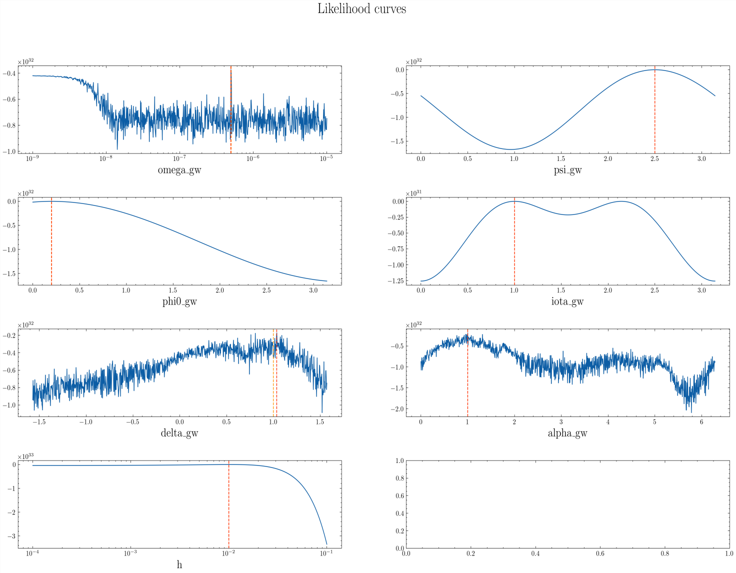
EE/GW meeting, May 19, 2023

Last time: Problems at small h, nasty likelihood curves

All likelihood curves for $\bar{\theta}_{\mathrm{GW}}$:



Drop PSR terms from model?

$$f_{\mathrm{measured}} = f_{\mathrm{emitted}}\, g(\tau; \bar{\theta}) \tag{1}$$

with

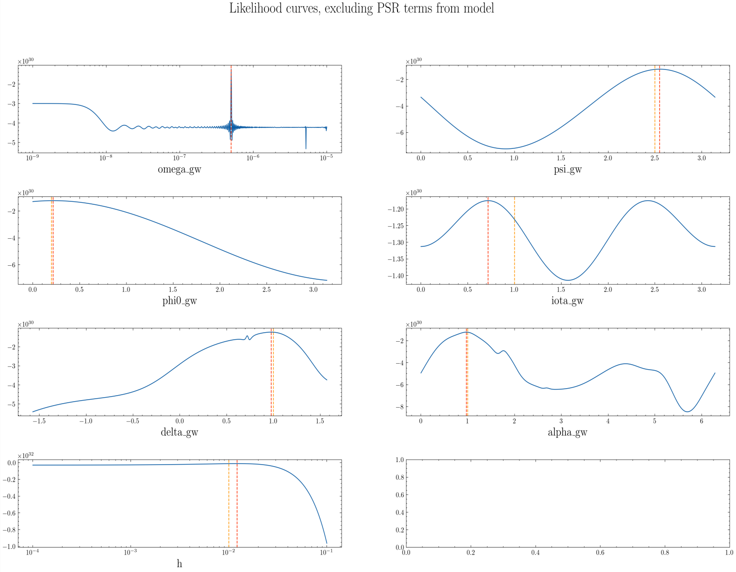
$$g(\tau; \bar{\theta}) = 1 - \frac{1}{2} \frac{H_{ij} q^i q^j}{(1 + \bar{n} \cdot \bar{q})} \left[\cos(-\Omega\tau + \Phi_0) - \cos(-\Omega\tau + \Phi_0 + \Omega(1 + \bar{n} \cdot \bar{q}) d) \right] \tag{2}$$

Generate the fake data using the full expression for $g(\tau; \bar{\theta})$, but use $g'(\tau; \bar{\theta})$ in the Kalman measurement model

$$g'(\tau; \bar{\theta}) = 1 - \frac{1}{2} \frac{H_{ij} q^i q^j}{(1 + \bar{n} \cdot \bar{q})} \left[\cos(-\Omega\tau + \Phi_0) \right] \tag{3}$$

Drop PSR terms from model?

