Follow up a signal hidden into noise: the challenge of all-sky continuous gravitational wave searches

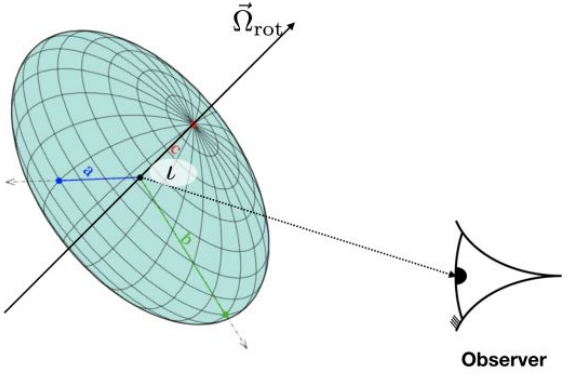
GEMMA 2 conference, Roma 16-19 Sept. 2024

Lorenzo Mirasola in collaboration with
P. Astone, S. D'Antonio, A. De Falco, C. Palomba, R. Tenorio

The continuous wave signal

Non axisymmetric rotating neutron stars as CW emitters

Figure taken from S. Mastrogiovanni PhD thesis

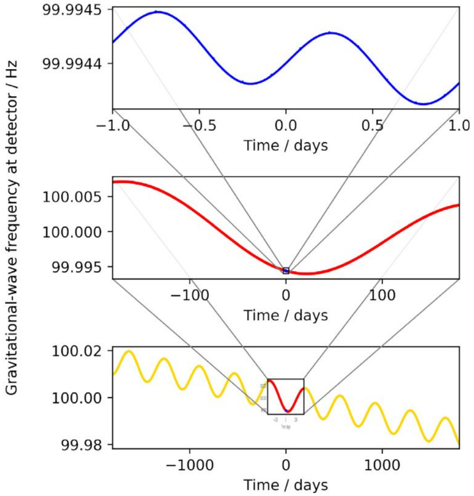


Doppler modulation due to earth rotation

Doppler modulation due to earth revolution

Frequency decrease due to spin-down effect

(c) Karl Wette CC BY-SA



CW frequency linked to rotational one

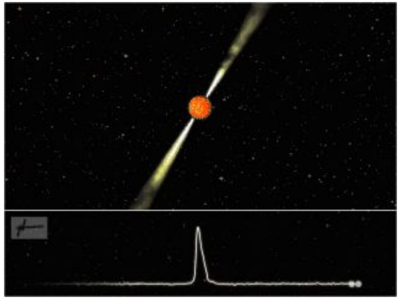
$$f_{GW}(t) = 2 \cdot f_{rot}(t)$$

Quasi-monochromatic signal **in the source frame**

Possible additional Doppler modulation due to binary motion

Continuous wave searches

Binary orbital parameters can significantly increase computational cost if not well constrained



Rotational and orbital parameters inferred through electromagnetic observations

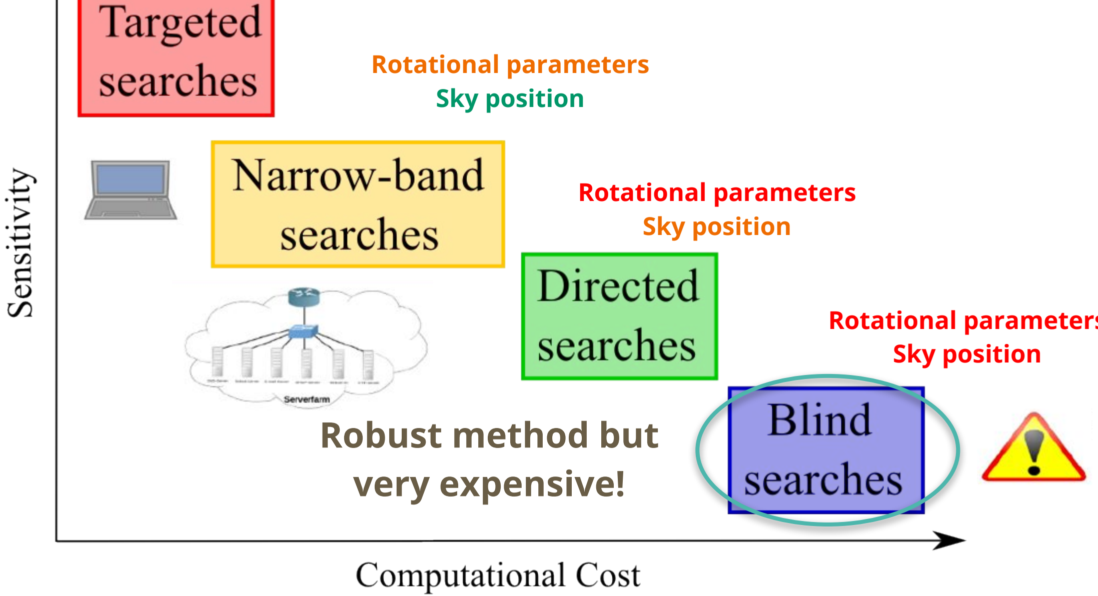


Figure taken from Sieniawska & Bejger (2019)

All-sky searches

Not all neutron stars can be seen with telescopes

↓

how can we look for them?

