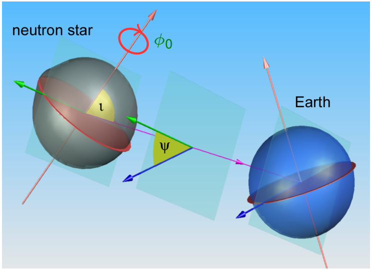
# The weak-signal Bayes-factor approximation


Reinhard Prix

Continuous-waves group
Albert-Einstein-Institute Hannover


Continuous gravitational waves and neutron stars @ Hannover
18 June 2024

# Signal amplitude parameters




☞ Amplitude parameters:

$$\mathcal{A} = \{h_0, \underbrace{\cos\iota, \psi, \phi_0}\}$$

"radius" ($h_0 \geq 0$) + 3 Euler angles

$\mathcal{A}$ *generally* unknown

Any spin orientation equally likely (isotropic):

$$P(\cos\iota | I) = \mathrm{unif}(-1, 1)$$
$$P(\psi | I) = \mathrm{unif}(0, 2\pi)$$

Any initial phase equally likely:

$$P(\phi_0 | I) = \mathrm{unif}(0, 2\pi)$$

☞ "physical prior" $P(h_0, \cos\iota, \psi, \phi_0 | I) = \mathrm{const} \times \underbrace{P(h_0|I)}_{?}$

# Part I: Pre-history

From $\mathcal{F}(x)$- to $\mathcal{B}(x)$-statistic

# JKS coordinates $\{\mathcal{A}^\mu\}$

[Jaranowski,Królak,Schutz PRD58(1998)]

"cartesian" JKS coordinates $\mathcal{A}^\mu = \mathcal{A}^\mu(h_0, \cos\iota, \psi, \phi_0)$

$$
\begin{aligned}
\mathcal{A}^1 &= \phantom{-}0.5h_0(1+\cos^2\iota)\cos 2\psi\,\cos\phi_0 - h_0\cos\iota\,\sin 2\psi\,\sin\phi_0 \\
\mathcal{A}^2 &= \phantom{-}0.5h_0(1+\cos^2\iota)\sin 2\psi\,\cos\phi_0 + h_0\cos\iota\,\cos 2\psi\,\sin\phi_0 \\
\mathcal{A}^3 &= -0.5h_0(1+\cos^2\iota)\cos 2\psi\,\sin\phi_0 - h_0\cos\iota\,\sin 2\psi\,\cos\phi_0 \\
\mathcal{A}^4 &= -0.5h_0(1+\cos^2\iota)\sin 2\psi\,\sin\phi_0 + h_0\cos\iota\,\cos 2\psi\,\cos\phi_0
\end{aligned}
$$

## ☞ Signal strain @ detector

$$
h(t;\mathcal{A};\lambda) = \sum_{\mu=1}^{4} \mathcal{A}^\mu\, h_\mu(t;\lambda)
$$

$$
h_1(t;\lambda) = a(t;\vec{n}_{\rm sky})\,\cos\phi(t;\lambda), \quad h_2(t;\lambda) = b(t;\vec{n}_{\rm sky})\,\cos\phi(t;\lambda)
$$

$$
h_3(t;\lambda) = a(t;\vec{n}_{\rm sky})\,\sin\phi(t;\lambda), \quad h_4(t;\lambda) = b(t;\vec{n}_{\rm sky})\,\sin\phi(t;\lambda)
$$

Phase-evolution parameters: $\lambda = \{\vec{n}_{\rm sky}, \textit{freq}, \textit{spindown}, \textit{binary}, \ldots\}$

(here: $\lambda$ assumed known)

# Likelihood ratio $\mathcal{L}(x; \mathcal{A}, \lambda)$

$\mathcal{H}_{\rm G}$: data = Gaussian noise: $x(t) = n(t)$
$\mathcal{H}_{\rm S}$: data = Gaussian noise + signal: $x(t) = n(t) + h(t; \mathcal{A}, \lambda)$

Goal: given $x(t)$ ☞ "optimally" decide between $\mathcal{H}_{\rm G}$ and $\mathcal{H}_{\rm S}$

☞ likelihood ratio: Gaussian in $\mathcal{A}^\mu$

$$\mathcal{L}(x; \mathcal{A}, \lambda) \equiv \frac{P\left(x|\mathcal{H}_{\rm S}\right)}{P\left(x|\mathcal{H}_{\rm G}\right)} = \exp\left[-\frac{1}{2}\mathcal{A}^\mu \mathcal{M}_{\mu\nu}\mathcal{A}^\nu + \mathcal{A}^\mu x_\mu\right]$$