All likelihood curves for $\bar{\theta}_{\mathrm{GW}}$:



Full NS search for all parameters

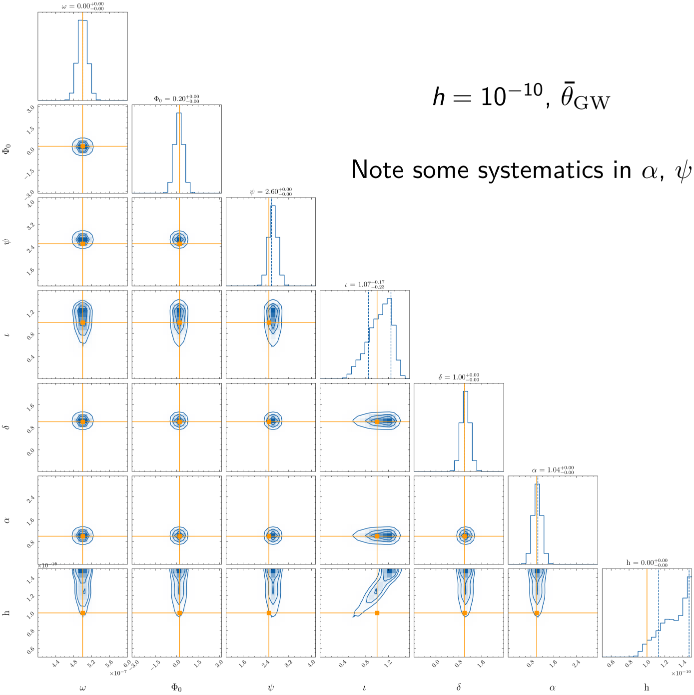
$\bar{\theta} = [\bar{\theta}_{\mathrm{GW}},\, f_0,\, \dot{f}_0]$

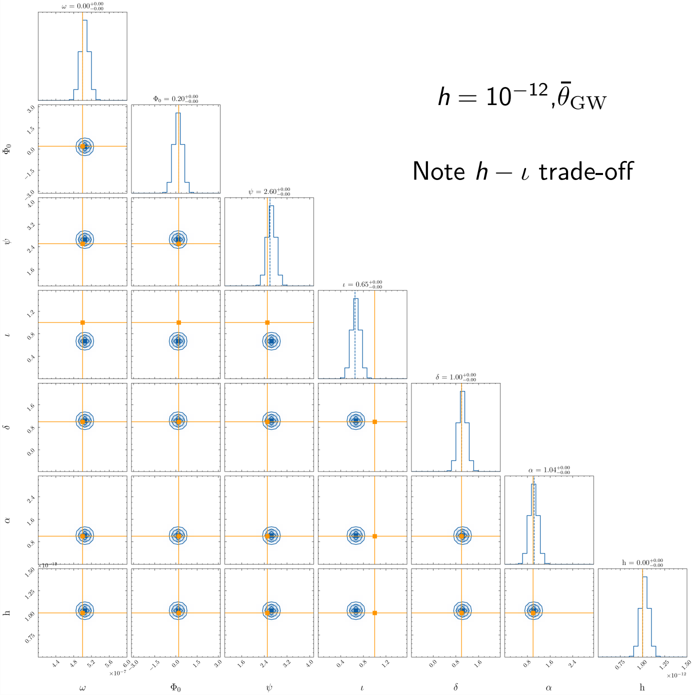
with $h = [10^{-10}, 10^{-12}]$, $\sigma_m = 10^{-11}$