↓

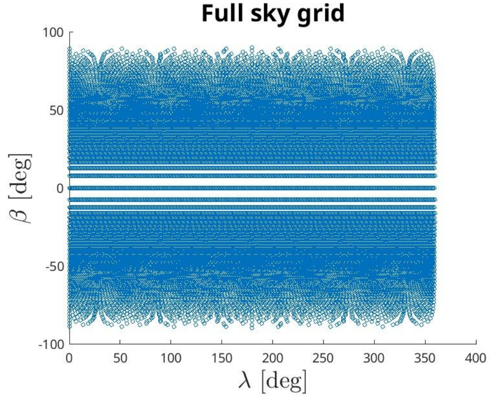
need to probe all the parameter's combinations

↓

each combination must be studied for all detectors

↓

Impressive computational cost



sky position (2 angles)

frequency

frequency time derivative

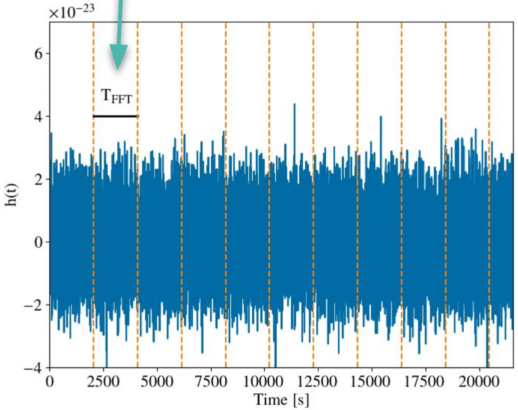
+ binary parameters

↓

Parameters' resolution linked to observing time/coherence time

Needed semi-coherent approaches to lower computational cost

Coherence time → Chunk of data analysed coherently → e.g. with FFT



Information from single chunks are combined to construct a detection statistic

See e.g. Wette AP 153 102880 (2023) for a review

All-sky searches (2)

Outliers are identified following selection criteria (threshold on the detection statistic, clustering, coincidences, …)

Injection campaign

Semi-analytical approach

Fixing false-alarm and false-dismissal probabilities

Typically $O(10^{8\text{-}9})$ outliers are identified

Selection of most significant ones

$O(10^{5\text{-}6})$ selected for the follow-up stage

How to deal with that many outliers?

How to efficiently discard noise outliers?

How can we achieve our first detection?

The crucial role of follow-ups

$$h_{\text{sens}} \propto (T_{\text{obs}}\, T_{\text{FFT}})^{-1/4}$$

Improve signal's significance → longer observing runs / longer chunks

longer chunks → where we can play!

Longer coherence time → Finer parameter space grid / Narrower signal peak

Finer parameter space grid → Too many templates?

Narrower signal peak → Too narrow to be detected?

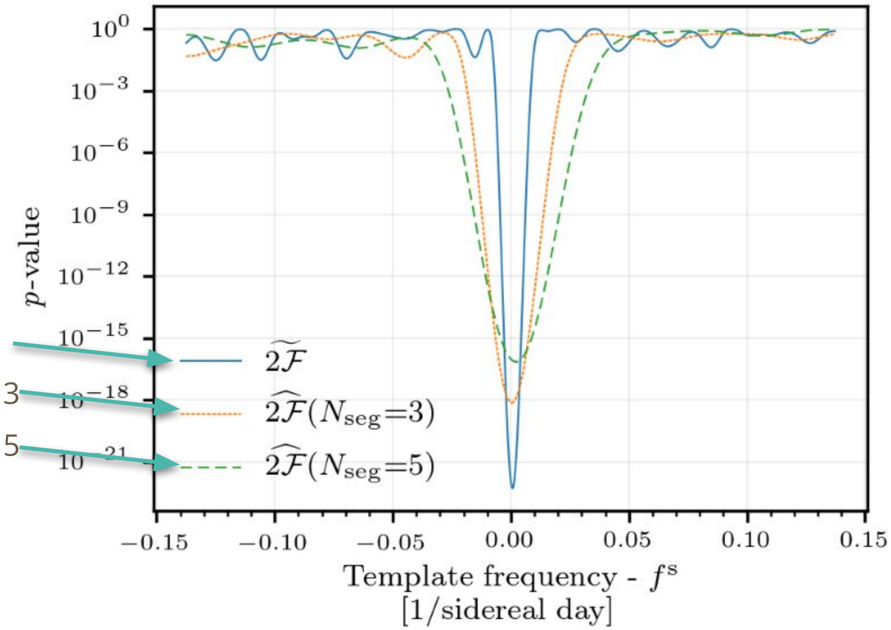
Good follow-up procedure → **Cheap** / **High detection efficiency**

Figure taken from Ashton & Prix PRD 97, 103020 (2018)

$T_{FFT} = T_{OBS}$

$T_{FFT} = T_{OBS} / 3$

$T_{FFT} = T_{OBS} / 5$



How can we follow-up an outlier?