4 matched-filter outputs $x_\mu(\lambda) \equiv (x|h_\mu(\lambda)) \in \mathbb{R}$
4x4 antenna-pattern matrix:

$$\mathcal{M}_{\mu\nu}(\vec{n}_{\rm sky}) \equiv (h_\mu|h_\nu) = \underbrace{(S_{\rm n}^{-1}T)}_{\gamma\equiv\text{"quality x quantity"}} \begin{pmatrix} A & C & & \\ C & B & & O \\ & O & A & C \\ & & C & B \end{pmatrix}$$

matched-filter scalar product: $(x|y) \equiv S_{\rm n}^{-1} \int x(t)\, y(t) dt = \gamma\ \langle xy \rangle$

$A \equiv \left\langle a^2 \right\rangle$, $B \equiv \left\langle b^2 \right\rangle$, $C \equiv \langle ab \rangle$

# 1998: Maximum-likelihood: $\mathcal{F}$-statistic

☞ analytically maximize $\mathcal{L}$ over $\mathcal{A}^\mu$: [Jaranowski,Królak,Schutz PRD58(1998)]

$$\mathcal{F}(x) \equiv \max_{\{\mathcal{A}^\mu\}} \ln \mathcal{L}(x; \mathcal{A}^\mu) = \frac{1}{2} x_\mu \left(\mathcal{M}^{-1}\right)^{\mu\nu} x_\nu$$

$$2\mathcal{F}(x) = 2D^{-1}\left[AB(\mathcal{F}_\mathrm{A}(x) + \mathcal{F}_\mathrm{B}(x)) - 2C^2\mathcal{F}_\mathrm{C}(x)\right]$$

$$2\mathcal{F}_\mathrm{A}(x) \equiv \frac{x_1^2 + x_3^2}{\gamma A}, \quad 2\mathcal{F}_\mathrm{B}(x) \equiv \frac{x_2^2 + x_4^2}{\gamma B}, \quad 2\mathcal{F}_\mathrm{C}(x) \equiv \frac{x_1 x_2 + x_3 x_4}{\gamma C}$$

sub-determinant $D = \det\begin{pmatrix} A & C \\ C & B \end{pmatrix} = AB - C^2$

☞ Long thought to be “optimal”, until . . .

# 2009: $\mathcal{B}$-statistic [Prix,Krishnan CQG26(2009)]

Neyman-Pearson-Searle theorem [Searle arXiv:0804.1161]

☞ Bayes factor is optimal statistic (if drawing from prior)

$$B_{\mathrm{S/G}}(x) \equiv \int \mathcal{L}(x;\mathcal{A})\,P(\mathcal{A}|I)\;d^4\mathcal{A}$$

- "$\mathcal{F}$-statistic prior" $P(\mathcal{A}^\mu) = \mathrm{const}$ ☞ $B_{\mathrm{S/G}} \propto e^{\mathcal{F}}$

  (unphysical prior: $P(h_0, \cos\iota, \psi, \phi_0) \propto {h_0}^3 (1 - \cos^2\iota)^3$)

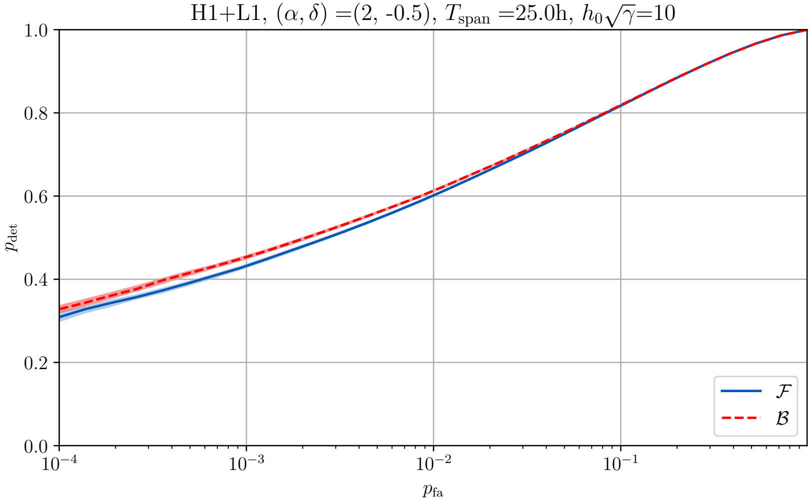
- "physical prior" $P(h_0, \cos\iota, \psi, \phi_0) = \mathrm{const} \times P(h_0)$
  - ☞ integrate $\phi_0$ ✓
  - ☞ if $P(h_0) = \mathrm{const} \Rightarrow$ integrate $h_0$! ✓