Note: pulsar distance is no longer part of model. γ doesn't matter

$h = 10^{-10}$, $\bar{\theta}_{\mathrm{GW}}$

Note some systematics in α, ψ

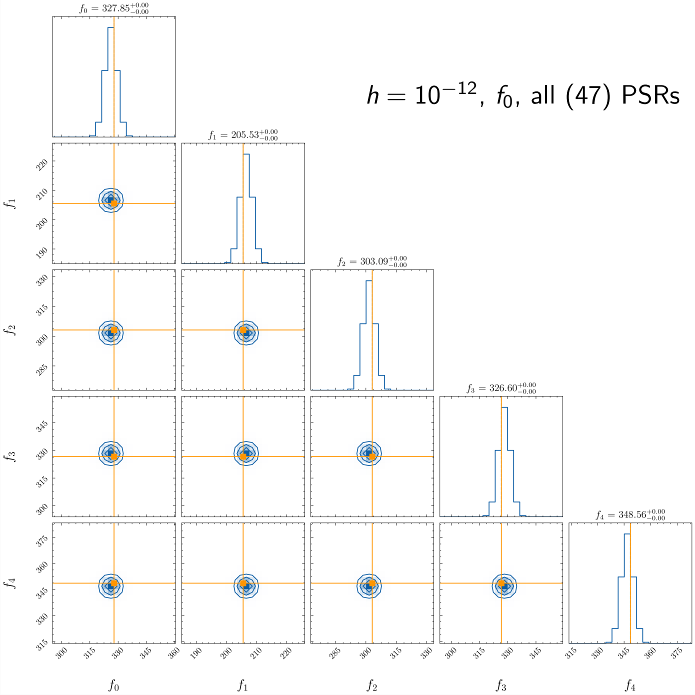




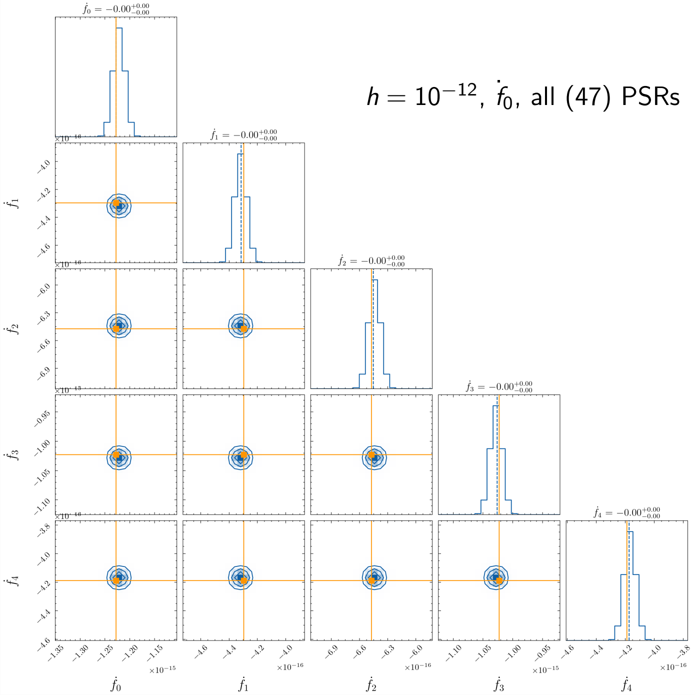
$$h = 10^{-12}, \bar{\theta}_{\rm GW}$$

Note $h - \iota$ trade-off

$h = 10^{-12}$, f_0, all (47) PSRs



$h = 10^{-12}$, $\dot{f}_0$, all (47) PSRs



Next steps:

- Bayes factors vs. strain (currently running)
- Process noise σ_p (next slides)
- $[h, \iota] \rightarrow [h_+, h_\times]$ (discussed, easy, helpful?)
- Can we add PSR terms back in? Do we need to?
- Robustness for different $\bar{\theta}_{\mathrm{GW}}$ (overkill? canonical example sufficient?)

Setting the process noise

Timing noise model of Shannon & Cordes, 2010

$$\ln \sigma_p^{\mathrm{TOA}}[\mu s] = \ln C + \alpha \ln \nu[s^{-1}] + \beta \ln \dot{\nu}[10^{-15} s^{-2}] + \gamma \ln T[years] \quad (4)$$

For MSP, $\ln C \sim -20$, $\alpha \sim 1$, $\beta \sim 2$, $\gamma \sim 2.4$

Taking $T = 1$ week, for our synthetic NANOGrav pulsars this puts:

$$\sigma_p^{\mathrm{TOA}}[s] \sim \left(10^{-19}, 10^{-17}, 10^{-13}\right), (\text{min/median/max})$$

Relate to frequency:

$$\sigma_p^{\mathrm{f}}[Hz] = f \frac{\sigma_p^{\mathrm{TOA}}}{T} \sim \left(10^{-23}, 10^{-21}, 10^{-16}\right), (\text{min/median/max})$$

Could hit float epsilon issues here... may need to scale state evolution equations, similar to heterodyning the measured data...