Gridded approach...

All-sky previous uncertainty on sky location

Increase coherence time → Typically 2-3 times the old one

Finer resolution

$$\delta f = \frac{1}{T_{FFT}}$$

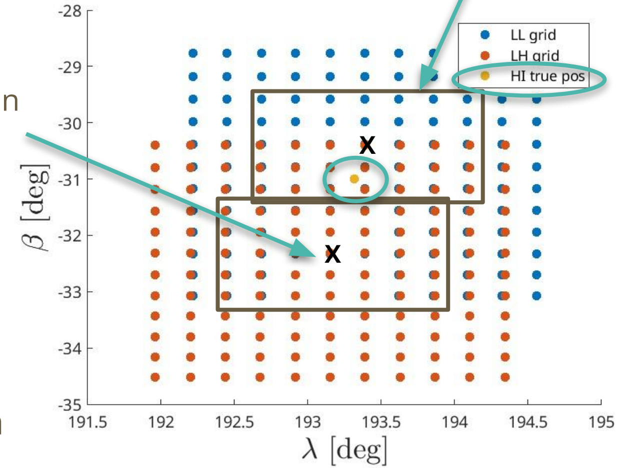
e.g. for frequency

"Brute-force" approach

Explore a "small" region around the outlier's parameter

- sky position (2 angles)
- frequency
- frequency time derivative
- + binary parameters

Starting points given by all-sky search



How large?

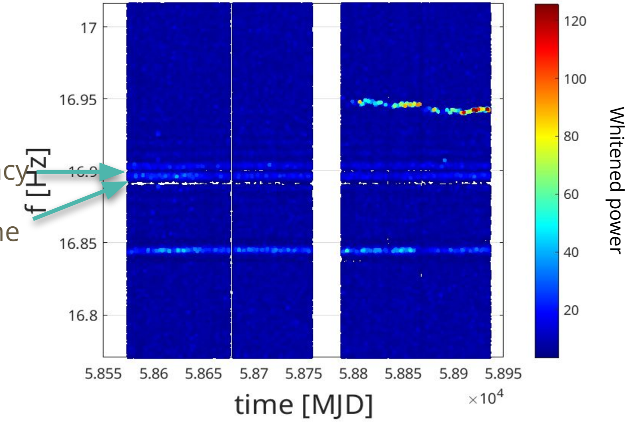
Depends on all-sky search parameter estimation

Injection campaign

See Abbott & Al. PRD 106, 102008 (2022) and references therein for a brief summary of all-sky searches and their follow-up methods

Gridded approach (2)…

Detectors' data are analysed independently → Identify template with the highest significance → Has the significance increased?
Compatible with noise lines?
Is the outlier present in both detectors?
…



Do we still have an interesting candidate?

Pros:
Explores systematically each point of the grid → **Efficient with signals close to the threshold**

Cons:
Explored parameter space can depend on computing power availability

See Abbott & Al. PRD 106, 102008 (2022) and references therein for a brief summary of all-sky searches and their follow-up methods

...vs Markov Chain Monte Carlo approach

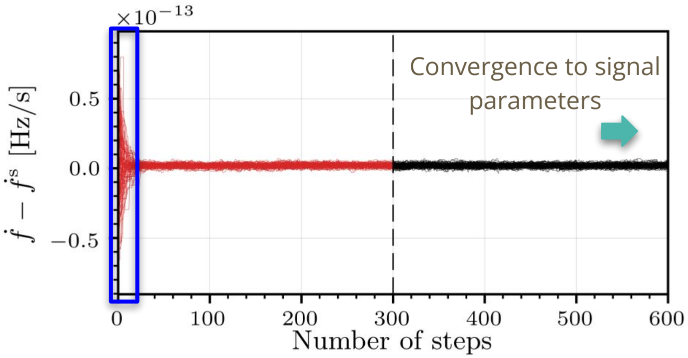
Likelihood maximisation with random approach

↓

Start simultaneously from N_w random positions

↓

Let the system evolve according to the likelihood



Input: candidate

↓

Increase coherence time

↓

Map the likelihood

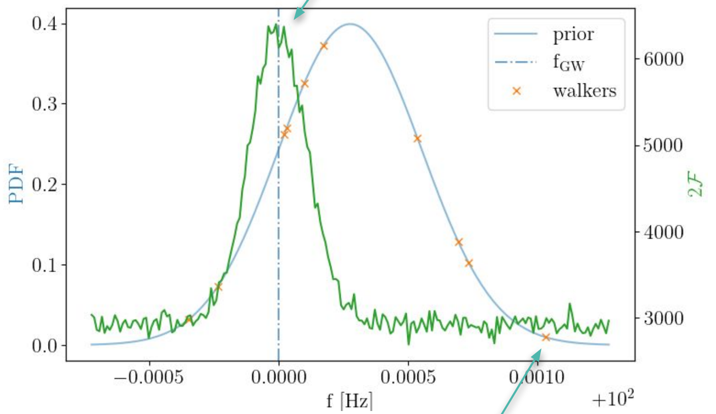
↓

Identify template with the highest significance

↓

Has the significance increased?
Compatible with noise lines?
...