"$\mathcal{B}$-statistic" ☞ 2D integral [Prix,Krishnan CQG26(2009)],[Whelan+ CQG31(2014)]

$$\mathcal{B}(x) \propto \int \frac{1}{\varrho}\, e^{\Theta}\, I_0(\Theta)\; d\cos\iota\; d\psi$$

$$\Theta(x;\cos\iota,\psi) \equiv \frac{q^2}{4\varrho^2},\; q = q(x;\cos\iota,\psi),\; \varrho = \varrho(\cos\iota,\psi)$$

# $\mathcal{B}(x)$ beats $\mathcal{F}(x)$




various attempts to solve/approximate $\mathcal{B}(x)$: [Dergachev PRD85(2012)][Whelan+ CQG31(2014)][Haris,Pai PRD96(2017)][Dhurandhar+ arXiv:1707.08163][Bero,Whelan CQG36(2019)][Wette Universe7(2021)]

# Part II: Recent surprises

## Short segments and weights

# Surprise 1: dominant-response $\mathcal{F}_{\mathrm{AB}}$-statistic [Covas,Prix PRD105(2022)]

Fully Coherent Search

Coherent + Coherent + Coherent + Coherent + Coherent

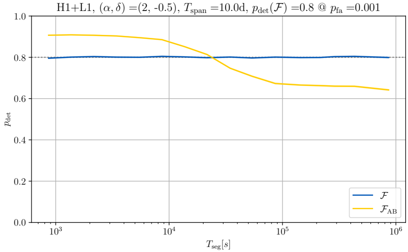
Sum = Semicoherent Search

All-sky binary search ☞ very short segments $T_{\mathrm{seg}} \sim \mathcal{O}(100s)$ ☞ degenerate $\mathcal{F}(x)$ in certain $\vec{n}_{\mathrm{sky}}$ where $D \to 0$, ☞ maximum-likelihood $\mathcal{A}^{\mu}{}_{\mathrm{ML}} = \mathcal{M}^{-1\,\mu\nu} x_{\mu}$ fails!

Fix: “dominant-response” statistic:

$$\mathcal{F}_{\mathrm{AB}}(x) \equiv \begin{cases} \mathcal{F}_{\mathrm{A}}(x) & \text{if } A > B \\ \mathcal{F}_{\mathrm{B}}(x) & \text{otherwise} \end{cases}$$

☞ up to $\sim 19\%$ more sensitive than $\mathcal{F}(x)$ for short segments!!

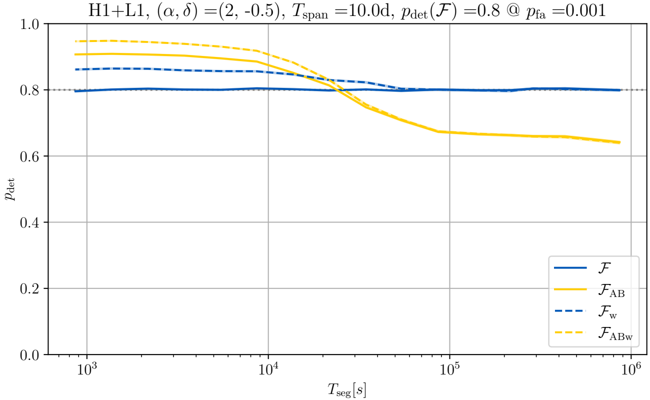


# Surprise 2: segment weighting [Covas,Prix PRD106(2022)]

Can improve statistics by (empirical) *per-segment weights* $w_\ell$:

("data factor" $\gamma = S_n^{-1} T$)

- $\widehat{\mathcal{F}}_w \equiv \sum_{\ell=1}^{N_{\mathrm{seg}}} \left(\gamma\,(A+B)\right)_\ell\, \mathcal{F}_\ell$
- $\widehat{\mathcal{F}}_{\mathrm{ABw}} \equiv \sum_{\ell=1}^{N_{\mathrm{seg}}} \left(\gamma\left(Q + \frac{C^2}{Q}\right)\right)_\ell \mathcal{F}_{\mathrm{AB},\ell}$ (with $Q = \max[A, B]$)




☞ $\mathcal{F}$-statistic far from optimal for short segments!

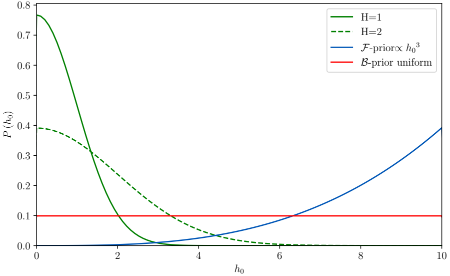
# Part III

Is there a ~~better~~ Bayesian way?