Rescaling to avoid float issues

State evolution equation (Ornstein–Uhlenbeck)

$$\frac{df}{dt} = -\gamma[f - f_{\mathrm{EM}}(t)] + \dot{f}_{\mathrm{EM}} + \xi(t; \sigma_p) \tag{5}$$

i.e.

$$df = a(f, t)dt + \sigma_p dW \tag{6}$$

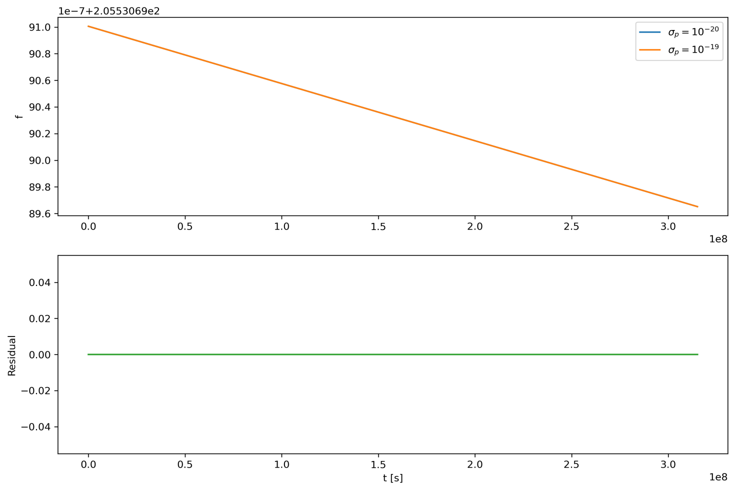
Can solve numerically via e.g. Euler-Maruyama

Partition interval $[0, T]$ into N intervals, each width $\Delta t = T/N$

$$f_{n+1} = f_n + a(f_n, \tau_n)\Delta t + \sigma_p \Delta W_n \tag{7}$$

$$\mathcal{O}(10^2) = \mathcal{O}(10^2) + \mathcal{O}(10^{-15}) \cdot \Delta t + \mathcal{O}(10^{-20})\Delta W_n \tag{8}$$

Practically: synthetic data (float64) generated with e.g. $\sigma_p = 10^{-19} =$ data with e.g. $\sigma_p = 10^{-20}$


1e−7+2.0553069e2
91.0
90.8
90.6
90.4
90.2
90.0
89.8
89.6
f
$\sigma_p = 10^{-20}$
$\sigma_p = 10^{-19}$
0.0
0.5
1.0
1.5
2.0
2.5
3.0
1e8
0.04
0.02
0.00
−0.02
−0.04
Residual
t [s]

Let $f' = f - f_{\mathrm{EM}}$

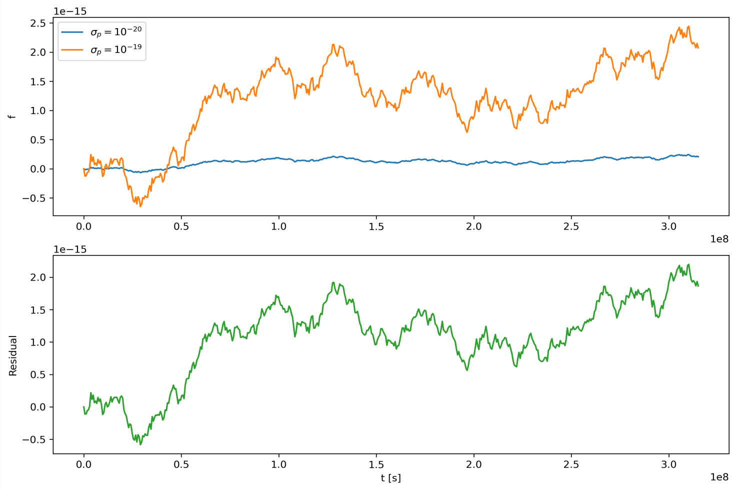
$$\frac{df'}{dt} = \frac{df}{dt} - \frac{df_{\mathrm{EM}}}{dt} \tag{9}$$
$$= \frac{df}{dt} - \frac{df_{\mathrm{EM}}}{dt} \tag{10}$$
$$= -\gamma f' + \xi \tag{11}$$