MCMC result



Initial random sampling of the prior

See Ashton & Prix PRD 97, 103020 (2018) and Tenorio+ PRD 104, 084012 (2021)

...vs Markov Chain Monte Carlo approach (2)

Focus on the open package Pyfstat

F-statistic + ptemcee sampler

3 hyperparameters to be tuned

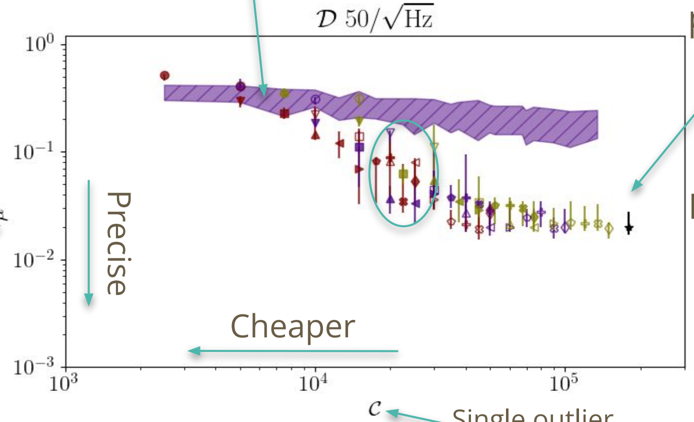
# of walkers	# of temps	# of steps
N_w	N_t	N_{tot}

Cost per outlier can be reduced significantly by choosing the appropriate setup

$$\mathcal{D} = \frac{\sqrt{S_n}}{h_0}$$

Threshold for $p_{fd} = 10^{-8}$
$p_{fa} = 10^{-1}$–10^{-2}

Filled/Unfilled = N_w
Colour = N_t
Shape = N_{step}



SNR fractional loss

Single outlier computational cost

Configuration of past searches

Isolated source O3 LL+LH data

First stage usually increases coherence time by **10 times**

Pro:
Can manage large parameter-space volumes

Cons:
Narrow signal peaks might be lost

See Mirasola & Tenorio arxiv (2024)

Impact on sensitivity

Increase number of analysed candidates

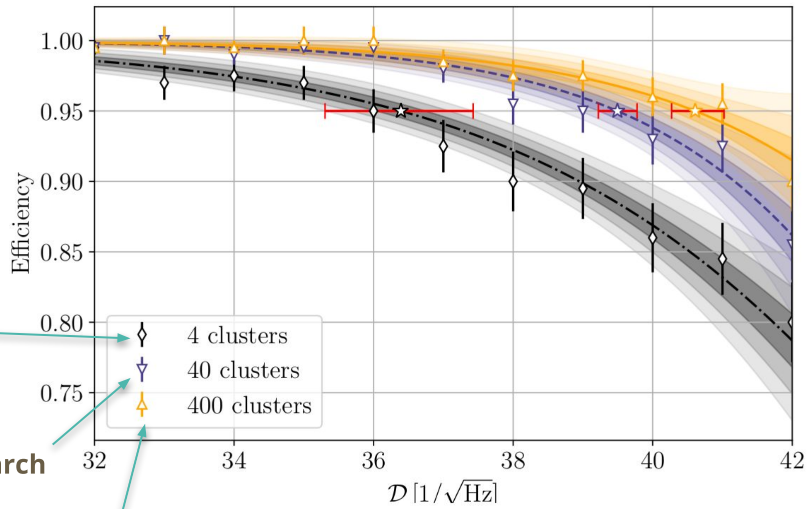
Sensitivity improves at the price of higher p_{fa}

Mimicked O3 injection campaign of SkyHough pipeline

Improvement from 5% to 20% depending on the search and frequency

Improvements can be different for other pipelines

Cluster = Outliers ascribed to the same cause



O3a-style search

O3-style search

Future?

SkyHough followed up clusters each 0.1 Hz

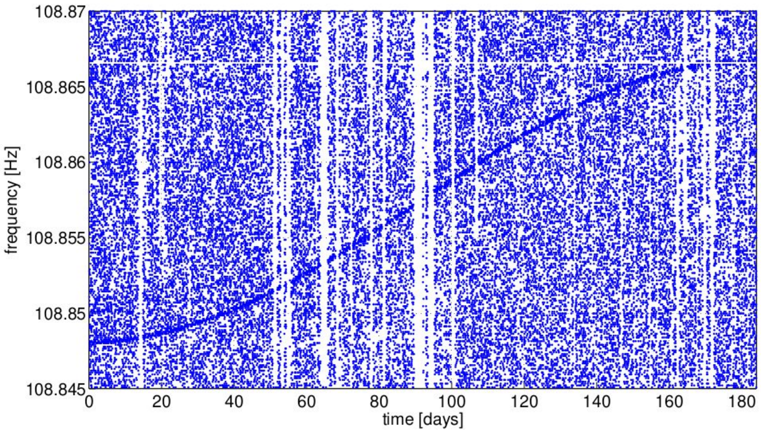
See Mirasola & Tenorio arxiv (2024)

