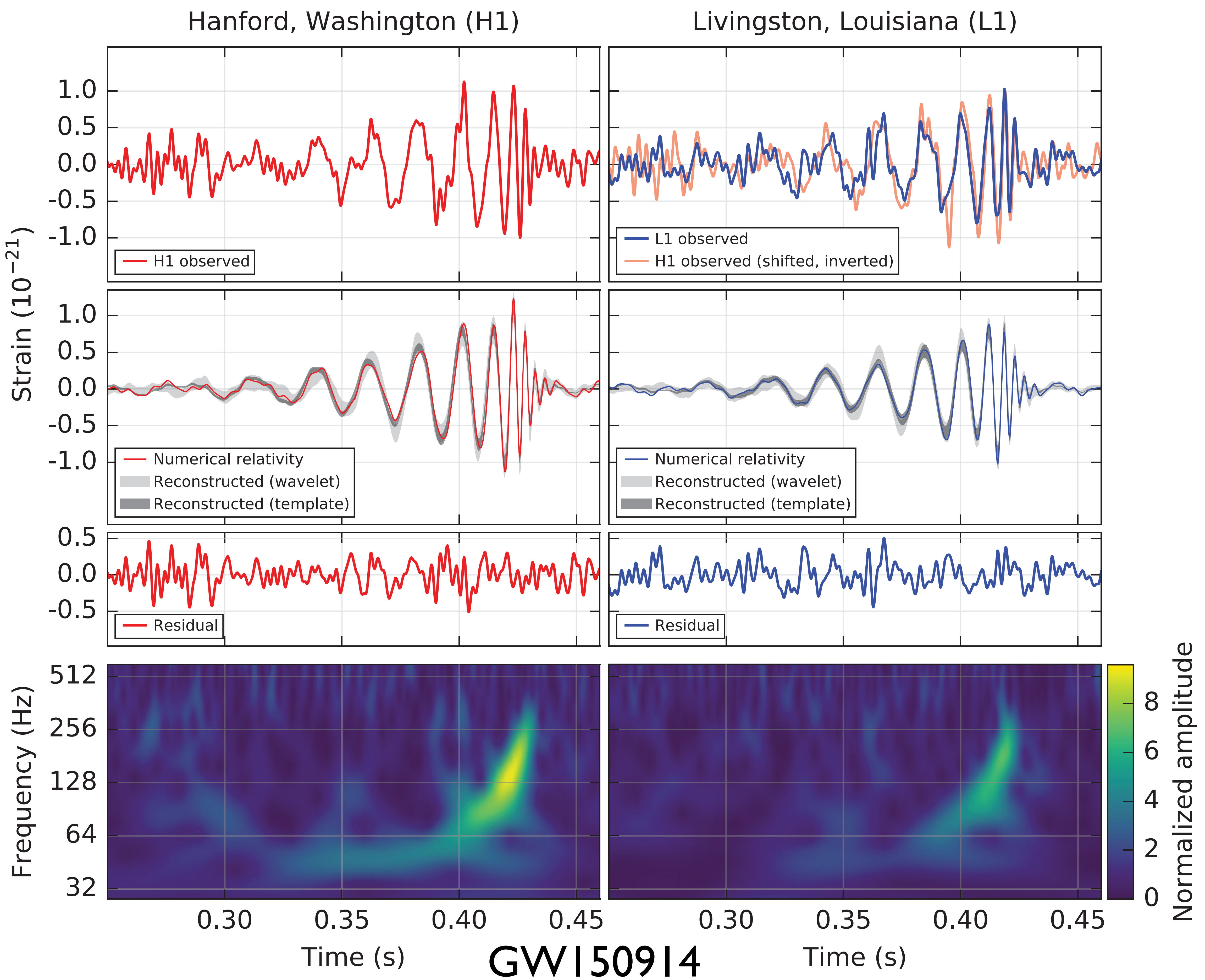
Gravitational wave data analysis
Part I

Sylvain Marsat (L2IT, Toulouse)

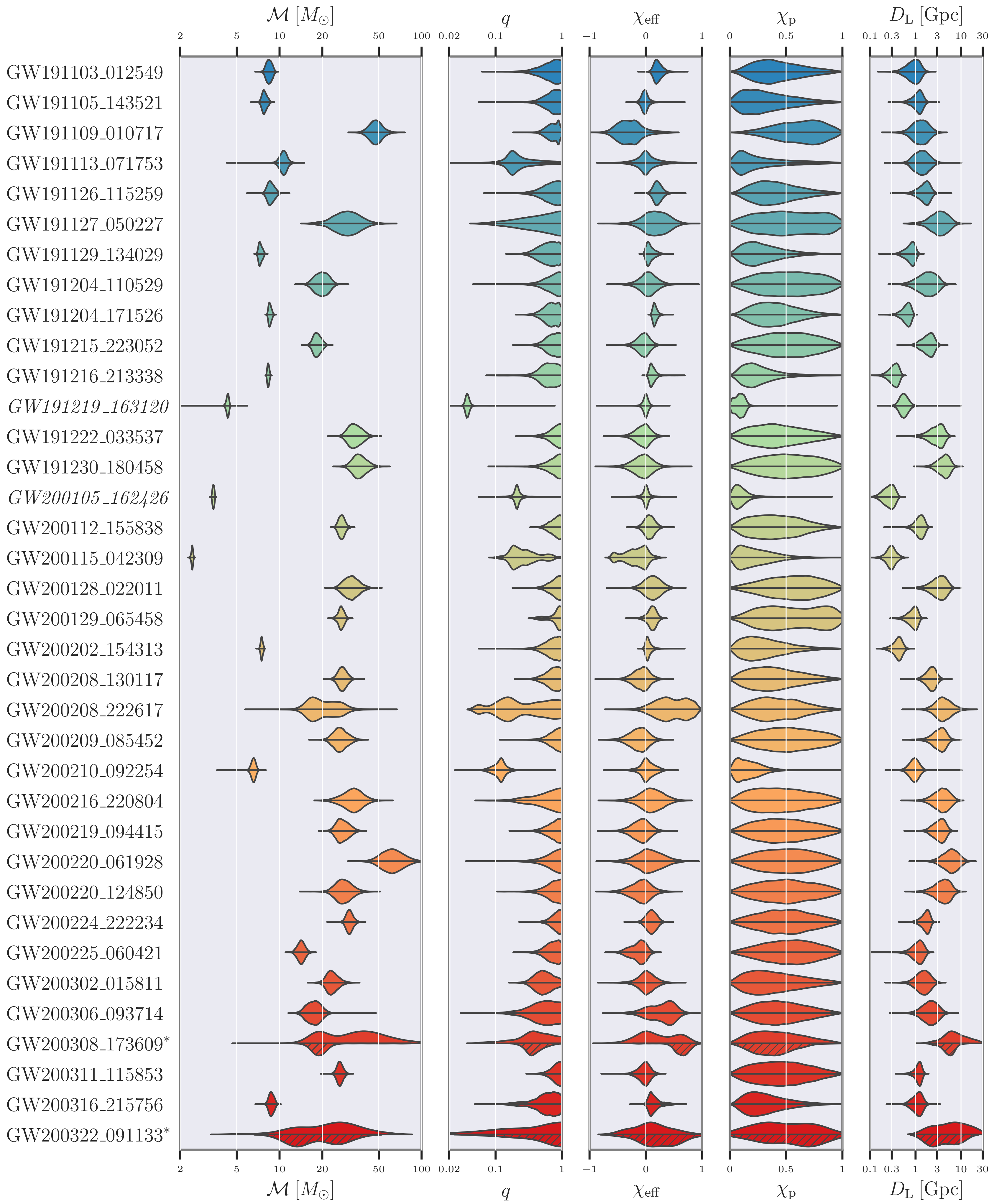
Introduction

From data…

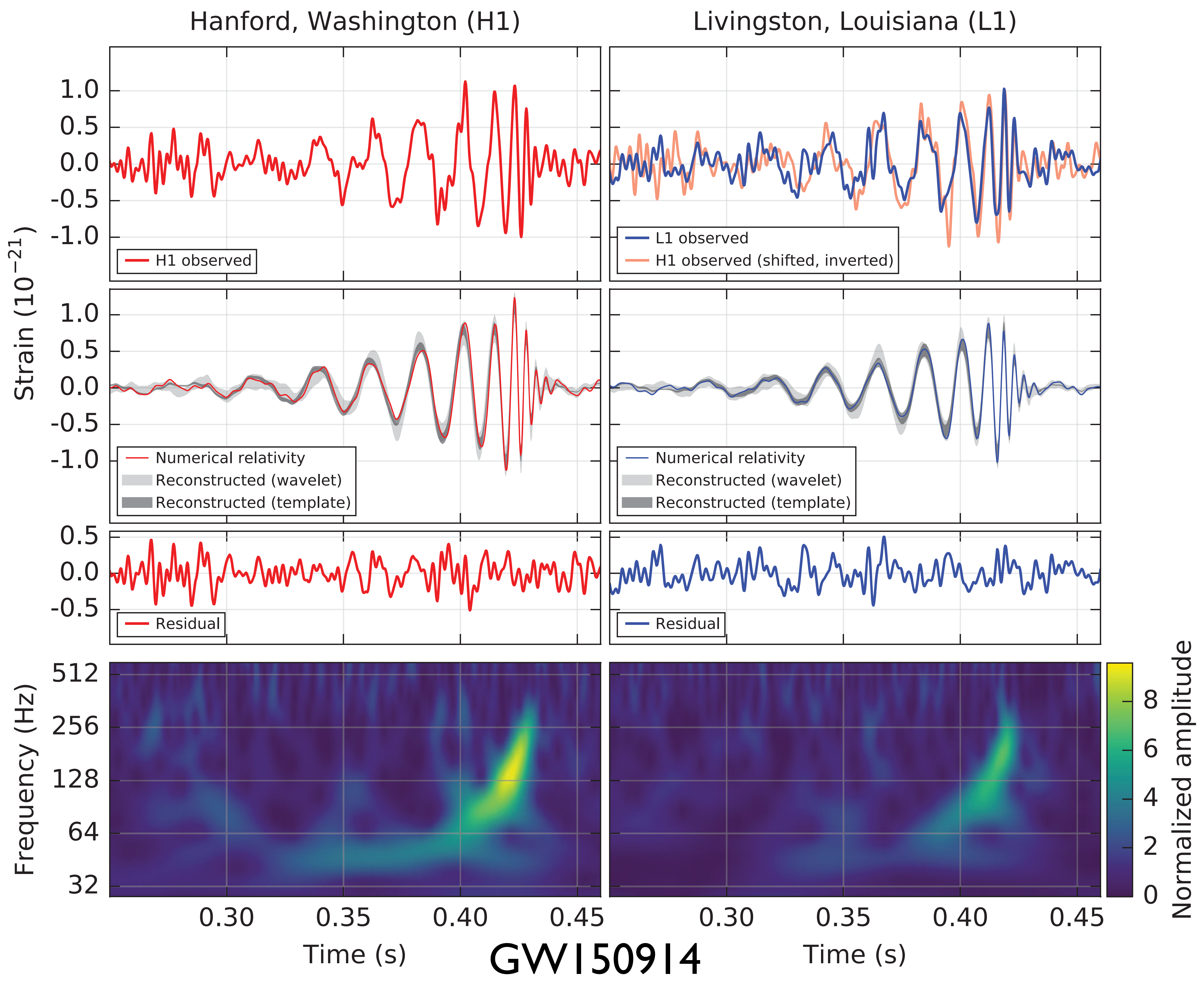


GW150914

… to science



GWTC-3b [LVK 2021]



Data & signal models

Our scope:

- Idealized detector: stationary Gaussian noise
- Simplified CBC signals

Realistic data and signal:

- Data quality for real detectors
- Data cleaning
- Glitch analysis and removal
- Full CBC signals, waveform modelling
- Continuous waves analysis
- Unmodeled signals (bursts)
- Stochastic backgrounds

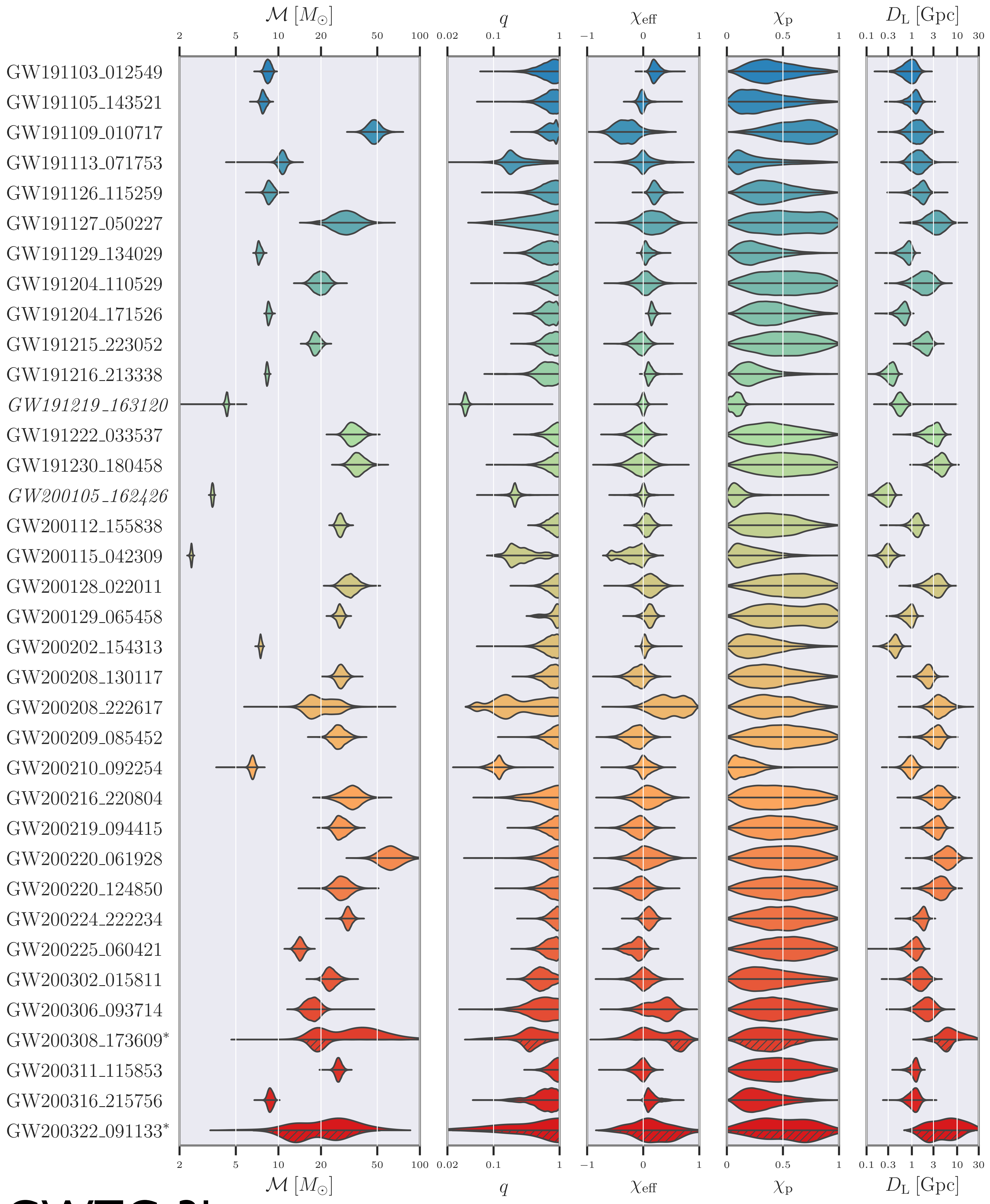
Science products

Our scope:

- Simplified detection with matched filtering
- Simplified parameter estimation

Full science products:

- Realistic detections, confidence and classification
- Realistic parameter estimation
- Evidence computation and model comparison
- Production of catalogs
- Cosmological analyses
- Population analyses
- Tests of GR analyses



GWTC-3b [LVK 2021]

$$\text{Data} = \text{Response} \cdot \text{Signal} + \text{Noise}$$

Detector response

- deterministic instrument transfer (exact)
- calibration: stochastic component

GW signal

- deterministic signals, waveform models
- models approx. GR
- stochastic background(s)

Noise

- stochastic process
- need modelling
- idealized process vs data artefacts ?