# New $h_0$-prior

## Half-normal distribution with $h_0$-“scale” $H$

$$P\left(h_0 | I, H\right) = \frac{\sqrt{2}}{\sqrt{\pi}\,H}\, e^{-\frac{{h_0}^2}{2H^2}} \quad \text{for } h_0 \geq 0$$




- ✓ favors weak signals over strong
- ✓ proper prior
- ✓ $H \to \infty$ ☞ reduces to uniform-$h_0$
- ✓ conjugate to $\mathcal{L}$: can integrate $h_0$!

# Generalized $\mathcal{B}$-statistic

"relative scale" $H_* \equiv H\sqrt{\gamma}$

$$\mathcal{B}(x; H_*) = \frac{1}{\pi} \int d\cos\iota \, d\psi \frac{1}{\sqrt{1 + H_*^2 \varrho^2}} \, e^{\Theta} \, I_0(\Theta)$$

$$\Theta(x; \cos\iota, \psi) \equiv \frac{H_*^2 \, q^2(x; \cos\iota, \psi)}{4(1 + H_*^2 \, \varrho^2(\cos\iota, \psi))}$$

Two interesting limits:

- strong-signal limit: $H_* \gg 1$ ☞ recover const-$h_0$ $\mathcal{B}(x)$
- weak-signal limit: $H_* \ll 1$: Taylor-expand in $H_*$!

(coherent sensitivity $h_0|^{p_{\mathrm{det}}=90\%}_{p_{\mathrm{fa}}=1\%} \sim 11.4/\sqrt{\gamma}$ ☞ $h_0\sqrt{\gamma} \sim \mathcal{O}(10)$)

# Analytic weak-signal limit $H_* \ll 1$

Taylor-expand $\mathcal{B}(x; H_*)$ in $\mathcal{O}(H_*)$:

$$\mathcal{B}(x) = 1 + \frac{H_*^2}{4} \left\langle q^2 - 2\varrho^2 \right\rangle_{\cos\iota,\psi} + \mathcal{O}\left(H_*^4\right)$$

☞ Weak-signal limit Bayes factor $\mathcal{B}_{\mathrm{weak}}(x)$

$$\ln \mathcal{B}_{\mathrm{weak}}(x) \equiv \frac{H^2}{10}\gamma\left[A\left(2\mathcal{F}_{\mathrm{A}}(x) - 2\right) + B\left(2\mathcal{F}_{\mathrm{B}}(x) - 2\right)\right]$$

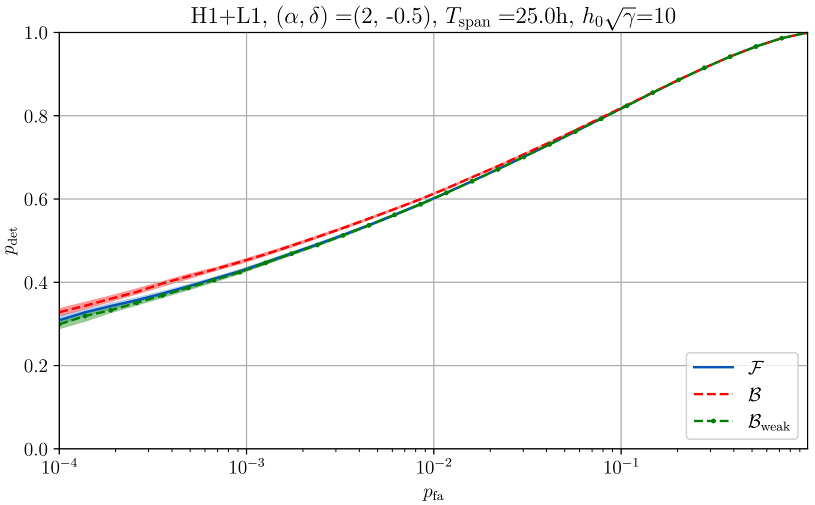
☞ generalized $\chi^2$-distribution with (2,2) degrees of freedom

☞ includes segment weights $(\gamma A, \gamma B)_\ell$ for free?