Conclusions

- All-sky searches as **robust** but computationally expensive method
  - semi-coherent methods used to reduce computational cost
  - sensitivity to CWs related to segment length
  - **identify *O*($10^{5\text{-}6}$) outliers**

$$h_{\text{sens}} \propto (T_{\text{obs}}\, T_{\text{FFT}})^{-1/4}$$

- Follow-up methods needed to improve sensitivity
  - methods must be **cheap** and **efficient**
  - increase coherence time to improve SNR and parameter estimation
- **Gridded** approaches as “**brute-force**” methods
  - probe **all** the parameter-space points around each outlier’s parameters
  - higher **chances** to detect signals **close to the threshold**
- **MCMC** approaches are based on a **maximisation likelihood** procedure
  - walkers randomly move towards **high posterior probability regions**
  - Here presented for local analyses, but broader regions are definitely possible (Covas+ 2024)
- Other approaches are being used (CNN, …) even if not mentioned
- Sensitivity to CWs can be improved also **analysing more candidates**
  - 5%-20% improvements depending on search/frequency/pipeline


108.87
108.865
108.86
108.855
108.85
108.845
frequency [Hz]
0
20
40
60
80
100
120
140
160
180
time [days]

STAY TUNED!

BACKUP

Where we are

O4 sensitivity

⬇

~2-4 time better than O3

⬇

Rough estimation that depends on frequency/pipeline

O3 upper limits

$$h_{sens} \propto (T_{obs} T_{FFT})^{-1/4}$$

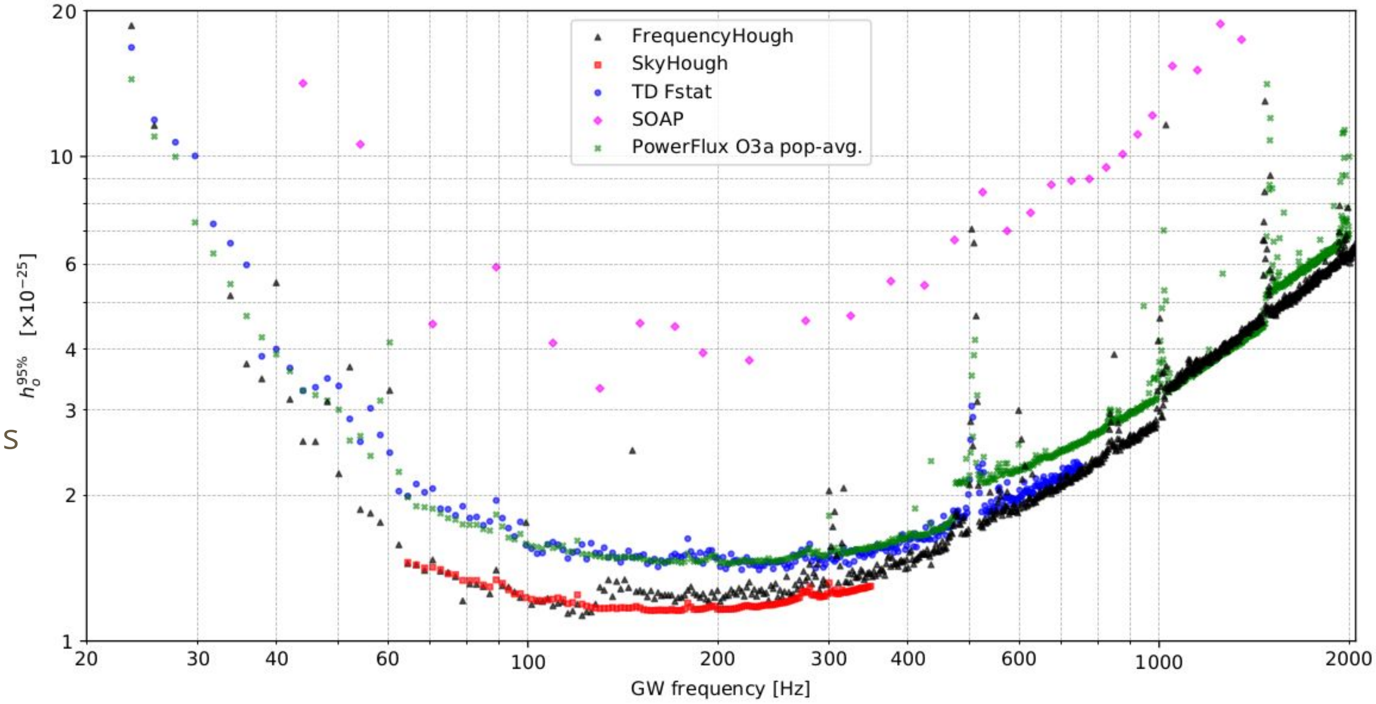
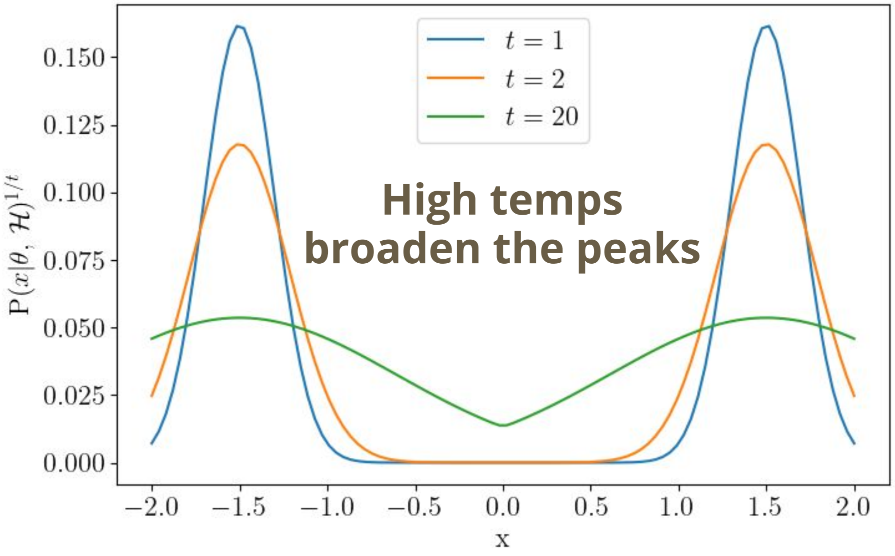