Now numerical solution looks like

$$\mathcal{O}(10^{-12}) = \mathcal{O}(10^{-13}) \cdot \mathcal{O}(10^{-12}) \cdot \Delta t + \mathcal{O}(10^{-20}) \tag{12}$$

$$\mathcal{O}(1) = \mathcal{O}(10^{-13}) \cdot \Delta t + \mathcal{O}(10^{-8}) \tag{13}$$

Practically: synthetic data (float64) generated with e.g. $\sigma_p = 10^{-19} \neq$ data with e.g. $\sigma_p = 10^{-20}$


1e−15
$\sigma_p = 10^{-20}$
$\sigma_p = 10^{-19}$
f
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
1e8
1e−15
Residual
2.0
1.5
1.0
0.5
0.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
t [s]
1e8

Heterodyne state and measurement?

- Heterodyne the state: $f' = f - f_{\mathrm{EM}}$
- Heterodyne the measurement: $f'_{\mathrm{M}} = f_{\mathrm{M}} - f^*_{\mathrm{EM}}$

f_{EM} is the exact, true solution parametrised by the unknowns $f_0, \dot{f}_0$

f^*_{EM} is a guess based on the pulsar ephemeris.

For our case where we have synthetic data we can set $f_{\mathrm{EM}} = f^*_{\mathrm{EM}}$, but obviously not true generally - we are trying to estimate f_{EM}!

We can then write the measurement equation as

$$f_{\mathrm{M}} = (1 - X) f + N_{\mathrm{G}} \tag{14}$$

$$f'_{\mathrm{M}} = (1 - X) f' - X f_{\mathrm{EM}} + (f_{\mathrm{EM}} - f^*_{\mathrm{EM}}) + N_{\mathrm{G}} \tag{15}$$

- Linear ✓(control vector has shifted from state equation to measurement equation)
- Scales: $\mathcal{O}(h) = \mathcal{O}(10^{-12}) - \mathcal{O}(h) \cdot \mathcal{O}(10^{2}) + \mathcal{O}(0) + \mathcal{O}(10^{-11})$ ✓

Insert corner plot here

Backup slides

Where did σ_m value come from?

The measurement noise is

$$\sigma_m = f \frac{\sigma_{\mathrm{TOA}}}{\mathrm{cadence}}$$

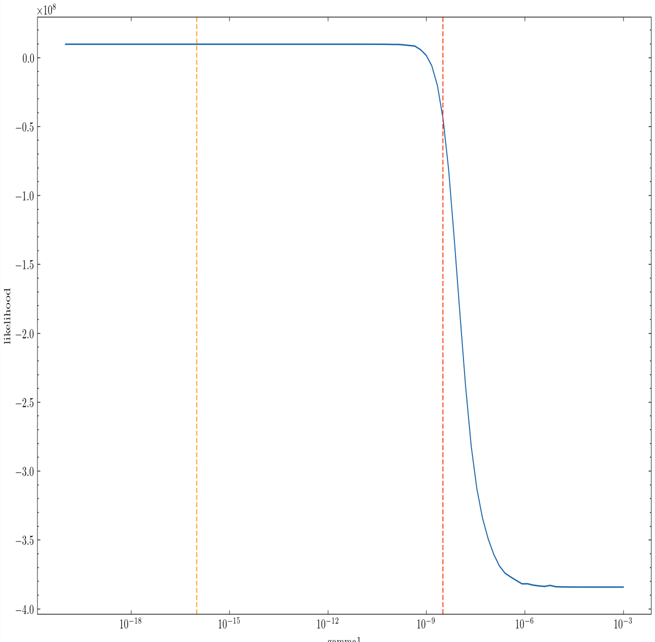
So for a MSP ($f \sim 100 Hz$) observed with a weekly cadence and $\sigma_{\mathrm{TOA}} \sim 1 \mu$s:

$$\sigma_m \sim 1.6 \times 10^{-10}$$

The very best pulsars might have $\sigma_{\mathrm{TOA}} \sim 10$ ns:

$$\sigma_m \sim 1.6 \times 10^{-12}$$

γ likelihood



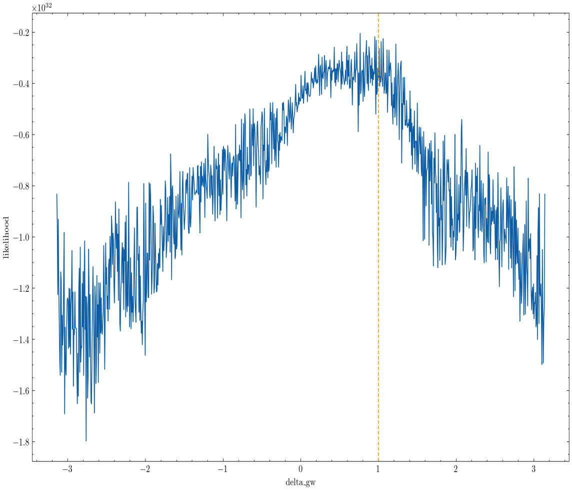
What changed?

AM: *"What changed such that now everything works?"*

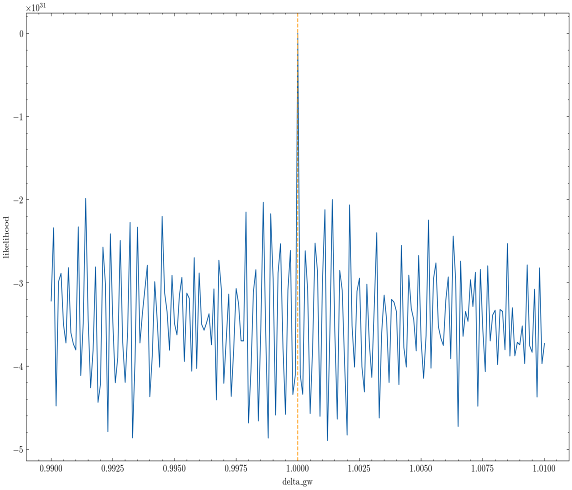
- Dropping PSR terms to smooth likelihoods
- *Increasing* measurement noise
- Don't need too many live points
  - $n_{live} >$ number of parameters
  - $n_{live} \times \frac{\text{mode volume}}{\text{prior volume}} > 1$ i.e. at least 1 live point per mode
  - Runtime scales $\mathcal{O}(n_{live})$
  - Posterior/Evidence uncertainties scale as $\mathcal{O}(1/\sqrt{n_{live}})$
- Non-default Bilby sampler settings
  - e.g. docs recommend 'sample=act-walk'or 'sample=acceptance-walk'
  - More success with 'sample=rwalk-dynesty' (not 'sample=rwalk' which is the Bilby implementation)
  - bound = multi(dynesty), not the Bilby version
- Be mindful of float errors - Python/Numpy/Bilby NaNs get propagated, rather than error e.g. Fortran.