Outline

Part I

- GW signals: the basics
- Noise as a stochastic process
- Introducing matched filtering
- Towards real CBC searches
- Other signals: continuous waves, stochastic backgrounds

Part II

- Bayesian parameter estimation basics, likelihood
- Parameter space and waveforms
- Fisher matrix approach
- Metropolis-Hastings MCMC, Parallel tempering and example PE
- PE toolbox
- PE results from LVK
- Future detectors and their challenges

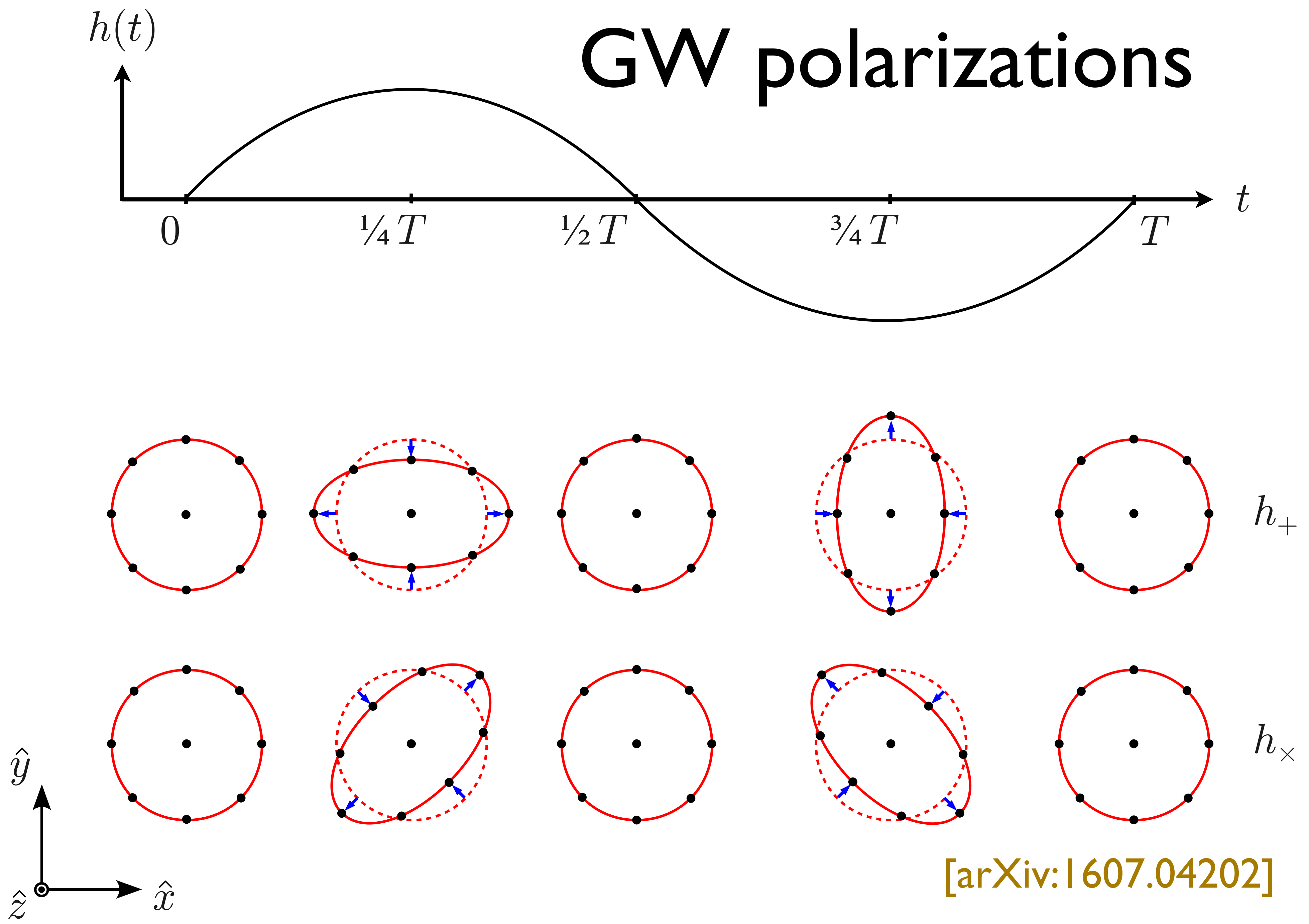
Outline

Part I

- **GW signals: the basics**
- Noise as a stochastic process
- Introducing matched filtering
- Towards real CBC searches
- Other signals: continuous waves, stochastic backgrounds

GW Signals: polarizations and strain

GW polarizations

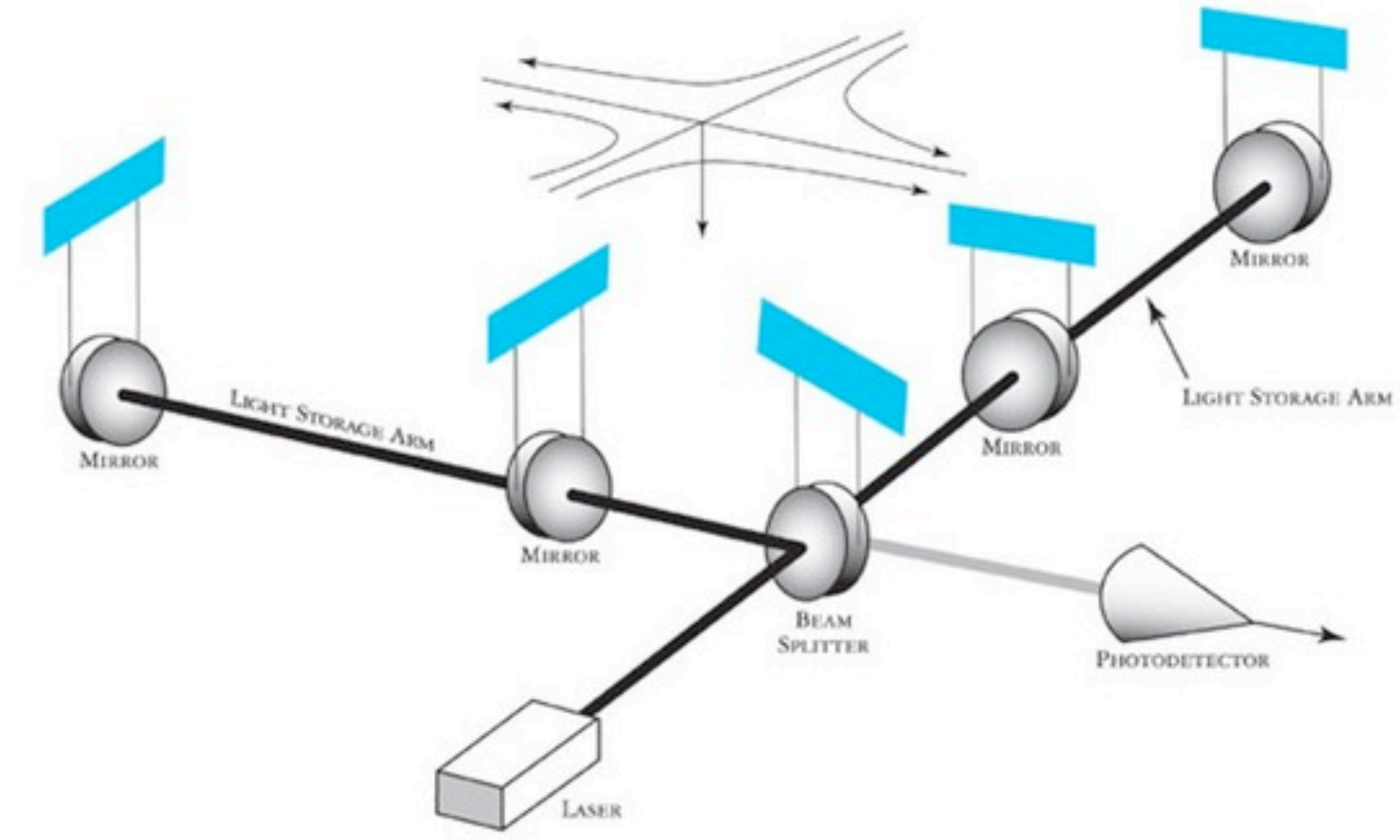


[arXiv:1607.04202]

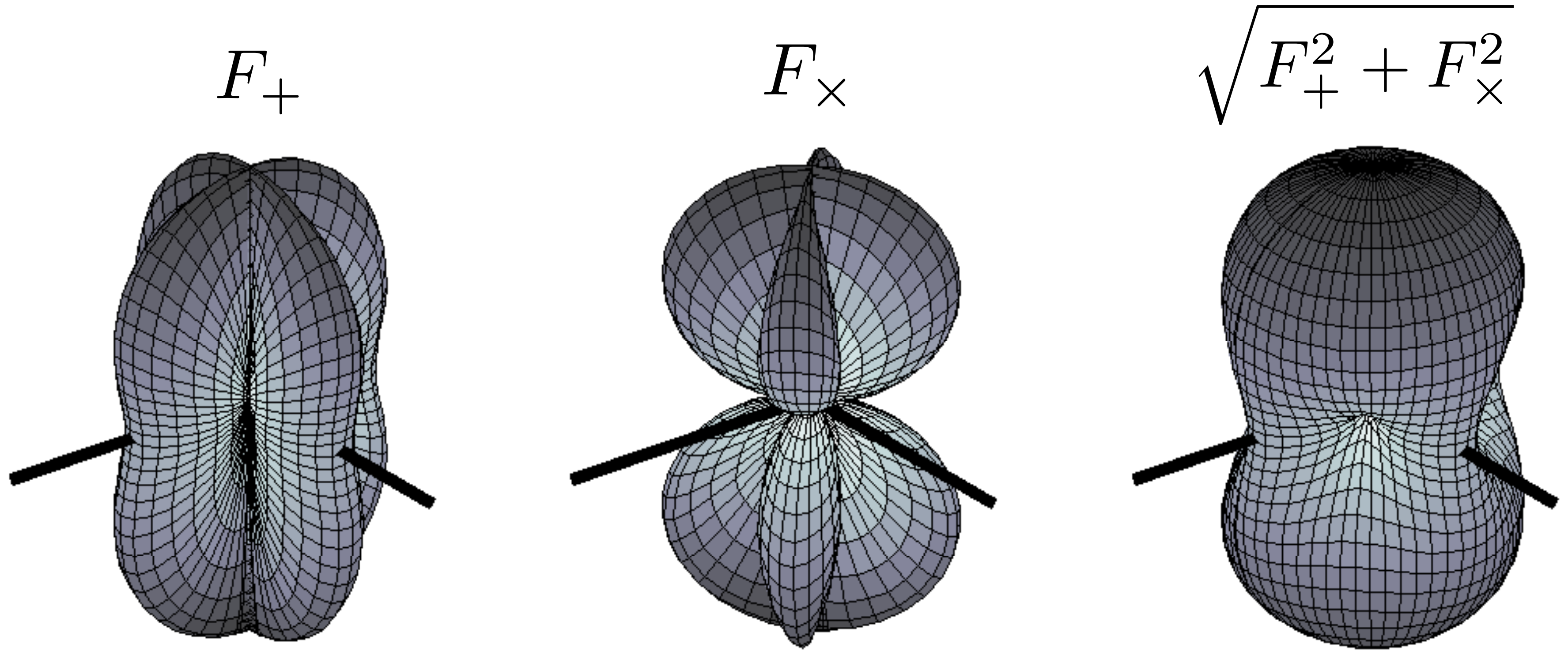
Response of an interferometer:

$$h = F_+ h_+ + F_\times h_\times$$

$F_{+,\times}(\theta, \phi, \psi)$ pattern functions, depend on sky and polarization



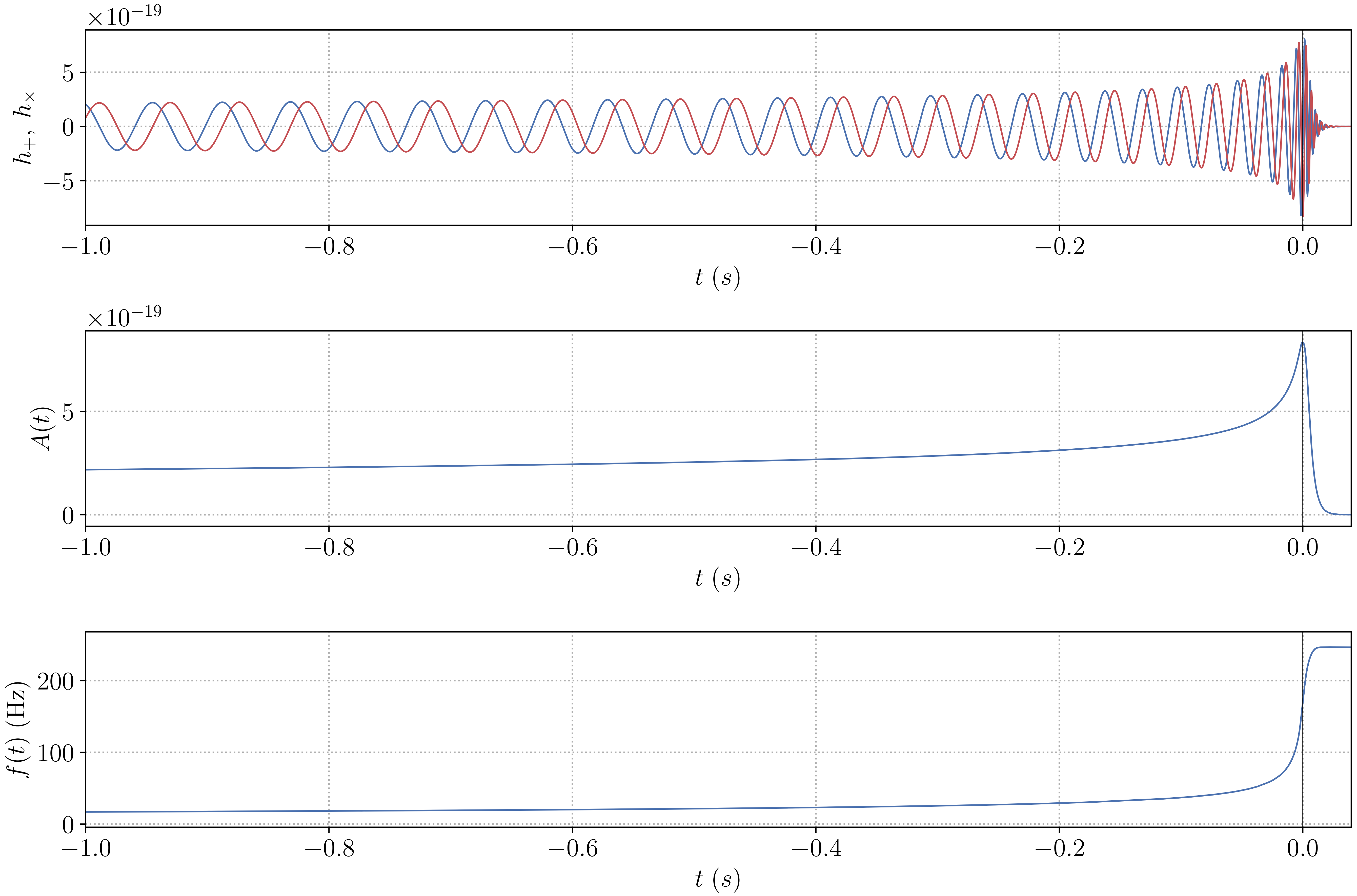
[https://www.ligo.caltech.edu]



[arXiv:0711.3041]

GW Signals: Compact Binary Coalescences - Fact sheet

Inspiral: analytical
Merger/Ringdown: numerical



- Dominant frequency: $f = 2f_{\mathrm{orb}}$
- Chirp mass: $\mathcal{M}_c = \dfrac{m_1^{3/5} m_2^{3/5}}{(m_1 + m_2)^{1/5}}$
- Inspiral frequency:

$$\omega_{\mathrm{orb}}(t) = \left(\frac{G\mathcal{M}_c}{c^3}\right)^{-5/8} \left(\frac{5}{256}\frac{1}{t_c - t}\right)^{3/8}$$

- BBH scale invariance:

$$G = c = 1 \qquad t \to t/M \quad f \to Mf \quad h \to rh/M$$

- End of inspiral:

$$r_{\mathrm{ISCO}} = 6M \quad f_{\mathrm{ISCO}} = 1/6^{3/2}/(\pi M)$$

- Effect of cosmology:

$$M \to (1+z)M \qquad 1/r \to 1/d_L$$

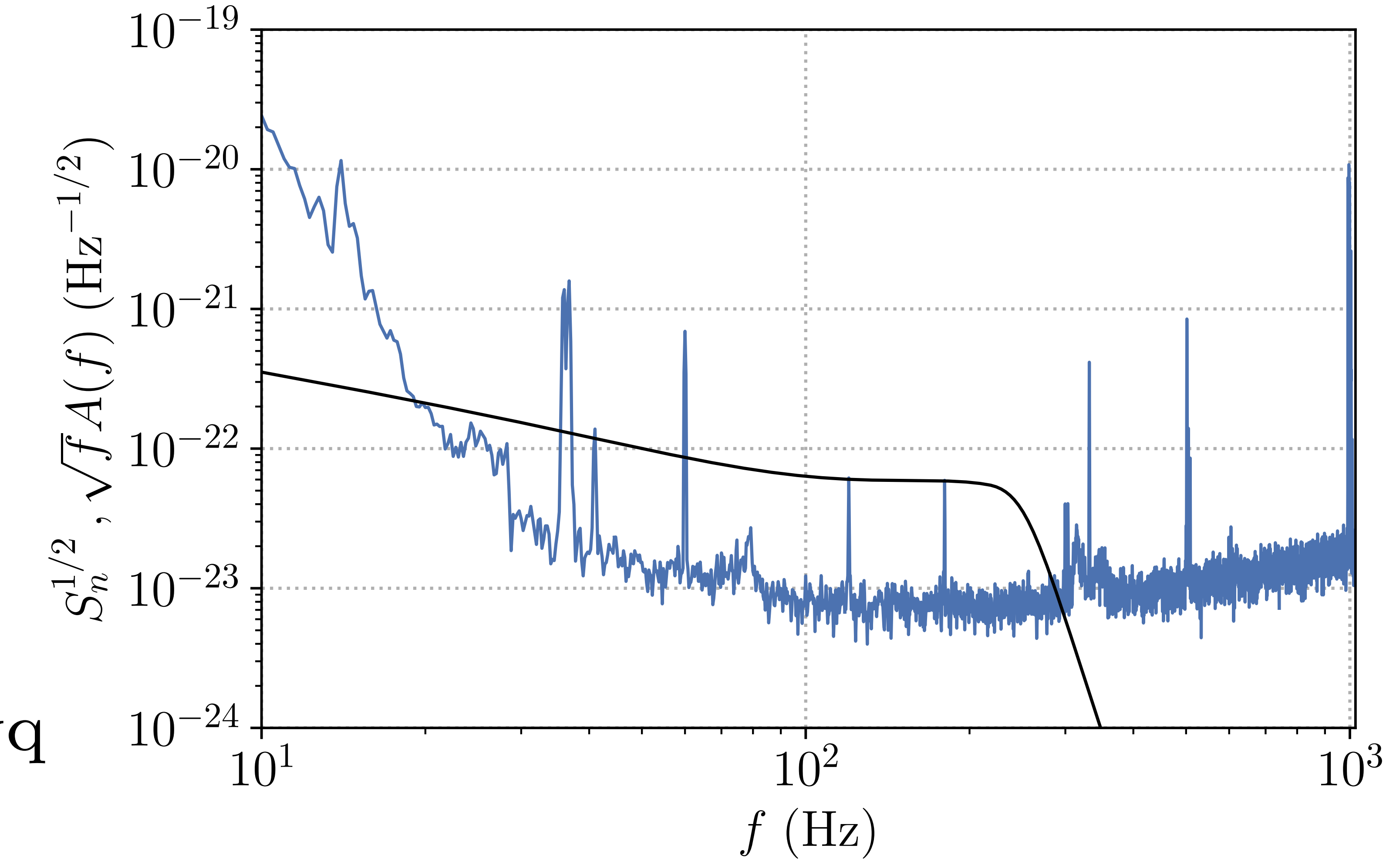
The Fourier domain

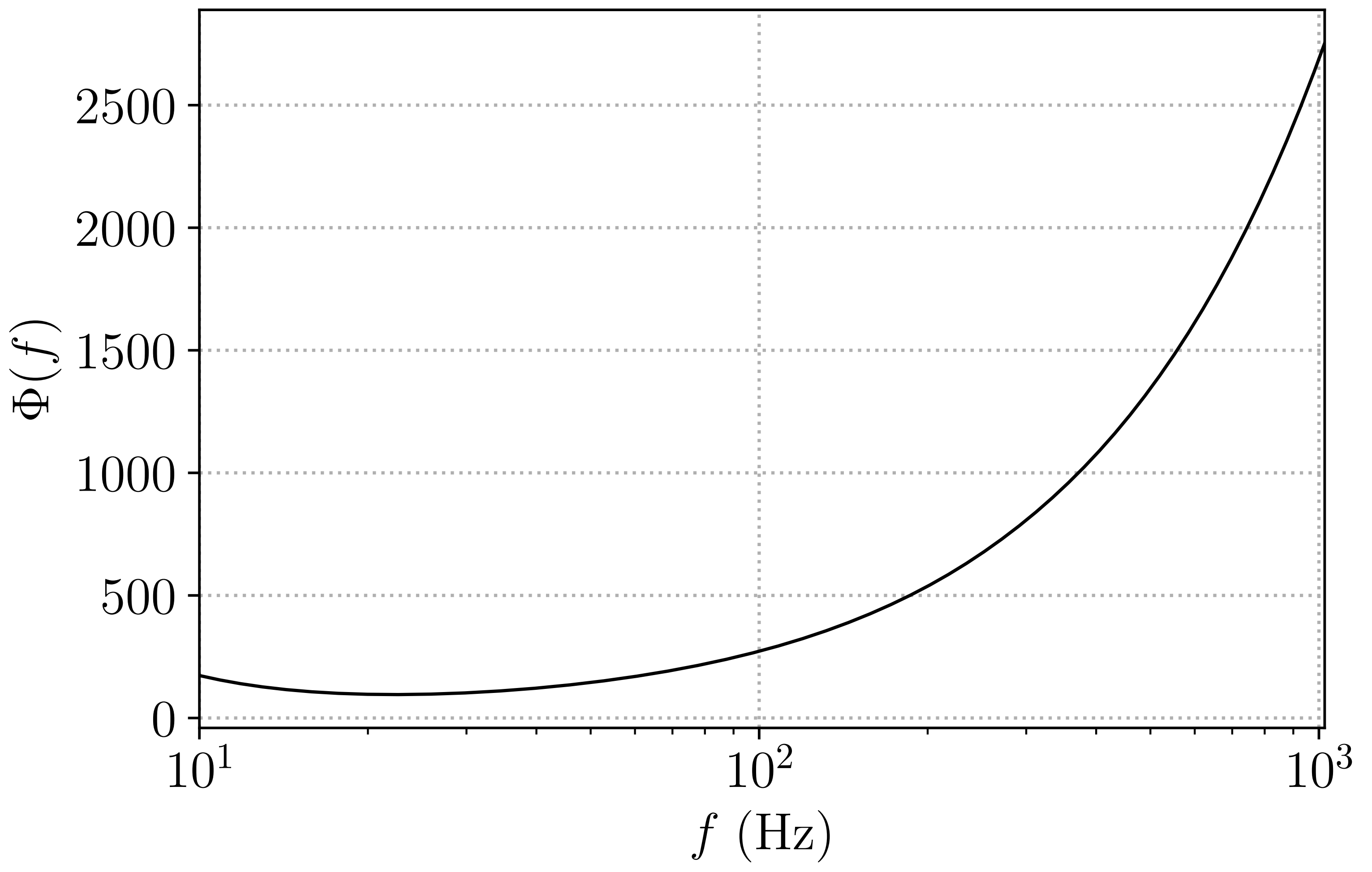
FT: $\tilde{F}(f) = \int dt\, e^{-2i\pi f t} F(t)$

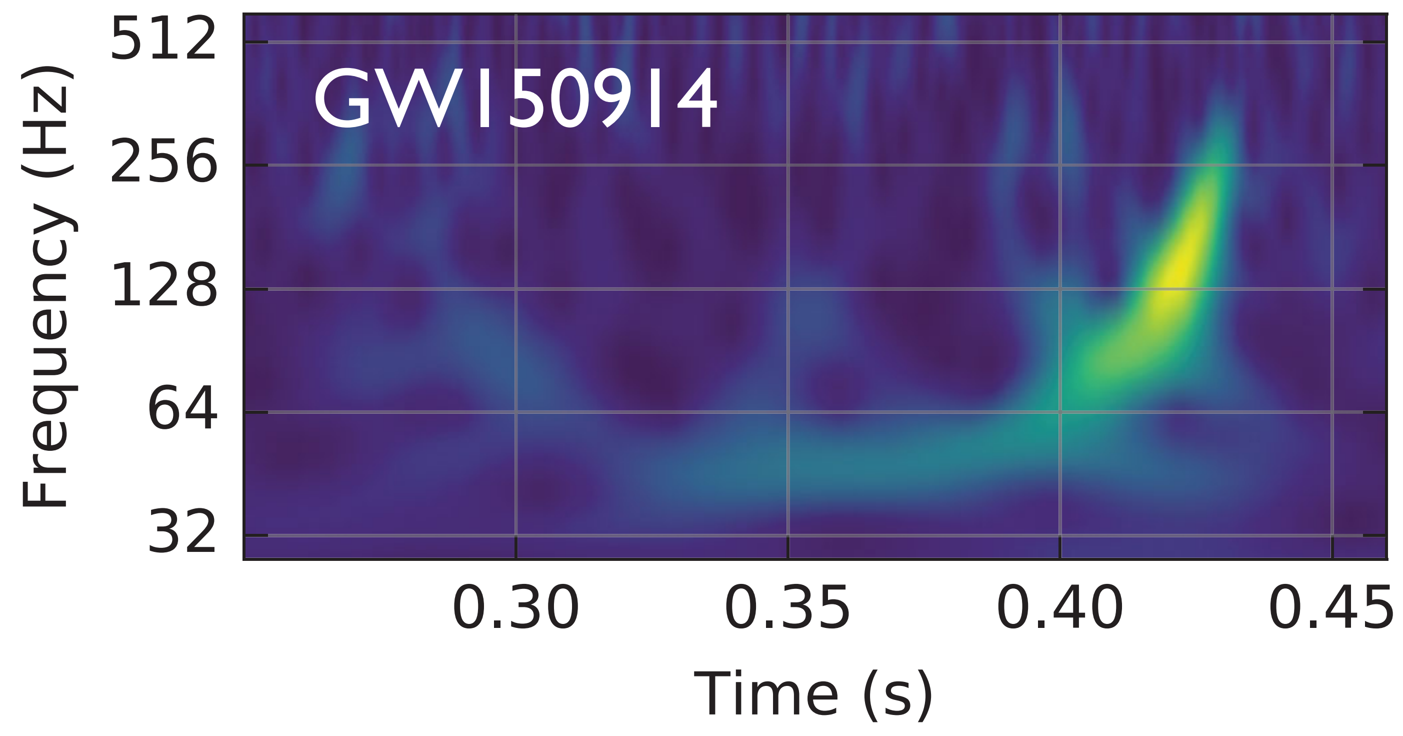
DFT: $\tilde{F}_k = \Delta t \sum e^{-2i\pi jk/N} F_j$

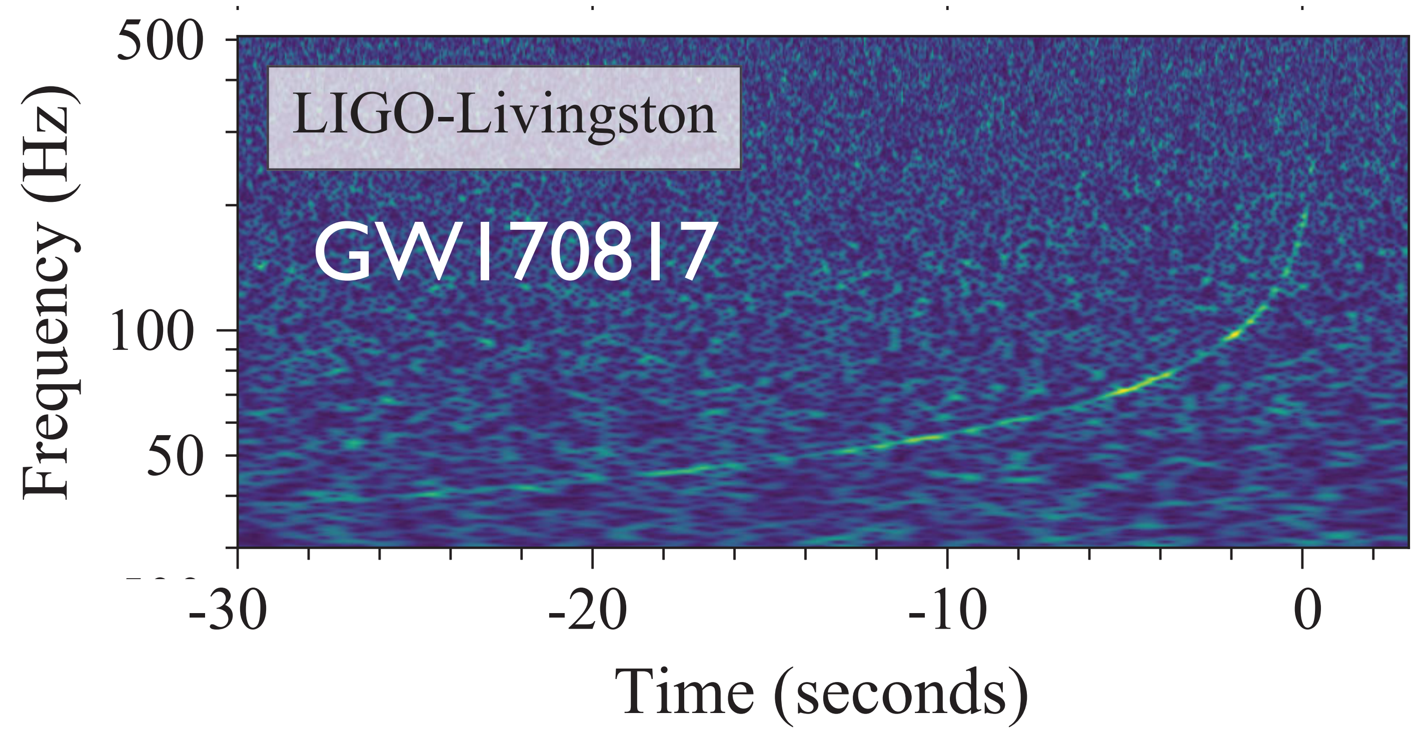
$\tilde{h} = 0$ for $f < f_{\rm Nyq}$ $\quad f_s = 2 f_{\rm Nyq}$

$\Delta t = 1/f_s \quad \Delta f = 1/T \quad N = f_s T$







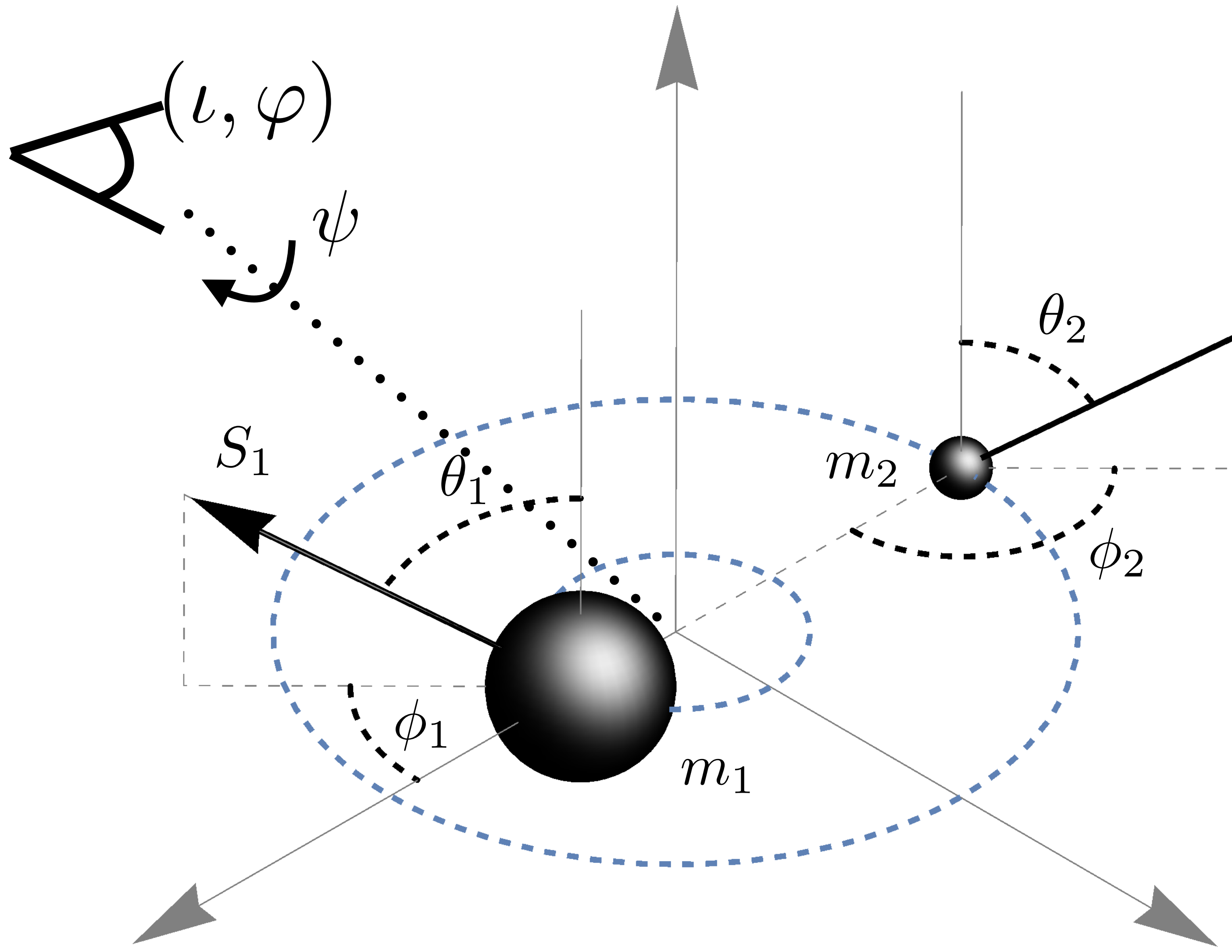


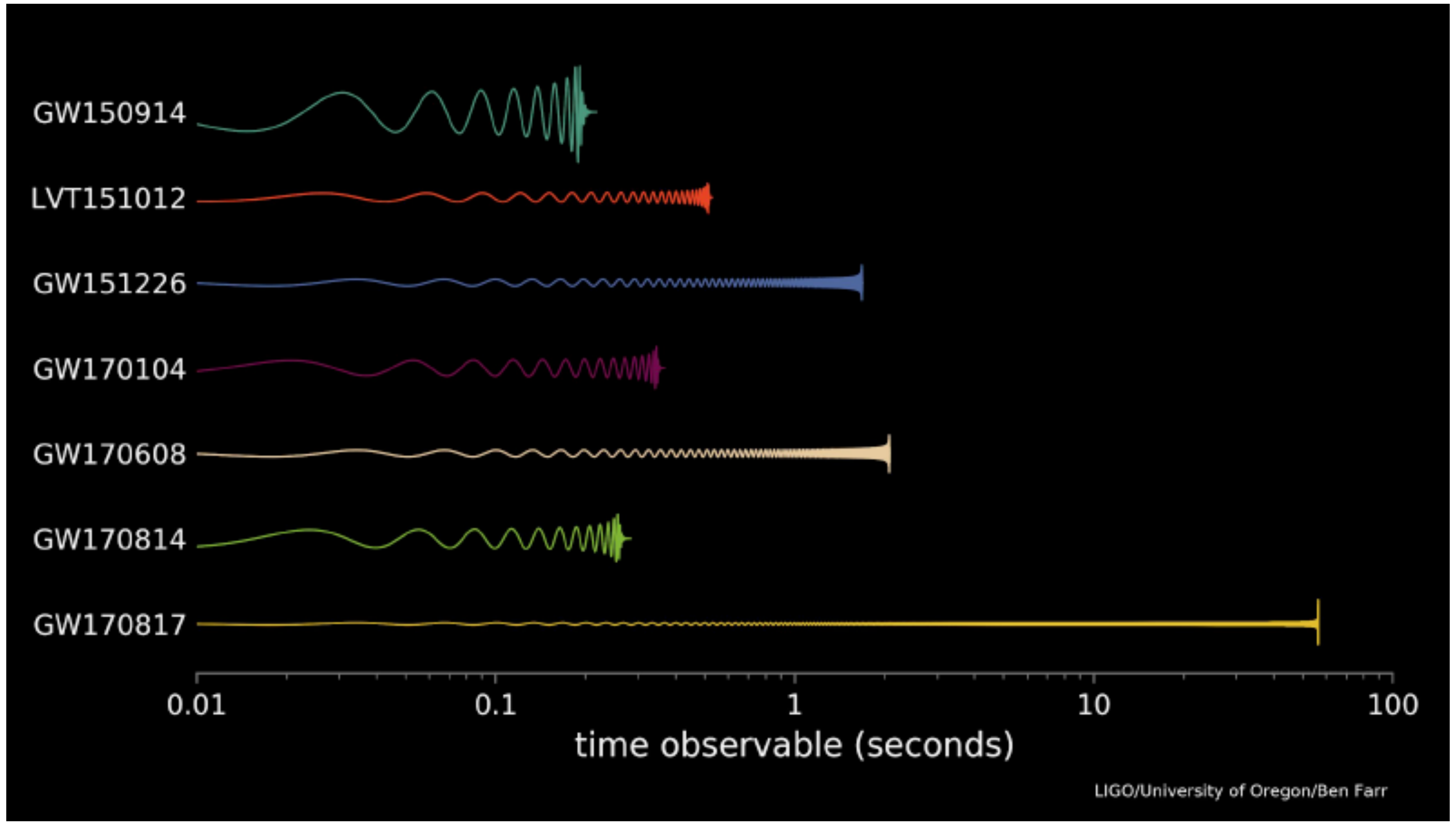
- Fourier domain: natural frequency window, simplifies the stationary noise covariance

- Inspiral: clear time-to-frequency correspondence, (Stationary Phase Approximation), merger not so

- Fourier domain not optimal for signal compression

GW Signals: CBC parameter space





[https://www.ligo.caltech.edu]

For CBC: 15+2+2 parameters

- intrinsic: 2 masses, 2*3 spin vectors
- distance: 1
- time of coalescence: 1
- direction to the observer: 2 angles
- sky position in observer's frame: 2 angles
- polarization angle: 1 angle
- +eccentricity, periastron: 2
- +tidal deformabilities BNS: 2

- BBH: massive, merger-ringdown
- BNS: inspiral dominated, tidal effects
- NSBH: high mass ratio, tidal effects ?