Figure taken from Phys. Rev. D 106, 102008

Parallel tempered MCMC

$$P(\theta|\boldsymbol{x}, t_i, \mathcal{H}) \propto P(\boldsymbol{x}|\theta, \mathcal{H})^{1/t_i} P(\theta|\mathcal{H})$$



Mismatch threshold: false-dismissal

CDF of $\chi^2(4N_{seg},\rho^2)$

$$p_{\mathrm{fd}}(2\mathcal{F}_{\mathrm{thr}}, \mathcal{D}) = \int_0^\infty \mathrm{d}\rho_0^2\, p(\rho_0^2)\, p(2\hat{\mathcal{F}} < 2\mathcal{F}_{\mathrm{thr}} | \rho^2 = \rho_0^2/\mathcal{D}^2)$$

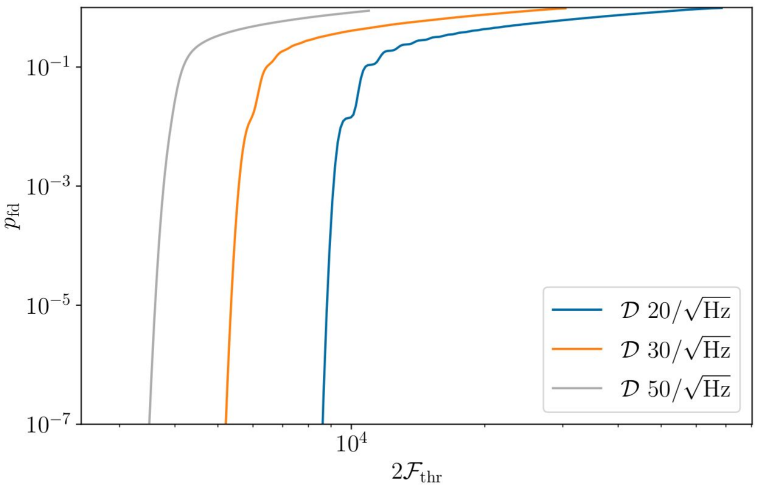
SNR distribution for D = 1/sqrt(Hz)

Sampled numerically

$$\mathcal{D} = \frac{\sqrt{S_{\mathrm{n}}}}{h_0}$$

We set the threshold using $p_{fd} = 10^{-8}$

Usual searches use p_{fd} = 1e-5



Released an open python package cows3

Evaluation of the MCMC performances

p_{fd} “assumes” perfect reconstruction of the parameters

need to introduce a mismatch

$$2\mathcal{F}^{\mu}_{\rm thr} = 2\mathcal{F}_{\rm thr} \cdot (1-\mu) + 4N_{\rm seg}\mu$$

The “effective” threshold is lowered

How much can we “afford” to lower?