More likelihood curves

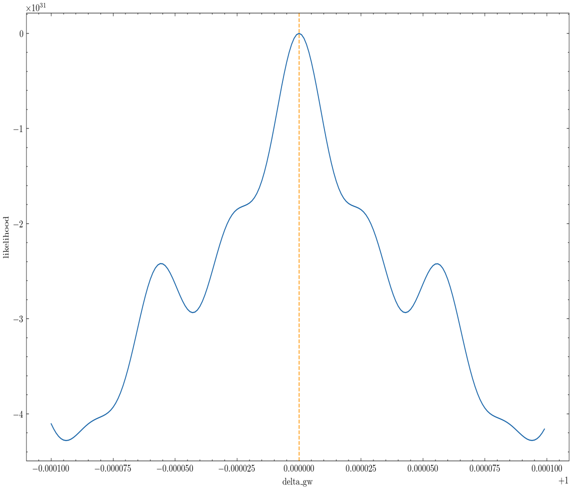
Example: $\mathcal{L}(\delta)$



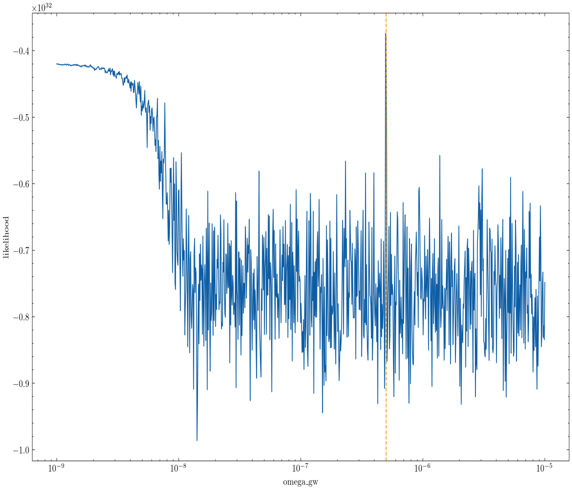
Example: $\mathcal{L}(\delta)$, zoomed 1



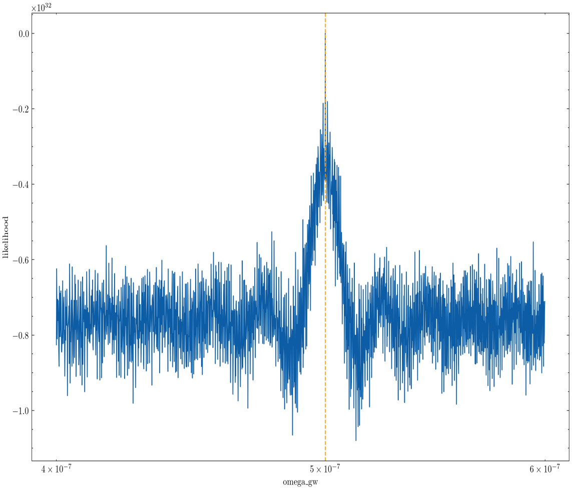
Example: $\mathcal{L}(\delta)$, zoomed 2



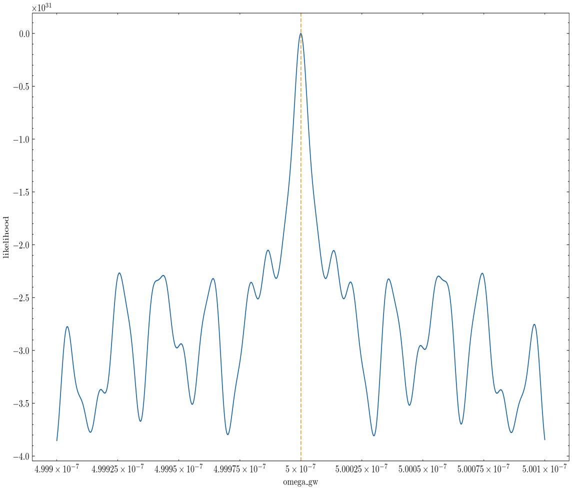
Example: $\mathcal{L}(\omega)$



Example: $\mathcal{L}(\omega)$, zoomed 1



Example: $\mathcal{L}(\omega)$, zoomed 2



- Problem = overly narrow,biased posteriors
- Cause: noisy, likelihoods which are "locally Gaussian". Sampling gets stuck in a local optima and infers posteriors widths based on that optima.

Potential solutions to noisy likelihoods

1. Better settings for sampler? E.g. more live points (how many?), massively parallel, optimized likelihood evaluations,...
2. Maximum likelihood / loss optimisation?
3. Drop PSR terms from model?
4. MCMC at higher temperatures (trades off evidence..., less efficient)
5. other...?