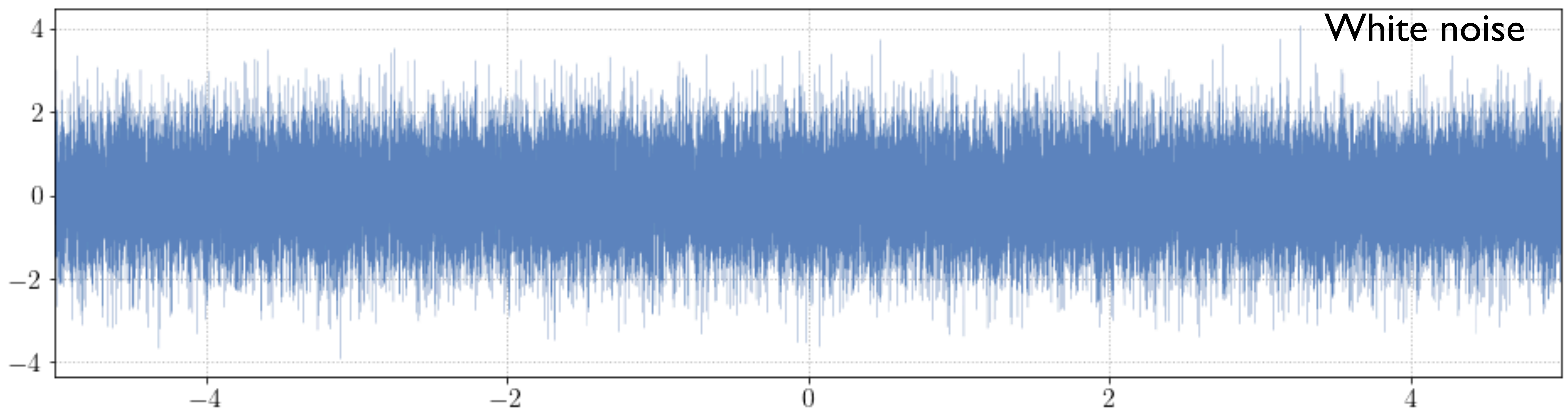
Outline

Part I

- GW signals: the basics
- **Noise as a stochastic process**
- Introducing matched filtering
- Towards real CBC searches
- Other signals: continuous waves, stochastic backgrounds

Noise

- How to understand noise as a stochastic process ?
- Ergodicity, stationarity, Gaussianity ?



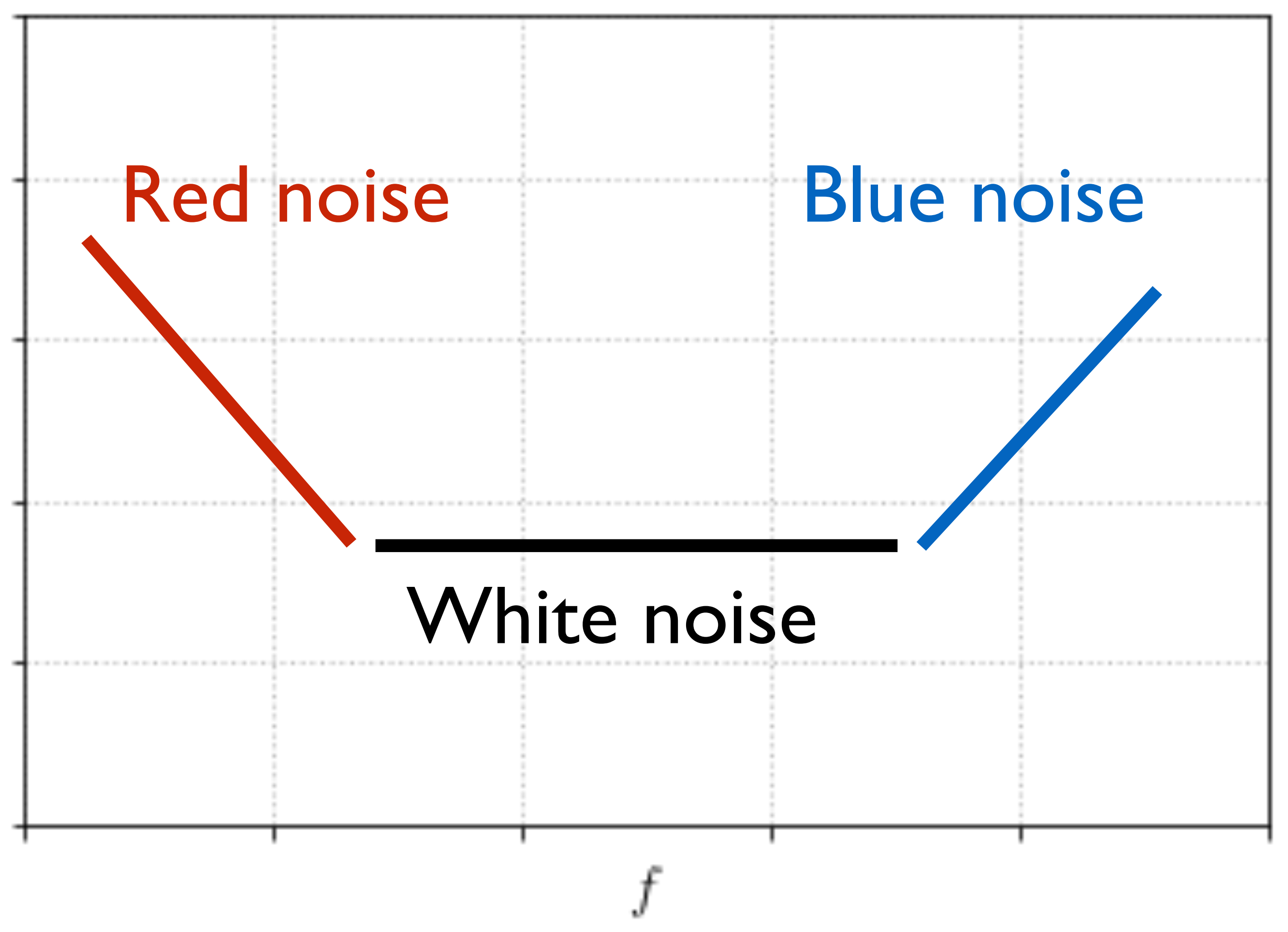
Noise autocorrelation:

$$C(t,t') = \langle n(t)n(t')\rangle$$

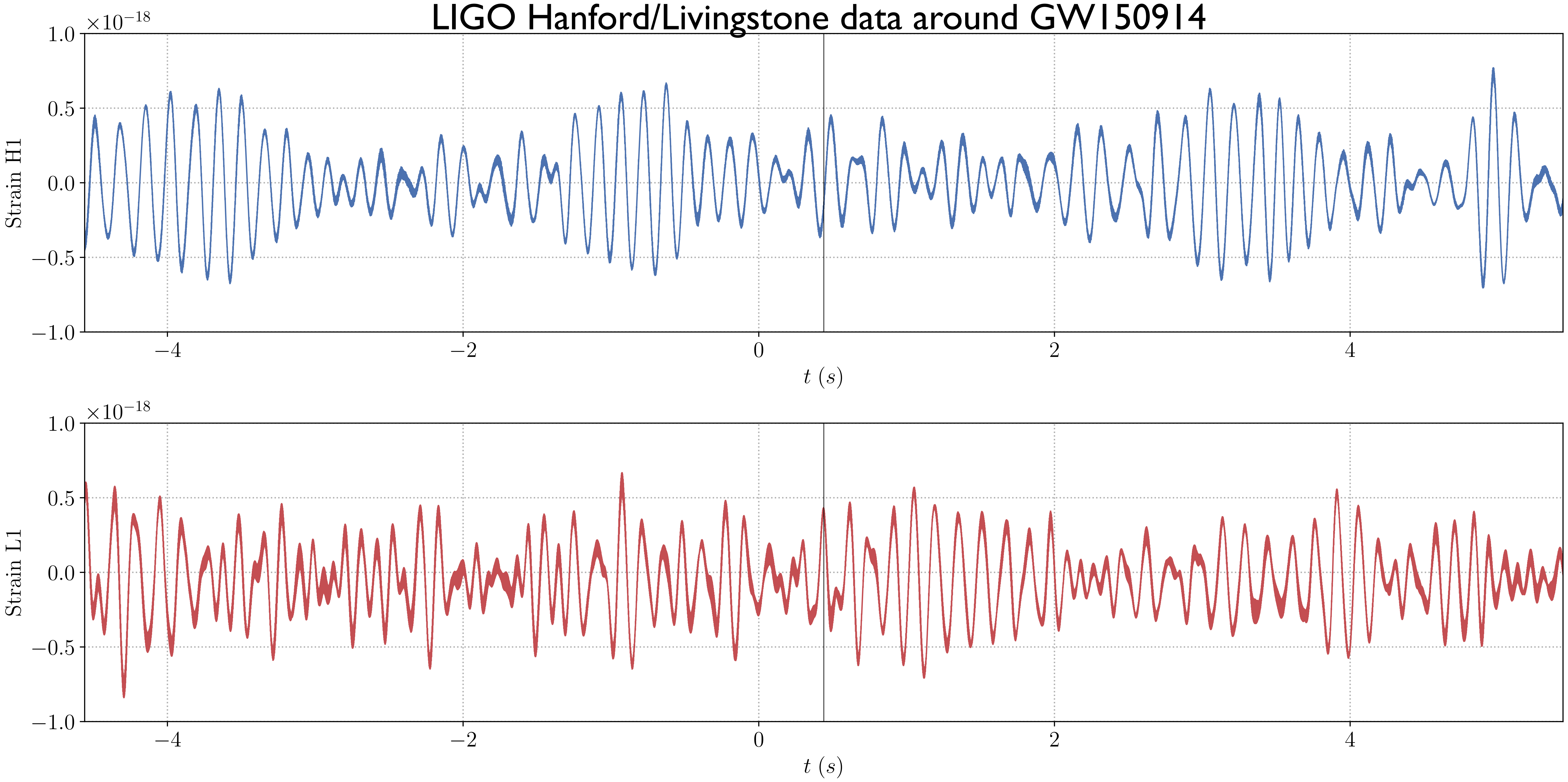
Stationary white noise:

$$C(t,t') = \mathrm{const}\,\delta(t-t')$$

Flat spectrum



Noise



- How to model real noise ?
- Ergodicity, stationarity, Gaussianity ?

Noise autocorrelation:

$$C(t,t') = \langle n(t) n(t') \rangle$$

Noise PSD

Mean power of the noise:

$$P_n = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} dt\, n(t)^2$$

Noise PSD

Mean power of the noise:

$$P_n = \lim_{T\to+\infty} \frac{1}{T}\int_{-T/2}^{T/2} dt\, n(t)^2$$

In the Fourier domain:

$$n_T(t) \equiv \chi_{[-T/2,T/2]} n(t)$$

$$\begin{aligned} P_n &= \lim_{T\to+\infty} \frac{1}{T}\int_{-\infty}^{+\infty} dt\, n_T(t)^2 \\ &= \lim_{T\to+\infty} \frac{1}{T}\int_{-\infty}^{+\infty} df\, |\tilde{n}_T(f)|^2 \\ &= \int_0^{+\infty} df\, S_n(f) \end{aligned}$$

Noise PSD

Mean power of the noise:

$$P_n = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} dt\, n(t)^2$$

In the Fourier domain:

$$n_T(t) \equiv \chi_{[-T/2,T/2]} n(t)$$

$$\begin{aligned} P_n &= \lim_{T \to +\infty} \frac{1}{T} \int_{-\infty}^{+\infty} dt\, n_T(t)^2 \\ &= \lim_{T \to +\infty} \frac{1}{T} \int_{-\infty}^{+\infty} df\, |\tilde{n}_T(f)|^2 \\ &= \int_0^{+\infty} df\, S_n(f) \end{aligned}$$

Noise Power Spectral Density (1-sided):

$$S_n(f) \equiv \lim_{T \to +\infty} \frac{2}{T} |\tilde{n}_T(f)|^2$$

Noise PSD

Mean power of the noise:

$$P_n = \lim_{T\to+\infty} \frac{1}{T}\int_{-T/2}^{T/2} dt\, n(t)^2$$

In the Fourier domain:

$$n_T(t) \equiv \chi_{[-T/2,T/2]} n(t)$$

$$\begin{aligned} P_n &= \lim_{T\to+\infty} \frac{1}{T}\int_{-\infty}^{+\infty} dt\, n_T(t)^2 \\ &= \lim_{T\to+\infty} \frac{1}{T}\int_{-\infty}^{+\infty} df\, |\tilde{n}_T(f)|^2 \\ &= \int_0^{+\infty} df\, S_n(f) \end{aligned}$$

Noise Power Spectral Density (1-sided):

$$S_n(f) \equiv \lim_{T\to+\infty} \frac{2}{T} |\tilde{n}_T(f)|^2$$

In the stationary case: $\langle\,\rangle \sim \frac{1}{T}\int dt$

$$C(t,t') = \langle n(t) n(t') \rangle$$

$$C(t,t') = C(0, t'-t) \equiv C(t'-t)$$

Noise PSD as the FT of the autocorrelation:

$$\frac{1}{2} S_n(f) = \int d\tau\, C(\tau) e^{-2i\pi f\tau}$$

Noise PSD

Mean power of the noise:

$$P_n = \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} dt\, n(t)^2$$

In the Fourier domain:

$$n_T(t) \equiv \chi_{[-T/2,T/2]} n(t)$$

$$\begin{aligned} P_n &= \lim_{T\to+\infty} \frac{1}{T} \int_{-\infty}^{+\infty} dt\, n_T(t)^2 \\ &= \lim_{T\to+\infty} \frac{1}{T} \int_{-\infty}^{+\infty} df\, |\tilde{n}_T(f)|^2 \\ &= \int_0^{+\infty} df\, S_n(f) \end{aligned}$$

Noise Power Spectral Density (1-sided):

$$S_n(f) \equiv \lim_{T\to+\infty} \frac{2}{T} |\tilde{n}_T(f)|^2$$

In the stationary case: $\langle\ \rangle \sim \frac{1}{T}\int dt$

$$C(t,t') = \langle n(t) n(t') \rangle$$

$$C(t,t') = C(0, t'-t) \equiv C(t'-t)$$

Noise PSD as the FT of the autocorrelation:

$$\frac{1}{2} S_n(f) = \int d\tau\, C(\tau) e^{-2i\pi f\tau}$$

The two definitions correspond

$$C(\tau) = \lim_{T\to+\infty} \frac{1}{T} \int dt\, n_T(t) n_T(t+\tau)$$

$$\begin{aligned} \int df\, e^{-2i\pi f\tau} C(\tau) &= \lim_{T\to+\infty} \frac{1}{T} |\tilde{n}_T(f)|^2 \\ &= \frac{1}{2} S_n(f) \end{aligned}$$

Noise stationarity

A consequence of noise stationarity: $C(t,t') = \langle n(t)n(t')\rangle$

$$C(t,t') \equiv C(t-t')$$

Noise stationarity

A consequence of noise stationarity: $C(t,t') = \langle n(t)n(t') \rangle$

$$C(t,t') \equiv C(t-t')$$

$$\langle \tilde{n}(f)\tilde{n}^*(f') \rangle = \int dt \int dt' \, e^{2i\pi f t} e^{-2i\pi f' t'} \langle n(t)n(t') \rangle$$

Noise stationarity

A consequence of noise stationarity: $C(t,t') = \langle n(t)n(t')\rangle$

$$C(t,t') \equiv C(t-t')$$

$$\langle \tilde{n}(f)\tilde{n}^*(f')\rangle = \int dt \int dt'\, e^{2i\pi f t} e^{-2i\pi f' t'} \langle n(t)n(t')\rangle$$

$$= \int dt \int d\tau\, e^{2i\pi f t} e^{-2i\pi f'(t+\tau)} \langle n(t)n(t+\tau)\rangle$$

Noise stationarity

A consequence of noise stationarity: $C(t,t') = \langle n(t)n(t')\rangle$

$$C(t,t') \equiv C(t-t')$$

$$\begin{aligned}
\langle \tilde{n}(f)\tilde{n}^*(f')\rangle &= \int dt \int dt'\, e^{2i\pi f t} e^{-2i\pi f' t'} \langle n(t)n(t')\rangle \\
&= \int dt \int d\tau\, e^{2i\pi f t} e^{-2i\pi f'(t+\tau)} \langle n(t)n(t+\tau)\rangle \\
&= \int d\tau\, C(\tau) e^{-2i\pi f'\tau} \int dt\, e^{2i\pi(f-f')t}
\end{aligned}$$

Noise stationarity

A consequence of noise stationarity: $C(t,t') = \langle n(t)n(t')\rangle$

$$C(t,t') \equiv C(t-t')$$

$$\begin{aligned}
\langle \tilde{n}(f)\tilde{n}^*(f')\rangle &= \int dt \int dt'\, e^{2i\pi f t} e^{-2i\pi f' t'} \langle n(t)n(t')\rangle \\
&= \int dt \int d\tau\, e^{2i\pi f t} e^{-2i\pi f'(t+\tau)} \langle n(t)n(t+\tau)\rangle \\
&= \int d\tau\, C(\tau) e^{-2i\pi f'\tau} \int dt\, e^{2i\pi (f-f')t} \\
&= \delta(f-f') \int d\tau\, C(\tau) e^{-2i\pi f'\tau}
\end{aligned}$$