Injected parameters

Offset introduced during the search

$$\mu(\Delta\lambda;\lambda_{\rm s}) = \frac{2\hat{\mathcal{F}}(\lambda_{\rm s}) - 2\hat{\mathcal{F}}(\lambda_{\rm s}+\Delta\lambda)}{2\hat{\mathcal{F}}(\lambda_{\rm s}) - 4N_{\rm seg}}$$

$\mathrm{int}(T_{obs}/T_{coh})$

mismatch = SNR^2 fractional loss

Noise distribution going to take over at some point

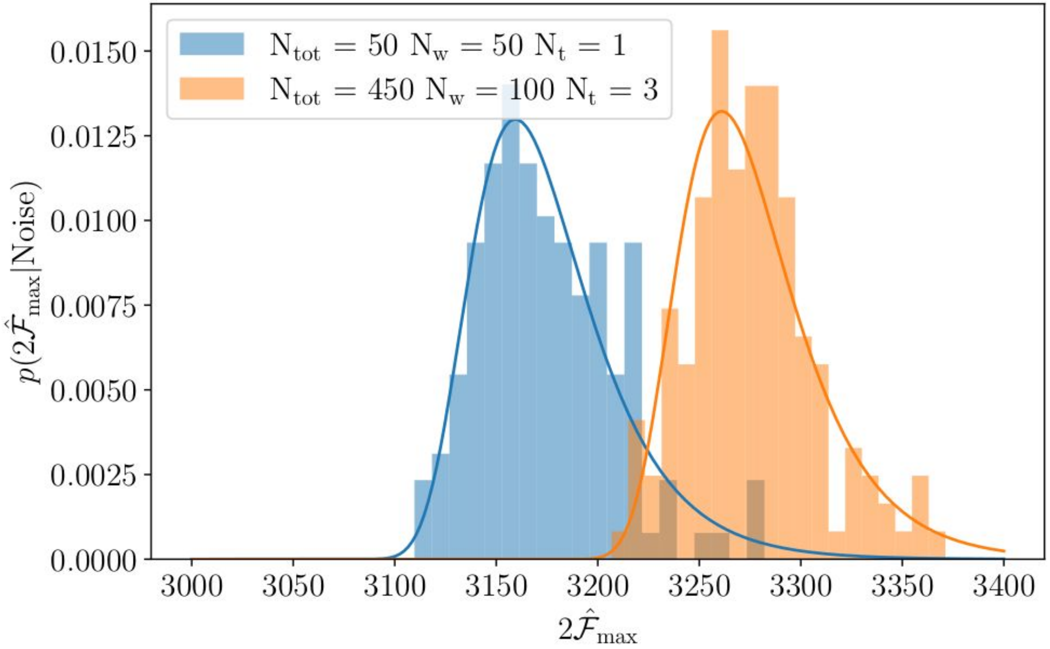
Mismatch related to false-alarm

Mismatch threshold: false-alarm and mismatch

Need to compare this new threshold with the noise distribution

MCMC noise distribution depends on its configuration

Novelty w.r.t. to past searches and explanation to very high Bayes factor seen previously

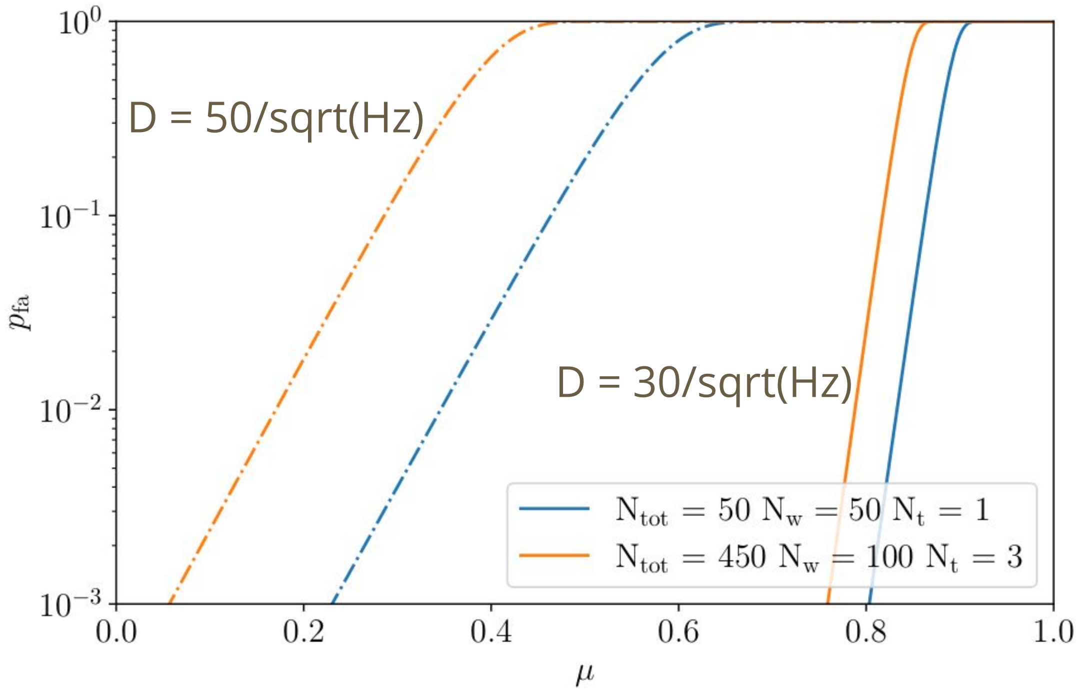


Mismatch threshold: false-alarm and mismatch (2)