Compare:

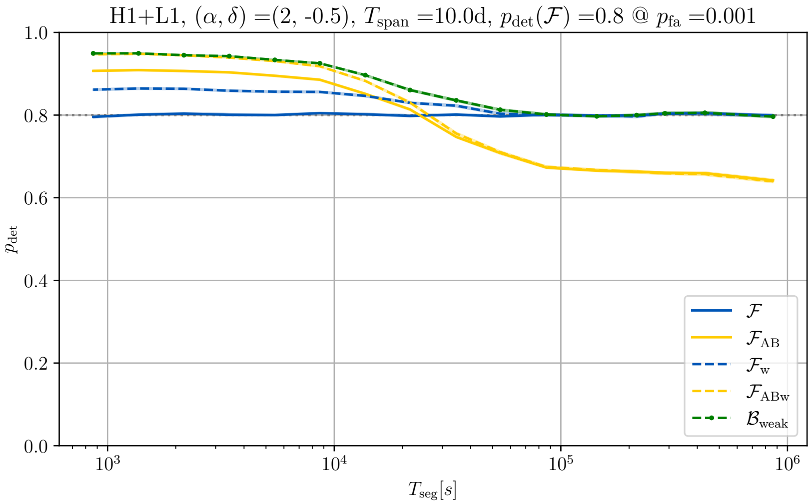
- $2\mathcal{F}(x) = 2D^{-1}\left[AB(\mathcal{F}_{\mathrm{A}}(x) + \mathcal{F}_{\mathrm{B}}(x)) - 2C^2\mathcal{F}_{\mathrm{C}}(x)\right] \sim \chi^2_{4\mathrm{dof}}$
- short-segment $\mathcal{F}_{\mathrm{AB}}(x) = \begin{cases} \mathcal{F}_{\mathrm{A}}(x) & \text{if } A > B \\ \mathcal{F}_{\mathrm{B}}(x) & \text{otherwise} \end{cases} \sim \chi^2_{2\mathrm{dof}}$

# Results

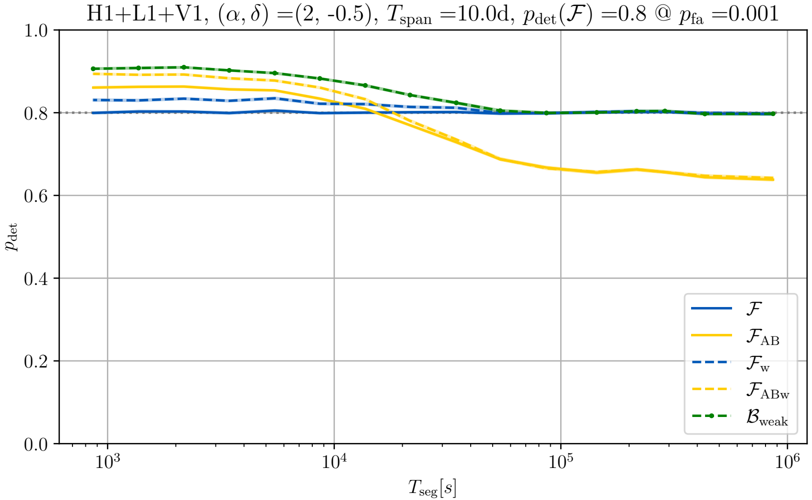


H1+L1, $(\alpha, \delta)$ =(2, -0.5), $T_{\rm span}$ =25.0h, $h_0\sqrt{\gamma}$=10
$p_{\rm det}$
$p_{\rm fa}$
1.0
0.8
0.6
0.4
0.2
0.0
$10^{-4}$
$10^{-3}$
$10^{-2}$
$10^{-1}$
$\mathcal{F}$
$\mathcal{B}$
$\mathcal{B}_{\rm weak}$

# Results (H1+L1)



# Results (H1+L1+V1)



# Summary

- Half-normal $h_0$-prior generalizes $\mathcal{B}$-statistic ☞ $\mathcal{B}(x; H)$
- "weak-signal" limit $H \ll 1/\sqrt{\gamma}$ ☞ new analytical solution
- for longer segments: similar to $\mathcal{F}_{\mathrm{w}}$, weaker than $\mathcal{B}$
- for short segments $\lesssim$ 1d: similar/better than $\mathcal{F}_{\mathrm{ABw}}$
- $\ln \mathcal{B}_{\mathrm{weak}} = \frac{H^2}{10}(\underbrace{\sum_\mu x_\mu^2} - \mathrm{Tr}\mathcal{M})$ vs $2\mathcal{F} = x_\mu (\mathcal{M}^{-1})^{\mu\nu} x_\nu$

  ☞ "5-vector statistic" [Astone+ CQG27(2010)],[D'Onofrio+ arXiv:2406.09236v]
- better line-robust statistics $B_{\mathrm{S/GL}}(H)$? (proper prior)