Noise stationarity

A consequence of noise stationarity: $C(t,t') = \langle n(t)n(t')\rangle$

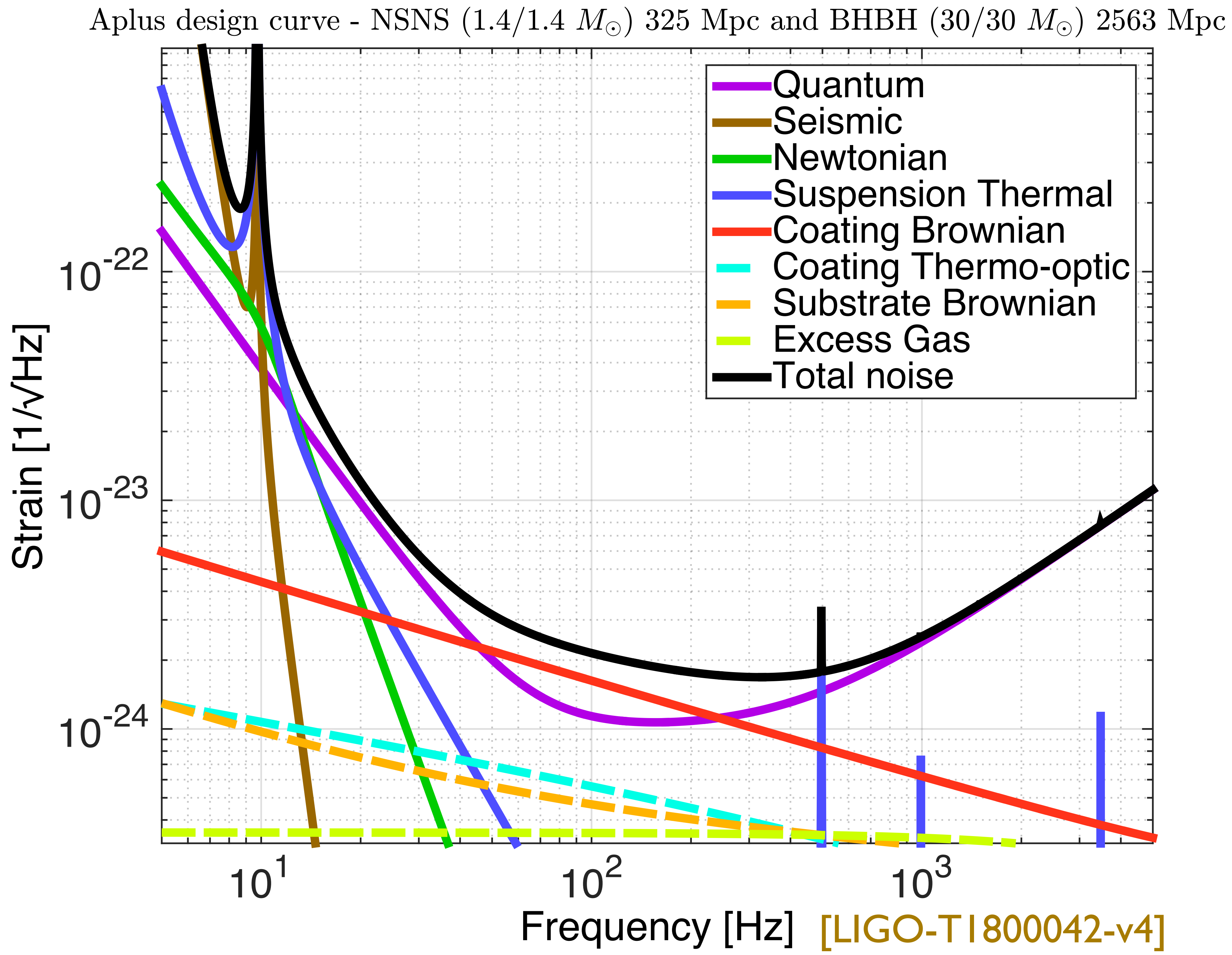
$$C(t,t') \equiv C(t-t')$$

$$\begin{aligned}
\langle \tilde{n}(f)\tilde{n}^*(f')\rangle &= \int dt \int dt'\, e^{2i\pi f t} e^{-2i\pi f' t'} \langle n(t)n(t')\rangle \\
&= \int dt \int d\tau\, e^{2i\pi f t} e^{-2i\pi f'(t+\tau)} \langle n(t)n(t+\tau)\rangle \\
&= \int d\tau\, C(\tau) e^{-2i\pi f'\tau} \int dt\, e^{2i\pi (f-f')t} \\
&= \delta(f-f') \int d\tau\, C(\tau) e^{-2i\pi f'\tau} \\
&= \delta(f-f') \frac{1}{2} S_n(f)
\end{aligned}$$

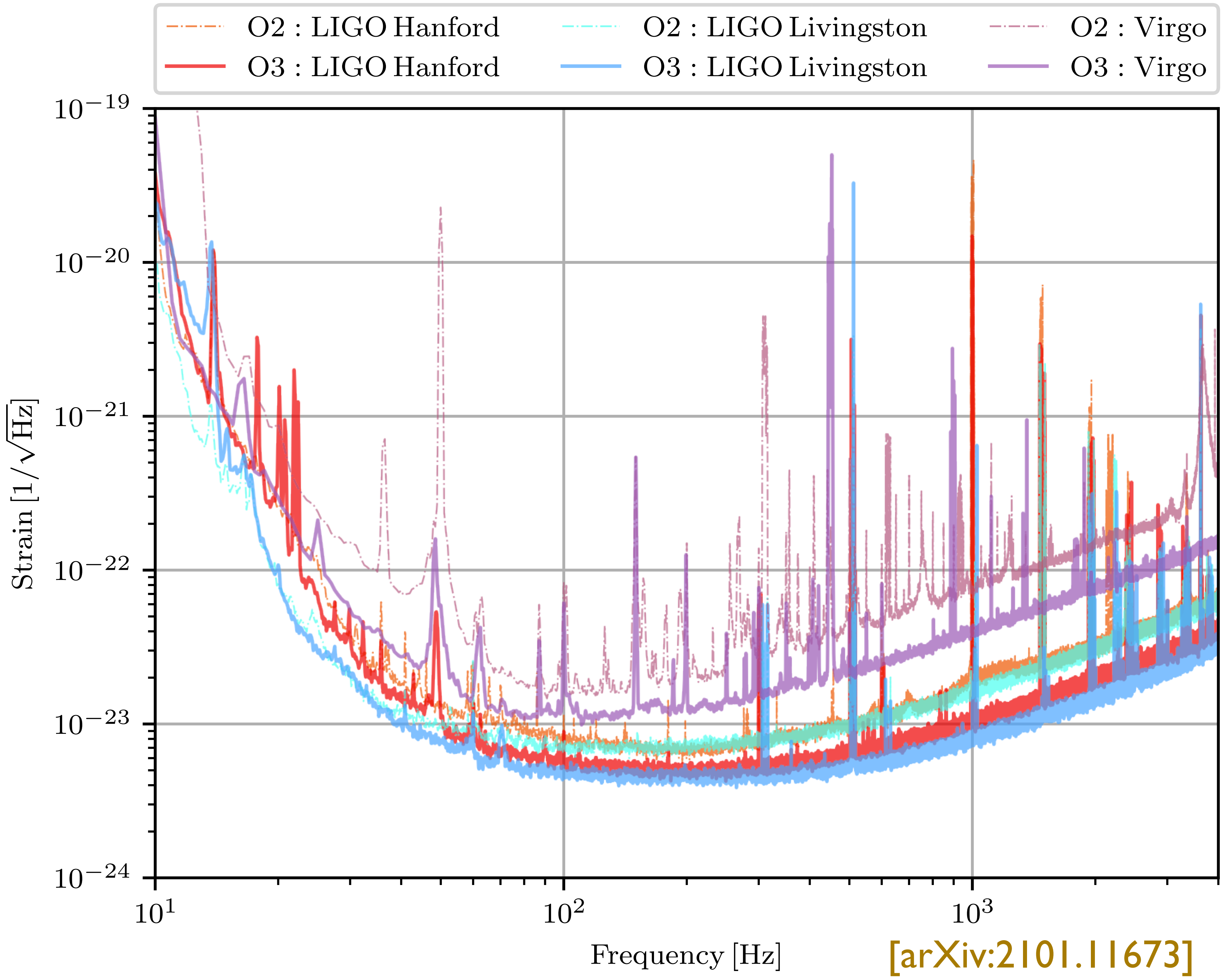
Noise stationarity means independence in Fourier domain !

In practice, stationarity is always approximate…

Noise PSD



[LIGO-T1800042-v4]

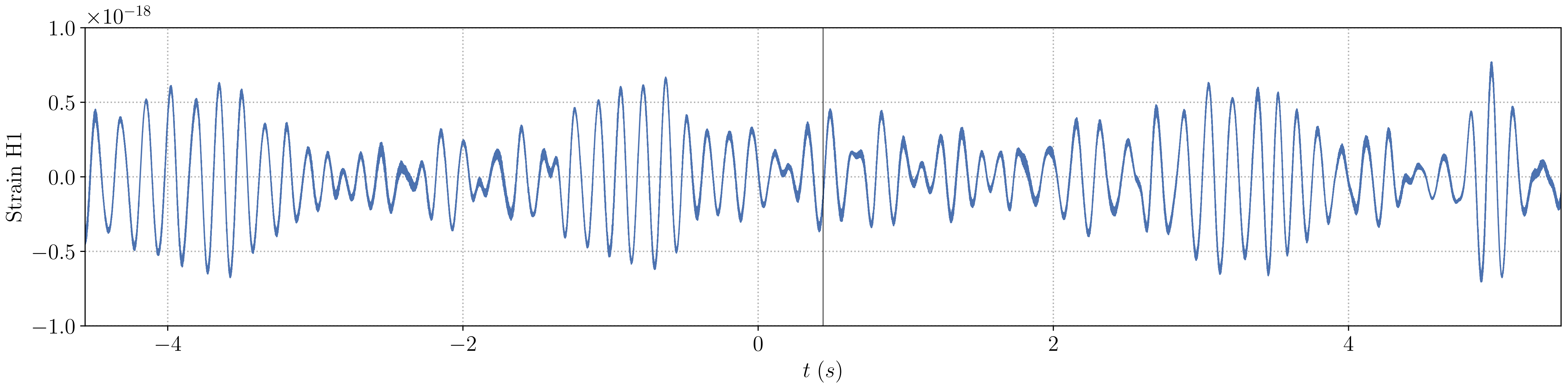


[arXiv:2101.11673]

- Different processes dominate red/white/blue noise
- PSD from real LVK data: lines, drifts over time
- PSD estimation method: average over segments (Welch)

Gaussian noise

$$n(t) \to \mathbf{n} \in \mathbb{R}^N$$



For a Gaussian process:

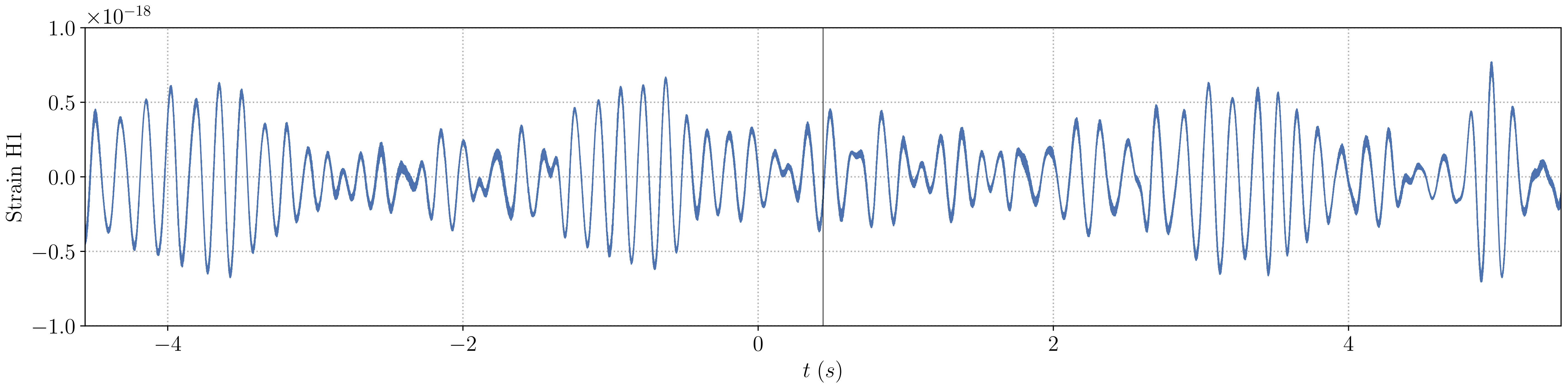
$$p(\mathbf{n}) = \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left[-\frac{1}{2}\mathbf{n}^T \cdot \boldsymbol{\Sigma}^{-1} \cdot \mathbf{n}\right]$$

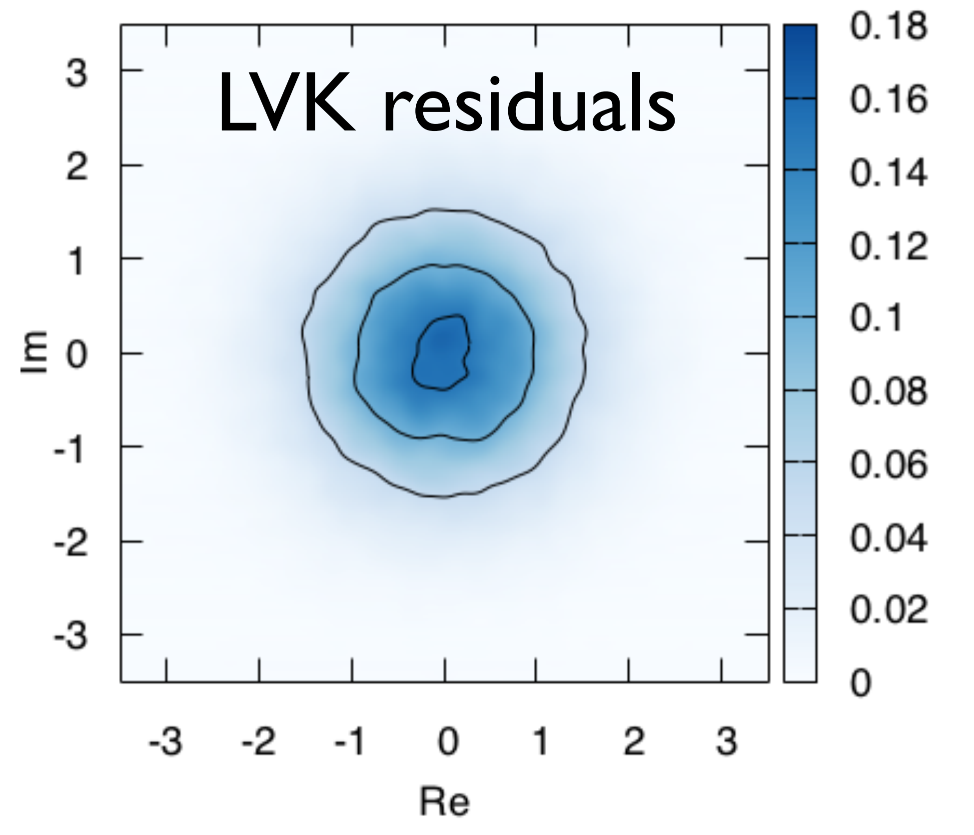
In Fourier domain (DFT):

$$p(\tilde{\mathbf{n}}) = \frac{1}{\sqrt{(2\pi)^N \det \tilde{\boldsymbol{\Sigma}}}} \exp\left[-\frac{1}{2}\tilde{\mathbf{n}}^T \cdot \tilde{\boldsymbol{\Sigma}}^{-1} \cdot \tilde{\mathbf{n}}\right]$$

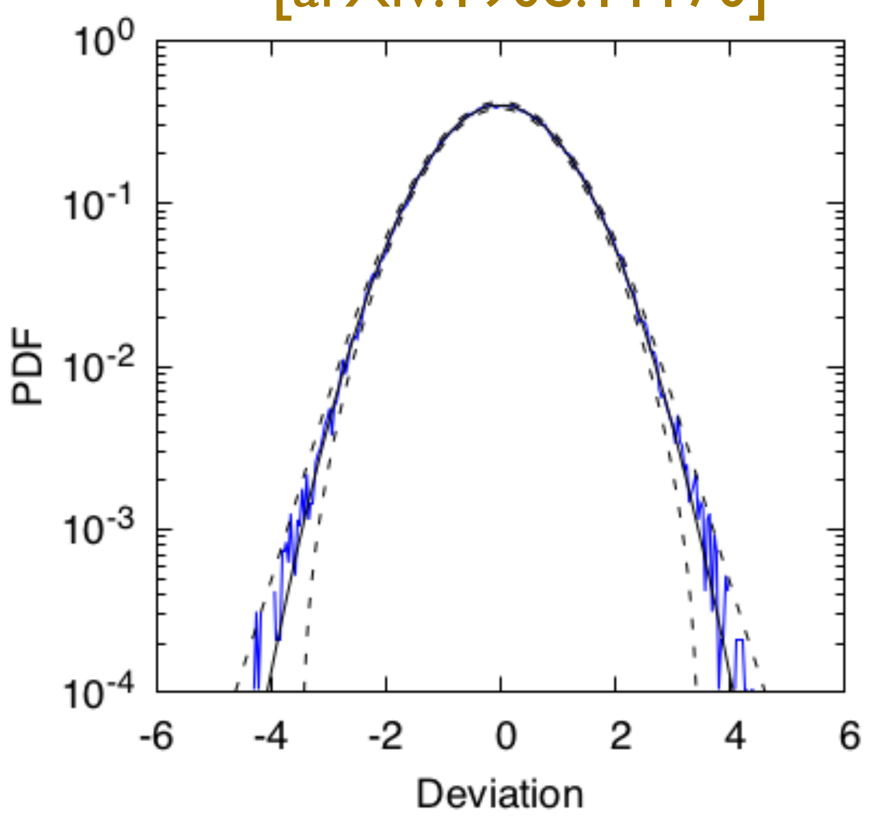
Gaussian noise

$$n(t) \to \mathbf{n} \in \mathbb{R}^N$$





[arXiv:1908.11170]



For a Gaussian process:

$$p(\mathbf{n}) = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{\Sigma}}} \exp\left[-\frac{1}{2}\mathbf{n}^T \cdot \mathbf{\Sigma}^{-1} \cdot \mathbf{n}\right]$$

In Fourier domain (DFT):

$$p(\tilde{\mathbf{n}}) = \frac{1}{\sqrt{(2\pi)^N \det \tilde{\mathbf{\Sigma}}}} \exp\left[-\frac{1}{2}\tilde{\mathbf{n}}^T \cdot \tilde{\mathbf{\Sigma}}^{-1} \cdot \tilde{\mathbf{n}}\right]$$

For a stationary Gaussian process: independence FD, diagonal covariance

$$\langle \tilde{n}_k \tilde{n}_l^* \rangle = \frac{1}{2\Delta f} S_n(f_k) \delta_{kl}$$

$$\mathrm{Re}\, \tilde{n}_k,\ \mathrm{Im}\, \tilde{n}_k \sim \mathcal{N}\left(0, \frac{1}{4\Delta f} S_n(f_k)\right)$$

From NxN to N !