Link between mismatch and false-alarm at fixed p_{fd}

↓

We used $p_{fa} = 10^{-1}$–10^{-2} from now on



Released an open python package cows3

MCMC optimisation: injection campaign

Randomly shifted from the injection point

	Interval	Prior width
f_0	$[50,\ 500]$ Hz	$\frac{1}{2T_{\rm coh}}$
f_1	$[-10^{-8},\ 10^{-9}]$ Hz/s	$\frac{1}{2T_{\rm coh}T_{\rm obs}}$
α	$[-\pi,\ \pi]$ rad	$\frac{10^4}{2fT_{\rm coh}}$
δ	$[-\pi/2,\ \pi/2]$ rad	$\frac{10^4}{2fT_{\rm coh}}$
P	$[7,\ 15]$ d	$\frac{\pi}{f\,a_p T_{\rm coh} T_{\rm obs} \Omega^3}$
a_p	$[5,\ 15]$ lt-s	$\frac{1}{2fT_{\rm coh}\Omega}$
$t_{\rm asc}$	$[t_{\rm mid} - \frac{P}{2},\ t_{\rm mid} + \frac{P}{2}]$	$\frac{5}{2f\ a_p\ T_{\rm coh}\ \Omega^2}$

Gaussian

Uniform

MCMC configurations

$N_{\rm w}$	50, 100
$N_{\rm t}$	1, 2, 3
$N_{\rm tot}$	50, 100, ..., 450, 500

100 injections for each configuration

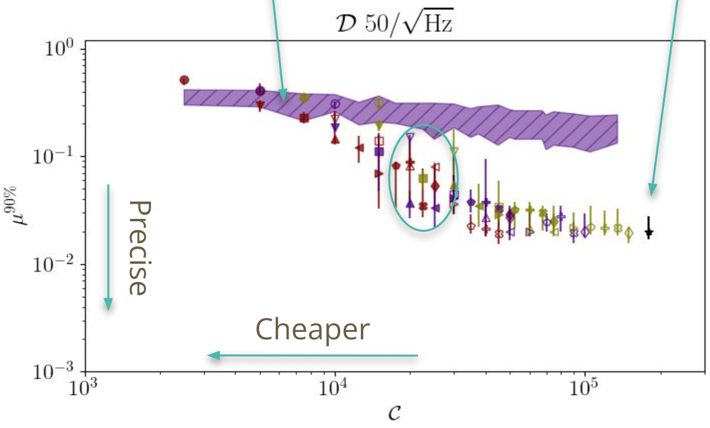
Single-stage MCMC

Threshold for $p_{fd} = 10^{-8}$

$p_{fa} = 10^{-1}–10^{-2}$

Configuration of past searches

Results in real data but similar in simulated Gaussian noise

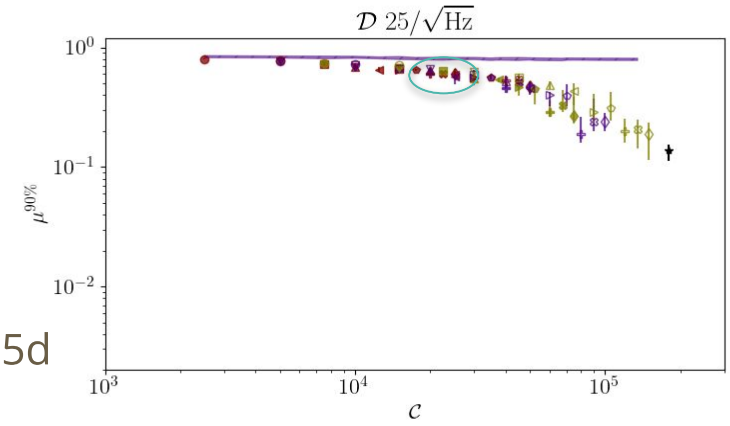


Isolated source

O3 LL+LH data

T_{coh}

$O(10^3 s) \rightarrow 0.5d$



Source in binary system

O3a LL+LH data

Cost can be reduced by an order of magnitude

Multi-stage MCMC: coherence times ladder

Previous ladder based on parameter space metric

$$\mathcal{N}_* \approx \frac{\sqrt{g\left(T_{\mathrm{coh}}^{(j+1)}\right)}}{\sqrt{g\left(T_{\mathrm{coh}}^{(j)}\right)}}$$

Number of templates w.r.t. previous stage

⬇

MCMC can deal with $10^{3\text{-}4}$

see [PRD 97, 103020 (2018)]

Metric that encapsulates the T_{coh} dependence

⬇

Cons: metric not always available

If outlier is consistent with a CW

⬇

Increase T_{coh}

Our proposal:
need only scaling with T_{coh}

$$\mathcal{N}_* = \left(\frac{T_{\mathrm{coh}}^{(j+1)}}{T_{\mathrm{coh}}^{(j)}}\right)^d$$

follow-ups probe small parameter regions

⬇

metrics ratio depends only on T_{coh} and # of resolved parameters