Outline

Part I

- GW signals: the basics
- Noise as a stochastic process
- **Introducing matched filtering**
- Towards real CBC searches
- Other signals: continuous waves, stochastic backgrounds

Matched filter I

Idea: correlating template with data

$$s(t) = h(t) + n(t) \qquad \langle n \rangle = 0$$

$$\frac{1}{T}\int dt\, h(t)s(t) = \underbrace{\frac{1}{T}\int dt\, h(t)^2}_{\substack{\text{coherent} \\ \sim\, \mathrm{const}}} + \underbrace{\frac{1}{T}\int dt\, h(t)n(t)}_{\substack{\text{incoherent} \\ \sim\, 1/\sqrt{T}}}$$

Matched filter I

Idea: correlating template with data

$$s(t) = h(t) + n(t) \qquad \langle n \rangle = 0$$

$$\frac{1}{T}\int dt\, h(t)s(t) = \frac{1}{T}\int dt\, h(t)^2 + \frac{1}{T}\int dt\, h(t)n(t)$$

coherent $\sim \mathrm{const}$ incoherent $\sim 1/\sqrt{T}$

In signal:

$$\hat{s} \equiv \int dt\, W(t)s(t)$$

$$S = \langle \hat{s} \rangle$$

In noise:

$$\hat{n} \equiv \int dt\, W(t)n(t)$$

$$N^2 = \langle \hat{n}^2 \rangle \qquad \langle \hat{n} \rangle = 0$$

Build filter $W(t)$ to optimize S/N

Matched filter I

Idea: correlating template with data

$$s(t) = h(t) + n(t) \qquad \langle n \rangle = 0$$

$$\frac{1}{T}\int dt\, h(t)s(t) = \frac{1}{T}\int dt\, h(t)^2 + \frac{1}{T}\int dt\, h(t)n(t)$$

coherent $\sim \mathrm{const}$ incoherent $\sim 1/\sqrt{T}$

In signal: In noise:

$$\hat{s} \equiv \int dt\, W(t)s(t) \qquad \hat{n} \equiv \int dt\, W(t)n(t)$$

$$S = \langle \hat{s} \rangle \qquad N^2 = \langle \hat{n}^2 \rangle \qquad \langle \hat{n} \rangle = 0$$

Build filter $W(t)$ to optimize S/N

$$S = \int dt\, W(t)h(t) + \cancel{\langle \int dt\, W(t)n(t) \rangle}$$

$$= \int df\, \tilde{W}^*(f)\tilde{h}(f)$$

Matched filter I

Idea: correlating template with data

$$s(t) = h(t) + n(t) \qquad \langle n \rangle = 0$$

$$\frac{1}{T}\int dt\, h(t)s(t) = \frac{1}{T}\int dt\, h(t)^2 + \frac{1}{T}\int dt\, h(t)n(t)$$

coherent $\sim \text{const}$ — incoherent $\sim 1/\sqrt{T}$

In signal:

$$\hat{s} \equiv \int dt\, W(t)s(t)$$

$$S = \langle \hat{s} \rangle$$

In noise:

$$\hat{n} \equiv \int dt\, W(t)n(t)$$

$$N^2 = \langle \hat{n}^2 \rangle \qquad \langle \hat{n} \rangle = 0$$

Build filter $W(t)$ to optimize S/N

$$S = \int dt\, W(t)h(t) + \cancel{\langle \int dt\, W(t)n(t) \rangle}$$

$$= \int df\, \tilde{W}^*(f)\tilde{h}(f)$$

$$N^2 = \langle \int dt\, dt'\, W(t)W(t')n(t)n(t') \rangle$$

Matched filter I

Idea: correlating template with data

$$s(t) = h(t) + n(t) \qquad \langle n \rangle = 0$$

$$\frac{1}{T}\int dt\, h(t)s(t) = \frac{1}{T}\int dt\, h(t)^2 + \frac{1}{T}\int dt\, h(t)n(t)$$

coherent $\sim \text{const}$ | incoherent $\sim 1/\sqrt{T}$

In signal:

$$\hat{s} \equiv \int dt\, W(t)s(t)$$

$$S = \langle \hat{s} \rangle$$

In noise:

$$\hat{n} \equiv \int dt\, W(t)n(t)$$

$$N^2 = \langle \hat{n}^2 \rangle \qquad \langle \hat{n} \rangle = 0$$

Build filter $W(t)$ to optimize S/N

$$S = \int dt\, W(t)h(t) + \cancel{\langle \int dt\, W(t)n(t) \rangle}$$

$$= \int df\, \tilde{W}^*(f)\tilde{h}(f)$$

$$N^2 = \langle \int dt\, dt'\, W(t)W(t')n(t)n(t') \rangle$$

$$= \int df\, df'\, \tilde{W}^*(f)\tilde{W}(f')\langle \tilde{n}(f)\tilde{n}^*(f') \rangle$$

Matched filter I

Idea: correlating template with data

$$s(t) = h(t) + n(t) \qquad \langle n \rangle = 0$$

$$\frac{1}{T}\int dt\, h(t)s(t) = \frac{1}{T}\int dt\, h(t)^2 + \frac{1}{T}\int dt\, h(t)n(t)$$

coherent $\sim \mathrm{const}$

incoherent $\sim 1/\sqrt{T}$

In signal:

$$\hat{s} \equiv \int dt\, W(t)s(t)$$

$$S = \langle \hat{s} \rangle$$

In noise:

$$\hat{n} \equiv \int dt\, W(t)n(t)$$

$$N^2 = \langle \hat{n}^2 \rangle \qquad \langle \hat{n} \rangle = 0$$

Build filter $W(t)$ to optimize S/N

$$S = \int dt\, W(t)h(t) + \cancel{\langle \int dt\, W(t)n(t) \rangle}$$
$$= \int df\, \tilde{W}^*(f)\tilde{h}(f)$$

$$N^2 = \langle \int dt\, dt'\, W(t)W(t')n(t)n(t') \rangle$$
$$= \int df\, df'\, \tilde{W}^*(f)\tilde{W}(f')\langle \tilde{n}(f)\tilde{n}^*(f') \rangle$$
$$= \int df\, df'\, \tilde{W}^*(f)\tilde{W}(f')\frac{1}{2}S_n(f)\delta(f-f')$$

Matched filter I

Idea: correlating template with data

$$s(t) = h(t) + n(t) \qquad \langle n \rangle = 0$$

$$\frac{1}{T}\int dt\, h(t)s(t) = \underbrace{\frac{1}{T}\int dt\, h(t)^2}_{\substack{\text{coherent} \\ \sim\, \mathrm{const}}} + \underbrace{\frac{1}{T}\int dt\, h(t)n(t)}_{\substack{\text{incoherent} \\ \sim\, 1/\sqrt{T}}}$$

In signal:

$$\hat{s} \equiv \int dt\, W(t)s(t)$$

$$S = \langle \hat{s} \rangle$$

In noise:

$$\hat{n} \equiv \int dt\, W(t)n(t)$$

$$N^2 = \langle \hat{n}^2 \rangle \qquad \langle \hat{n} \rangle = 0$$

Build filter $W(t)$ to optimize S/N

$$S = \int dt\, W(t)h(t) + \cancel{\langle \int dt\, W(t)n(t) \rangle}$$
$$= \int df\, \tilde{W}^*(f)\tilde{h}(f)$$

$$N^2 = \langle \int dt\, dt'\, W(t)W(t')n(t)n(t') \rangle$$
$$= \int df\, df'\, \tilde{W}^*(f)\tilde{W}(f')\langle \tilde{n}(f)\tilde{n}^*(f') \rangle$$
$$= \int df\, df'\, \tilde{W}^*(f)\tilde{W}(f')\frac{1}{2}S_n(f)\delta(f-f')$$
$$= \int df\, \frac{1}{2}S_n(f)|\tilde{W}(f)|^2$$

Matched filter II

Introduce a noise-weighted inner product:

$$\boxed{(a|b) \equiv 4\mathrm{Re} \int_0^{+\infty} \frac{df}{S_n(f)} \tilde{a}(f)\tilde{b}^*(f)}$$

Redefine:

$$\tilde{u}(f) \equiv \frac{1}{2} S_n(f) \tilde{W}(f)$$

Matched filter II

Introduce a noise-weighted inner product:

$$\boxed{(a|b) \equiv 4\mathrm{Re} \int_0^{+\infty} \frac{df}{S_n(f)} \tilde{a}(f)\tilde{b}^*(f)}$$

Redefine:

$$\tilde{u}(f) \equiv \frac{1}{2} S_n(f)\tilde{W}(f)$$

Simpler expressions:

$$S = (u|h) \qquad N^2 = (u|u)$$

$$\frac{S}{N} = \frac{(u|h)}{\sqrt{(u|u)}}$$

Optimization, Wiener filter:

$$u \propto \tilde{h}$$

$$\tilde{W}(f) \equiv 2\tilde{h}(f)/S_n(f)$$

Matched filter II

Introduce a noise-weighted inner product:

$$\boxed{(a|b) \equiv 4\mathrm{Re} \int_0^{+\infty} \frac{df}{S_n(f)} \tilde{a}(f)\tilde{b}^*(f)}$$

Redefine:

$$\tilde{u}(f) \equiv \frac{1}{2} S_n(f) \tilde{W}(f)$$

Simpler expressions:

$$S = (u|h) \qquad N^2 = (u|u)$$

$$\frac{S}{N} = \frac{(u|h)}{\sqrt{(u|u)}}$$

Optimization, Wiener filter:

$$u \propto \tilde{h}$$

$$\tilde{W}(f) \equiv 2\tilde{h}(f)/S_n(f)$$

Using the Wiener filter on data:

$$N^2 = (h|h)$$

$$\hat{s} = (h|s)$$

Matched filter SNR:

$$\rho = \frac{\hat{s}}{N} = \frac{(h|s)}{\sqrt{(h|h)}}$$

Matched filter II

Introduce a noise-weighted inner product:

$$\boxed{(a|b) \equiv 4\mathrm{Re}\int_0^{+\infty} \frac{df}{S_n(f)}\tilde{a}(f)\tilde{b}^*(f)}$$

Redefine:

$$\tilde{u}(f) \equiv \frac{1}{2}S_n(f)\tilde{W}(f)$$

Simpler expressions:

$$S = (u|h) \qquad N^2 = (u|u)$$

$$\frac{S}{N} = \frac{(u|h)}{\sqrt{(u|u)}}$$

Optimization, Wiener filter:

$$u \propto \tilde{h}$$

$$\tilde{W}(f) \equiv 2\tilde{h}(f)/S_n(f)$$

Using the Wiener filter on data:

$$N^2 = (h|h)$$

$$\hat{s} = (h|s)$$

Matched filter SNR:

$$\rho = \frac{\hat{s}}{N} = \frac{(h|s)}{\sqrt{(h|h)}}$$

In noise: $\rho = \hat{n}/N \sim \mathcal{N}(0,1)$

For signal: $\rho \sim \mathcal{N}(\overline{\rho},1)$ (perfect template)

$\overline{\rho} = \sqrt{(h|h)}$ optimal SNR

Matched filter II

Introduce a noise-weighted inner product:

$$\boxed{(a|b) \equiv 4\mathrm{Re} \int_0^{+\infty} \frac{df}{S_n(f)} \tilde{a}(f)\tilde{b}^*(f)}$$

Redefine:

$$\tilde{u}(f) \equiv \frac{1}{2} S_n(f)\tilde{W}(f)$$

Simpler expressions:

$$S = (u|h) \qquad N^2 = (u|u)$$

$$\frac{S}{N} = \frac{(u|h)}{\sqrt{(u|u)}}$$

Optimization, Wiener filter:

$$u \propto \tilde{h}$$

$$\tilde{W}(f) \equiv 2\tilde{h}(f)/S_n(f)$$

Using the Wiener filter on data:

$$N^2 = (h|h)$$

$$\hat{s} = (h|s)$$

Matched filter SNR:

$$\rho = \frac{\hat{s}}{N} = \frac{(h|s)}{\sqrt{(h|h)}}$$

In noise: $\rho = \hat{n}/N \sim \mathcal{N}(0, 1)$

For signal: $\rho \sim \mathcal{N}(\overline{\rho}, 1)$ (perfect template)

$\overline{\rho} = \sqrt{(h|h)}$ optimal SNR

In practice, multiple templates and optimize over time and phase

$$\tilde{h}_{\Delta t, \Delta\phi}(f) = e^{-2i\pi f \Delta t} e^{i\phi} \tilde{h}(f)$$

Matched filtering SNR

Optimization over phase, 2 d.o.f.:

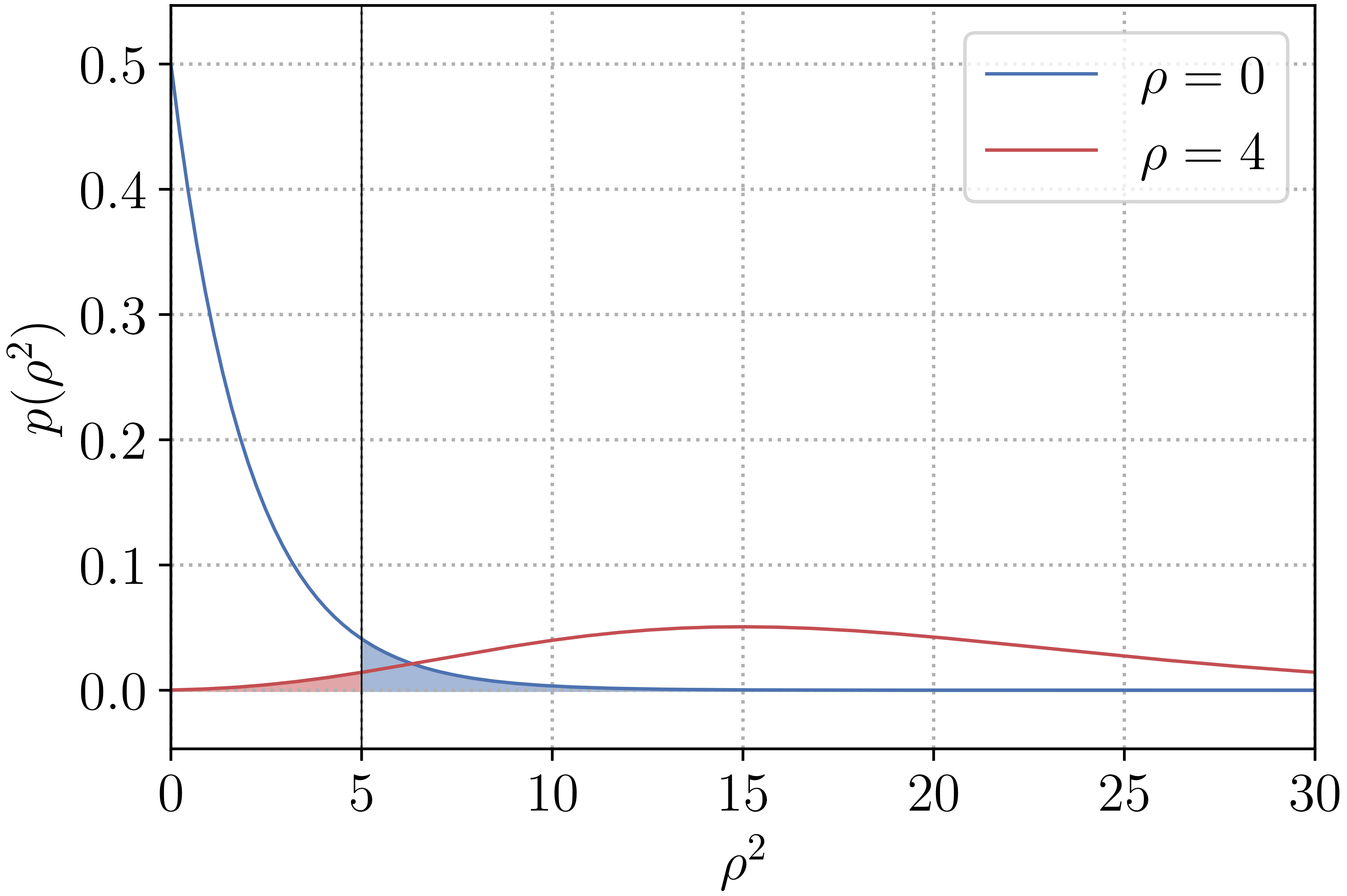
$$\rho^2 = \rho_c^2 + \rho_p^2$$

Distribution (chi2, noncentered):

$$R \equiv \rho^2$$

$$p(R|\overline{R}) = \frac{1}{2} e^{-(R+\overline{R})/2} I_0\left(\sqrt{R\overline{R}}\right)$$

$$p(R|0) = \frac{1}{2} e^{-R/2}$$



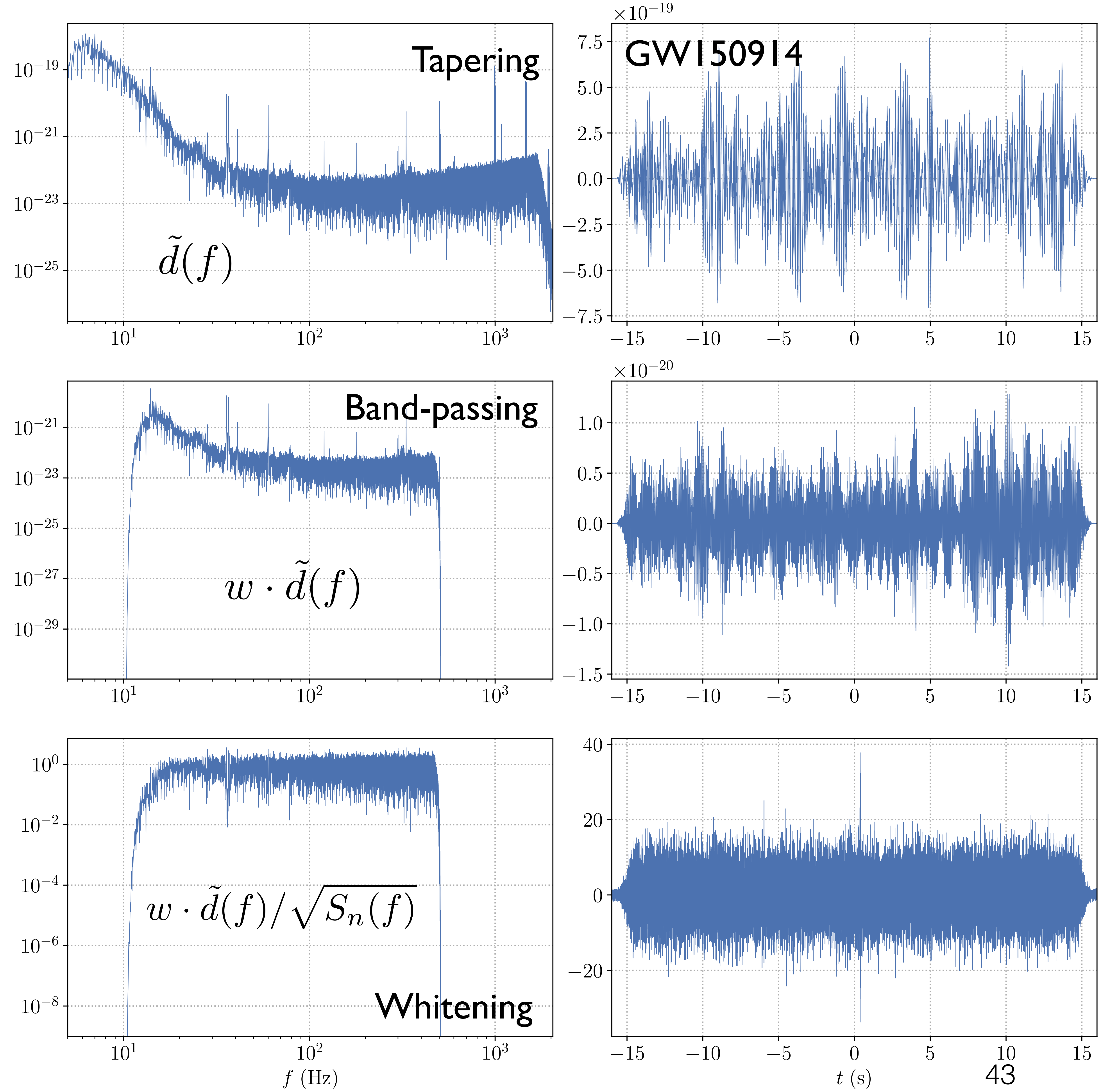
Rough estimates:

- templates in bank: $\sim 10^5$
- values of time / yr: $\sim 10^{10}$
- for a FAR ~ 1/yr: $\rho_t \sim 8$ single det.

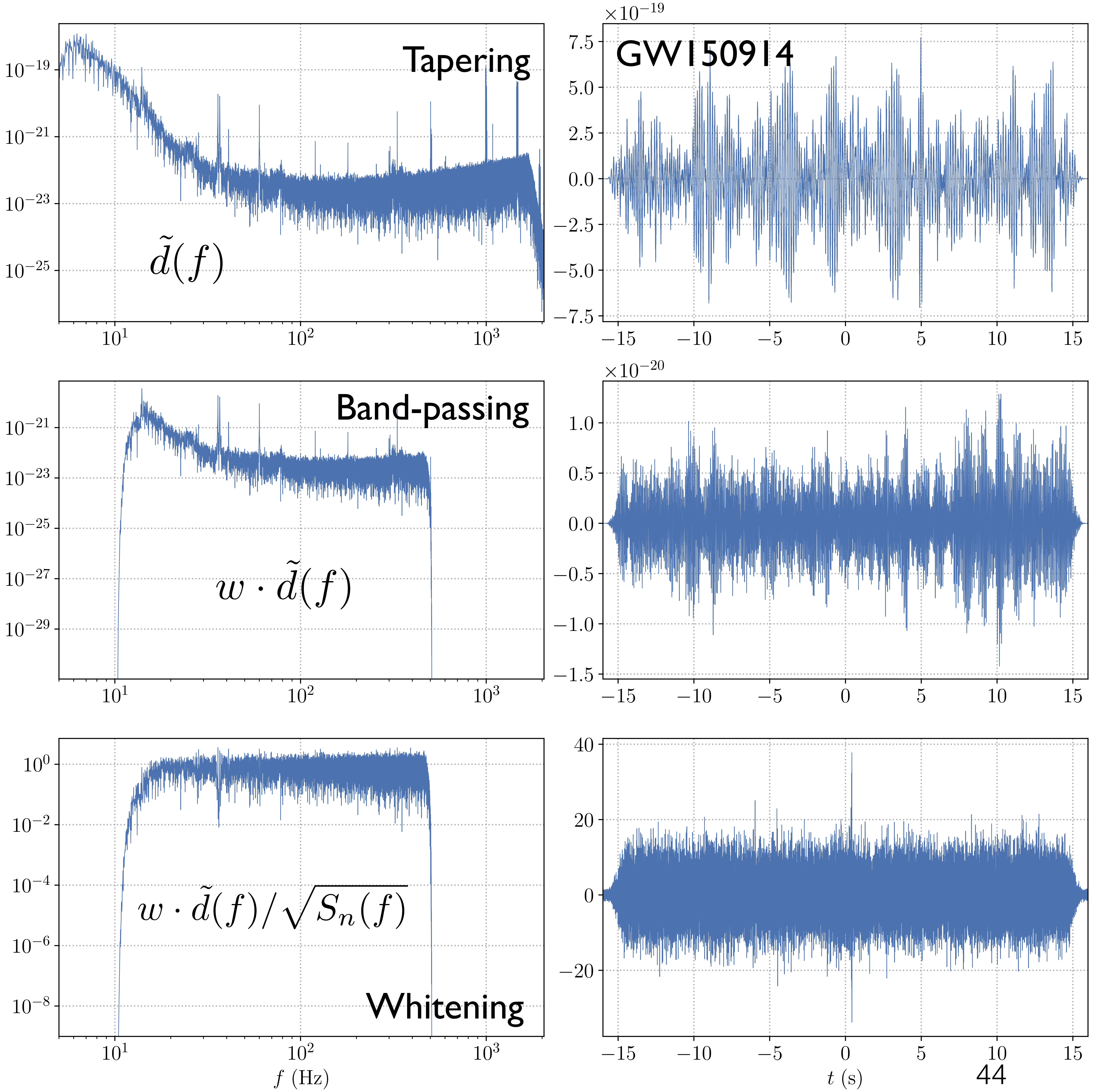
  $\rho_t \sim 5.5$ two det.

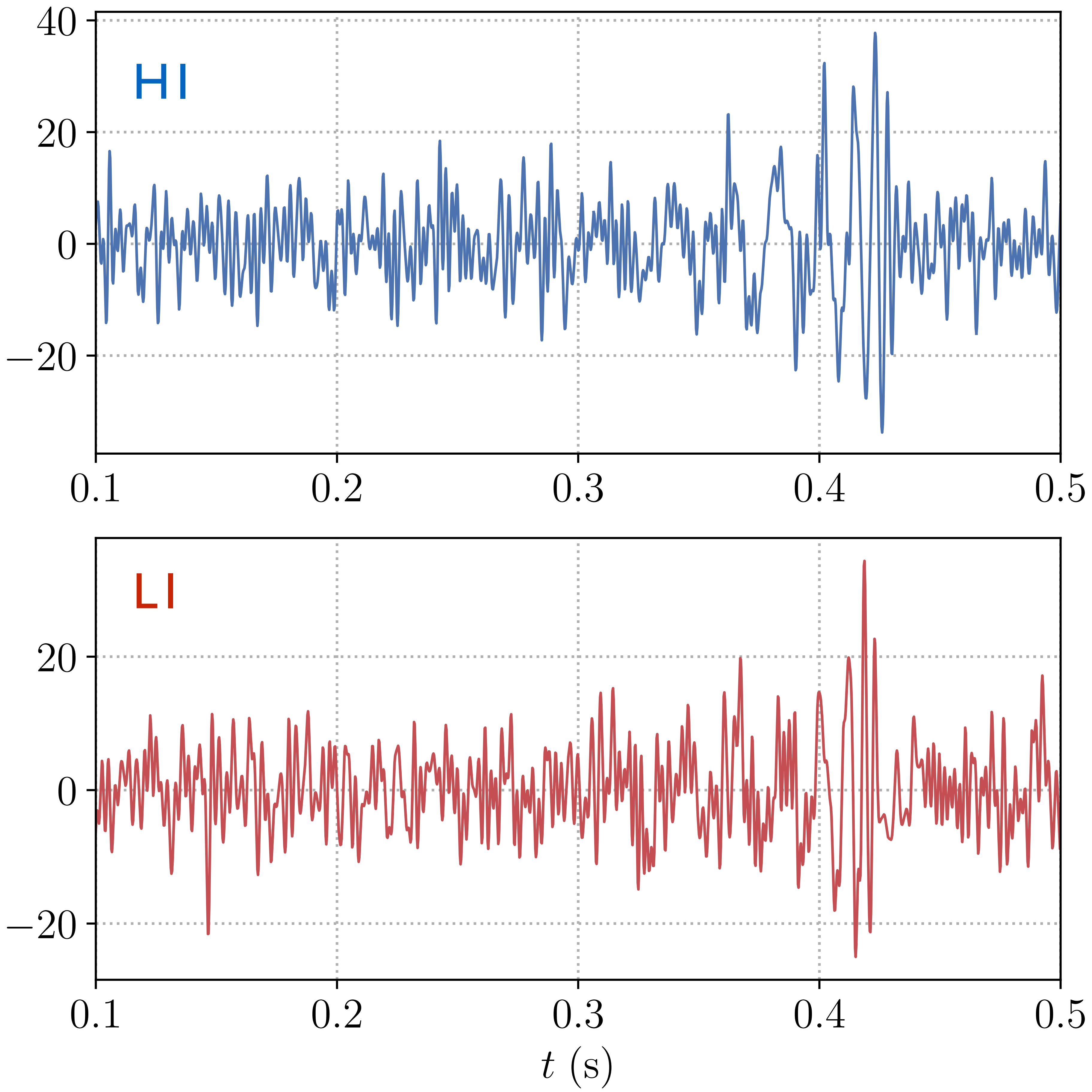
Thresholding: tradeoff between false alarms and false dismissals

Whitening, band-passing



Whitening, band-passing





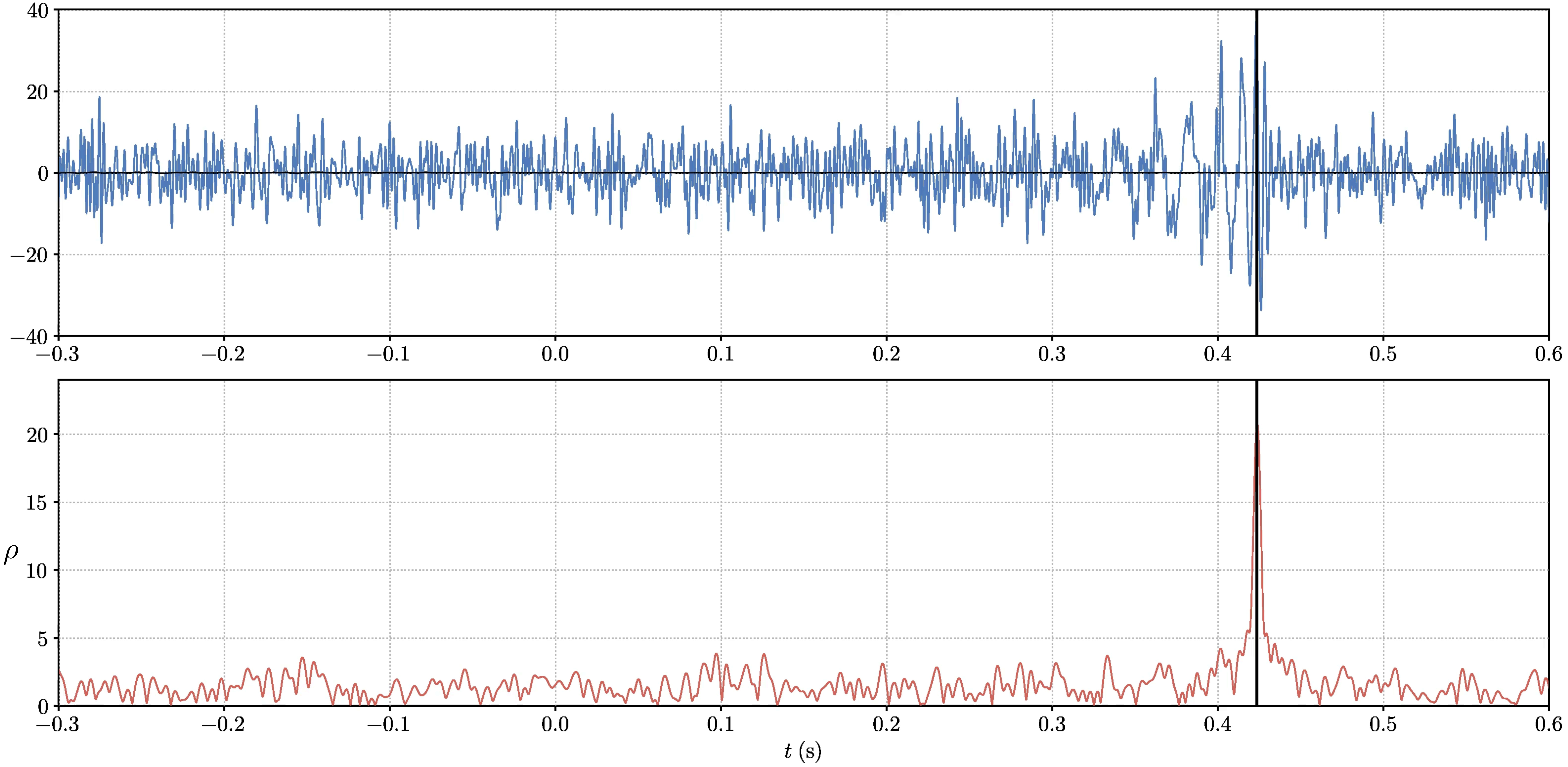
- w.b.p data: close to white noise
- GW150914 is visible in w.b.p data ! (only loud massive systems)

Matched filtering example

Try a fixed template

Optimize over phase: $\max_{\alpha} \mathrm{Re}\left(e^{i\alpha} h|s\right) = |(h|s)|$

Optimize over time: $\int df\, e^{2i\pi f \Delta t} \tilde{h}\tilde{s}^* / S_n = \mathrm{IFFT}(\tilde{h}\tilde{s}^*/S_n)$



See also GWOSC tutorials [https://gwosc.org]

Outline

Part I

- GW signals: the basics
- Noise as a stochastic process
- Introducing matched filtering
- **Towards real CBC searches**
- Other signals: continuous waves, stochastic backgrounds

Template banks

Match with nearest template:

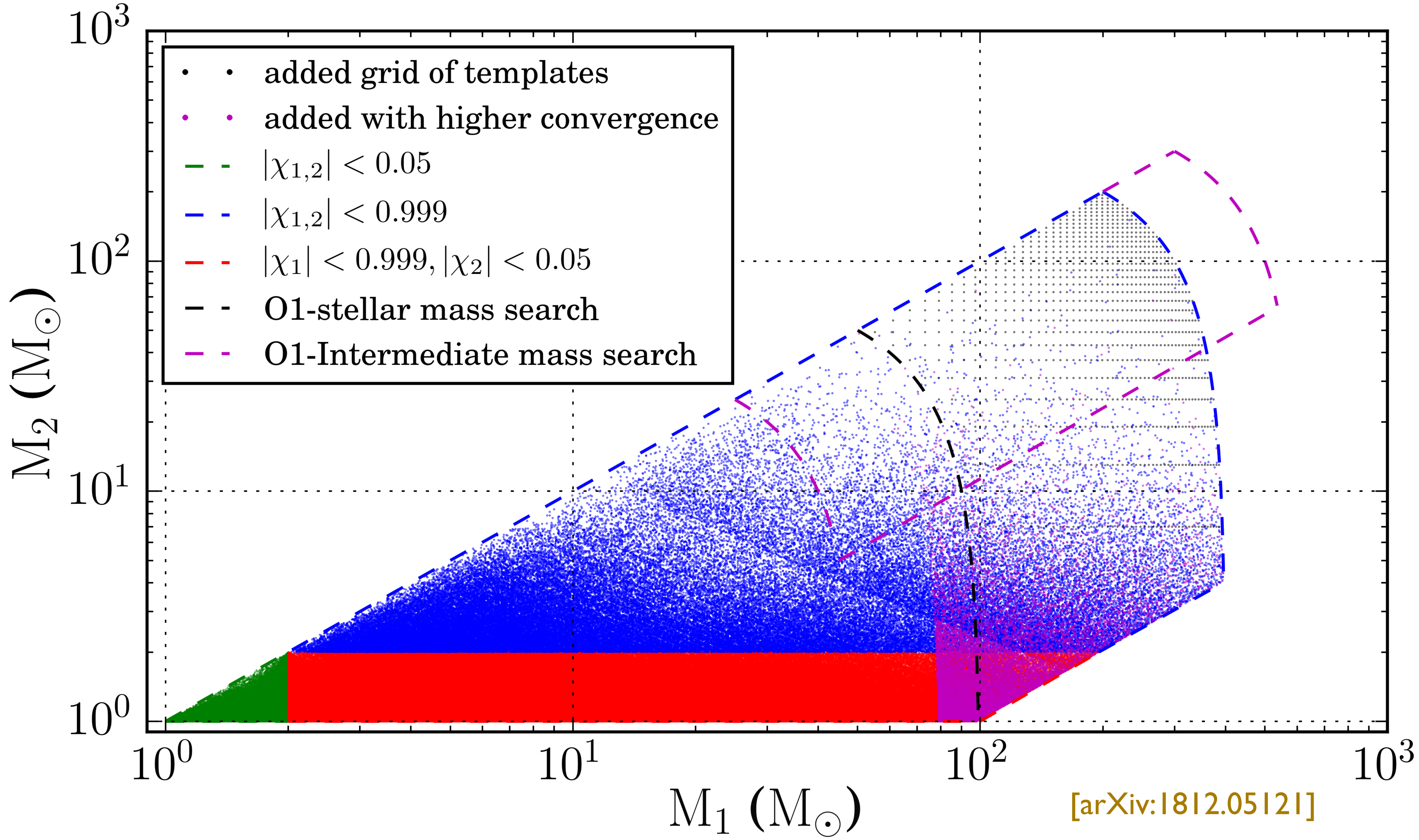
$$\max_{\Delta t,\phi} \frac{(h_t|h)}{\sqrt{(h_t|h_t)}\sqrt{(h|h)}}$$

- Effectualness criterion: match > 0.97

- Methods to build a template bank: geometric (metric based on match), stochastic, hybrid

- Trade-off between effectualness (template bank size) and FAR

- Simplified physics (no precession)

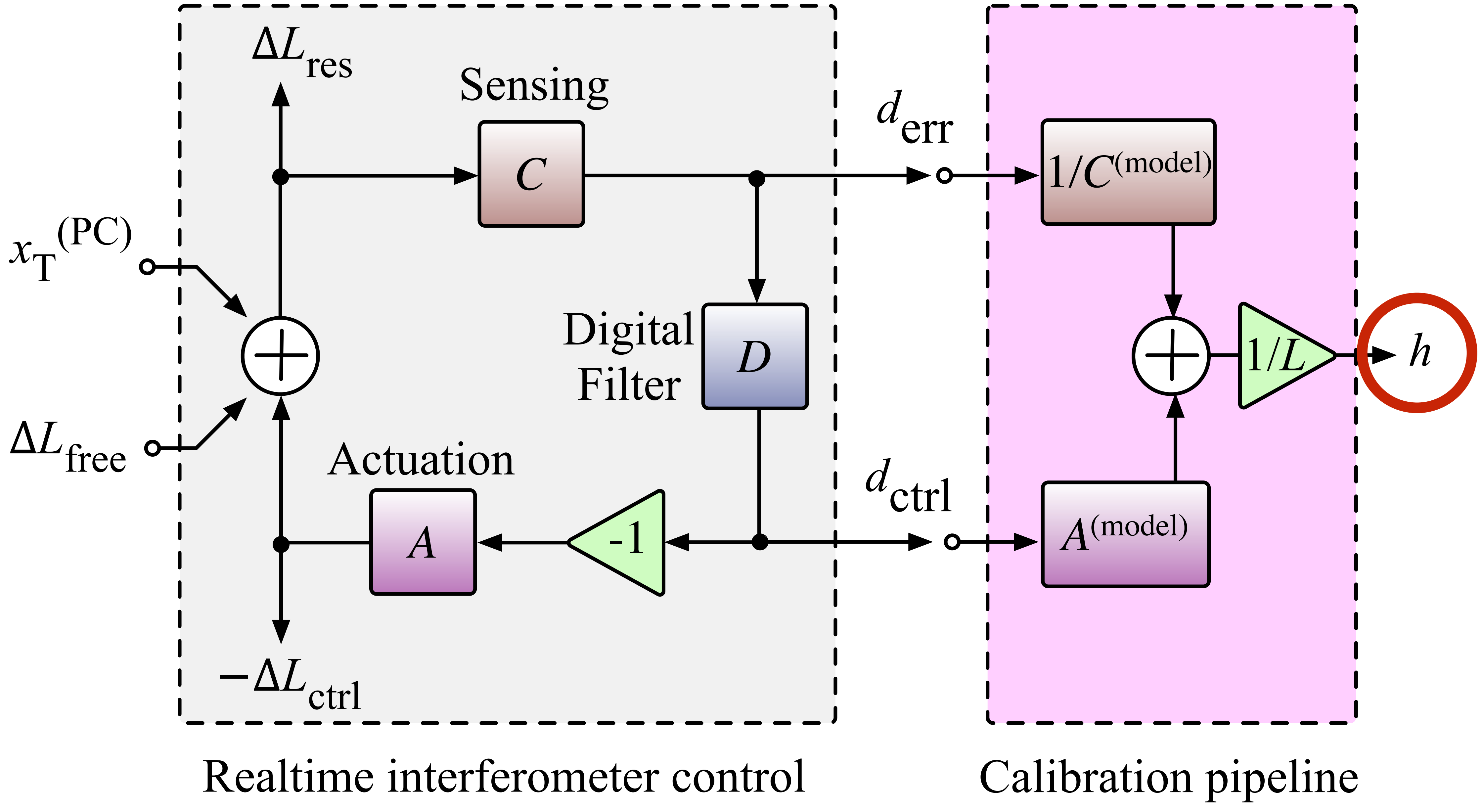


[arXiv:1812.05121]

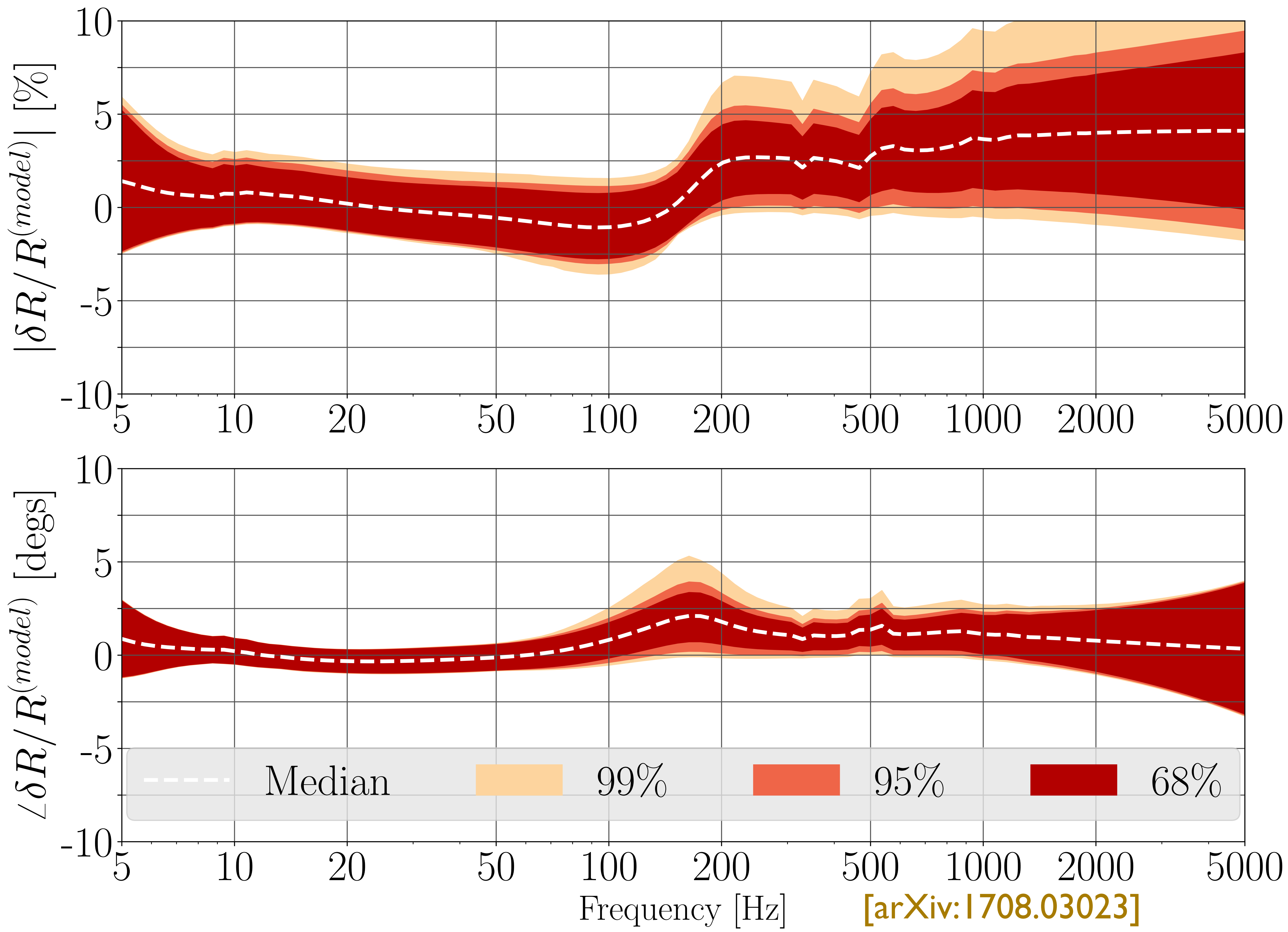
Templates are more orthogonal at low masses, with many wave cycles

Real data: calibration

[arXiv:1708.03023]



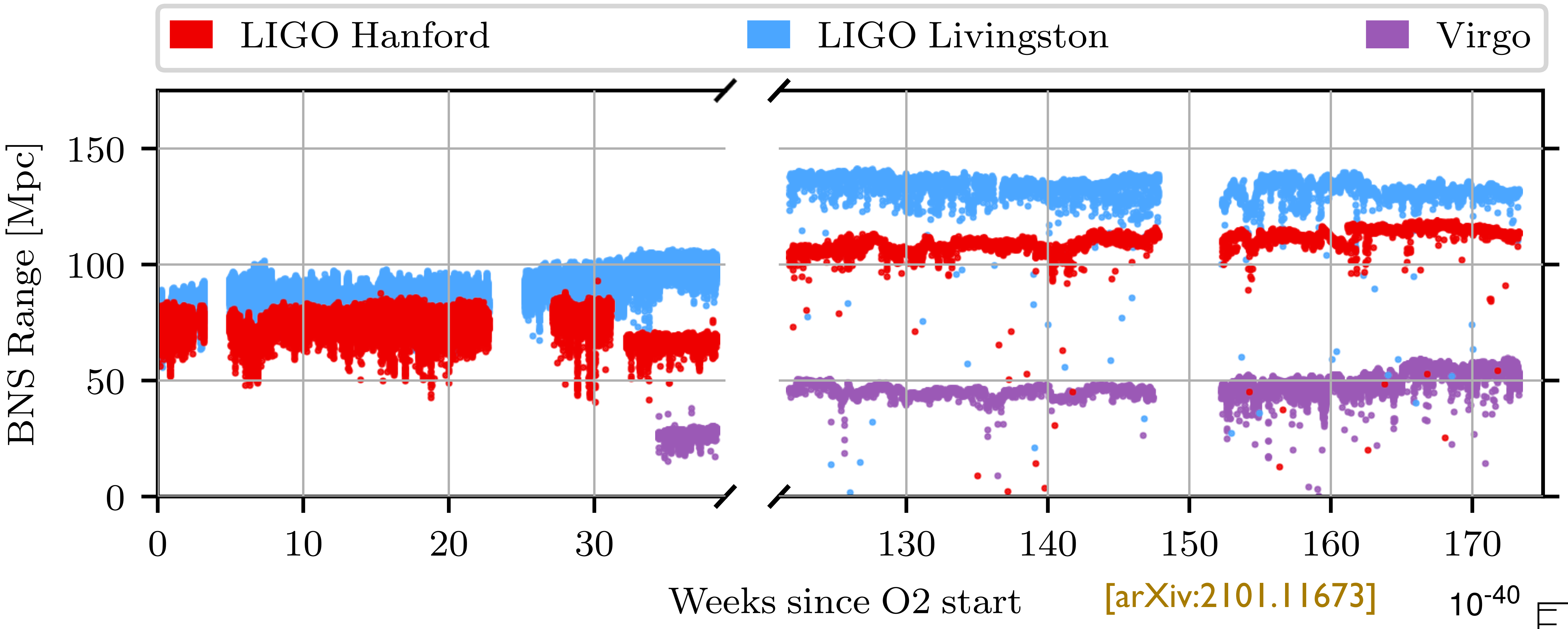
- Calibration: the output strain is the results of a complex control loop



[arXiv:1708.03023]

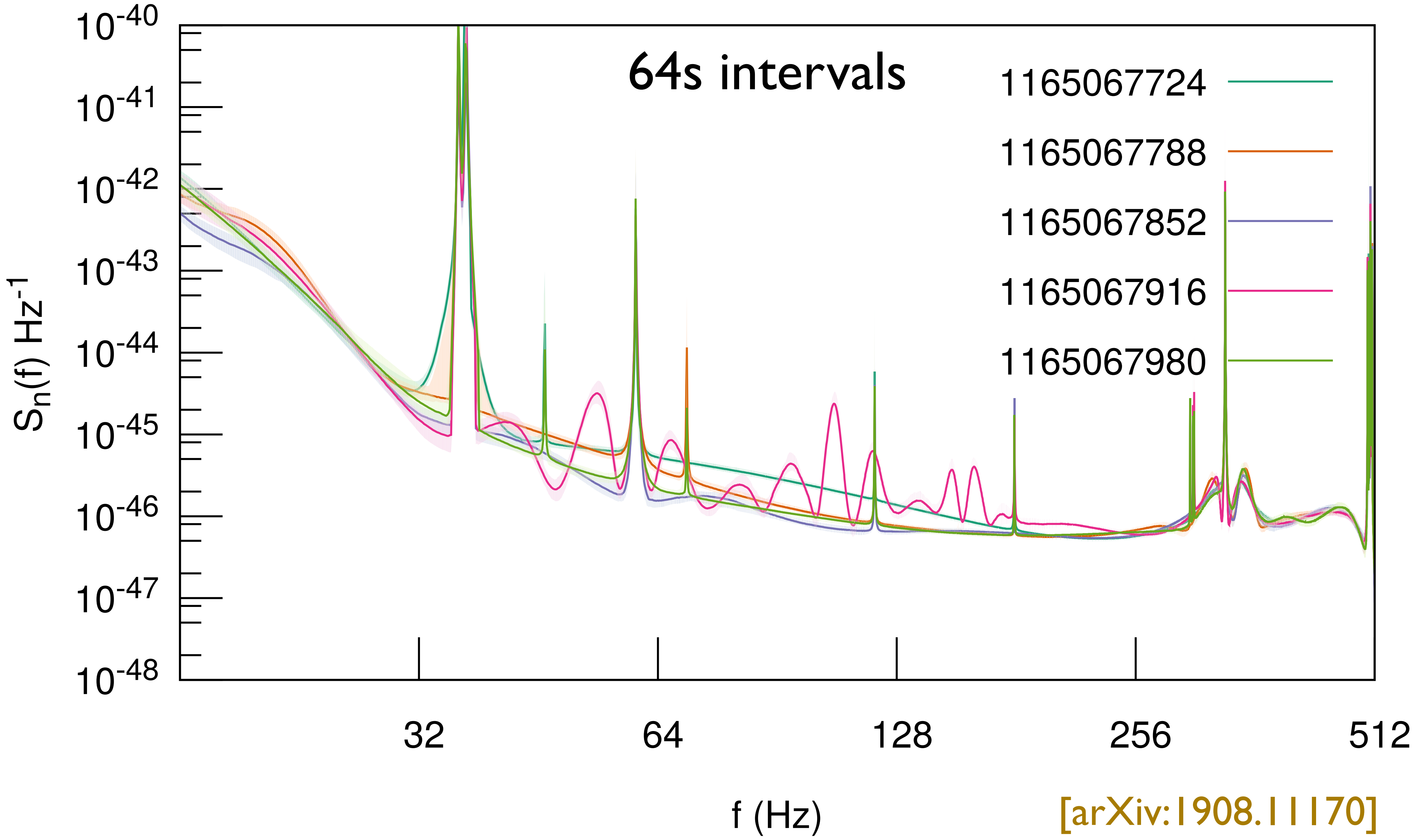
- Calibration is in part stochastic: amplitude and phase splines, with value at nodes randomly distributed in envelope

Real data and artefacts: glitches, non-stationarity



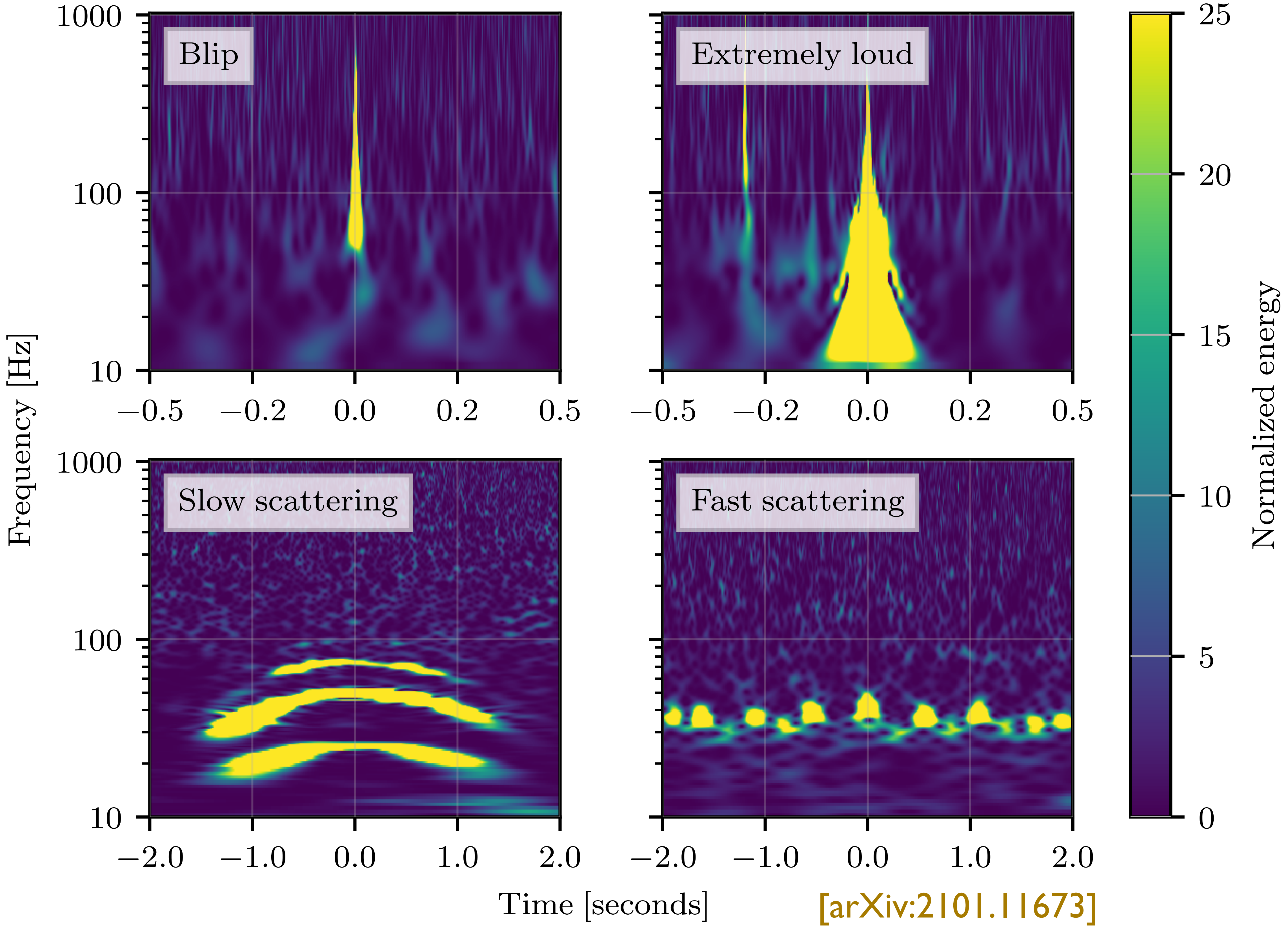
[arXiv:2101.11673]

- Detectors evolve over time, varying duty cycle
- Long-term variations of sensitivity
- PSD can show non-stationarity on short time scales

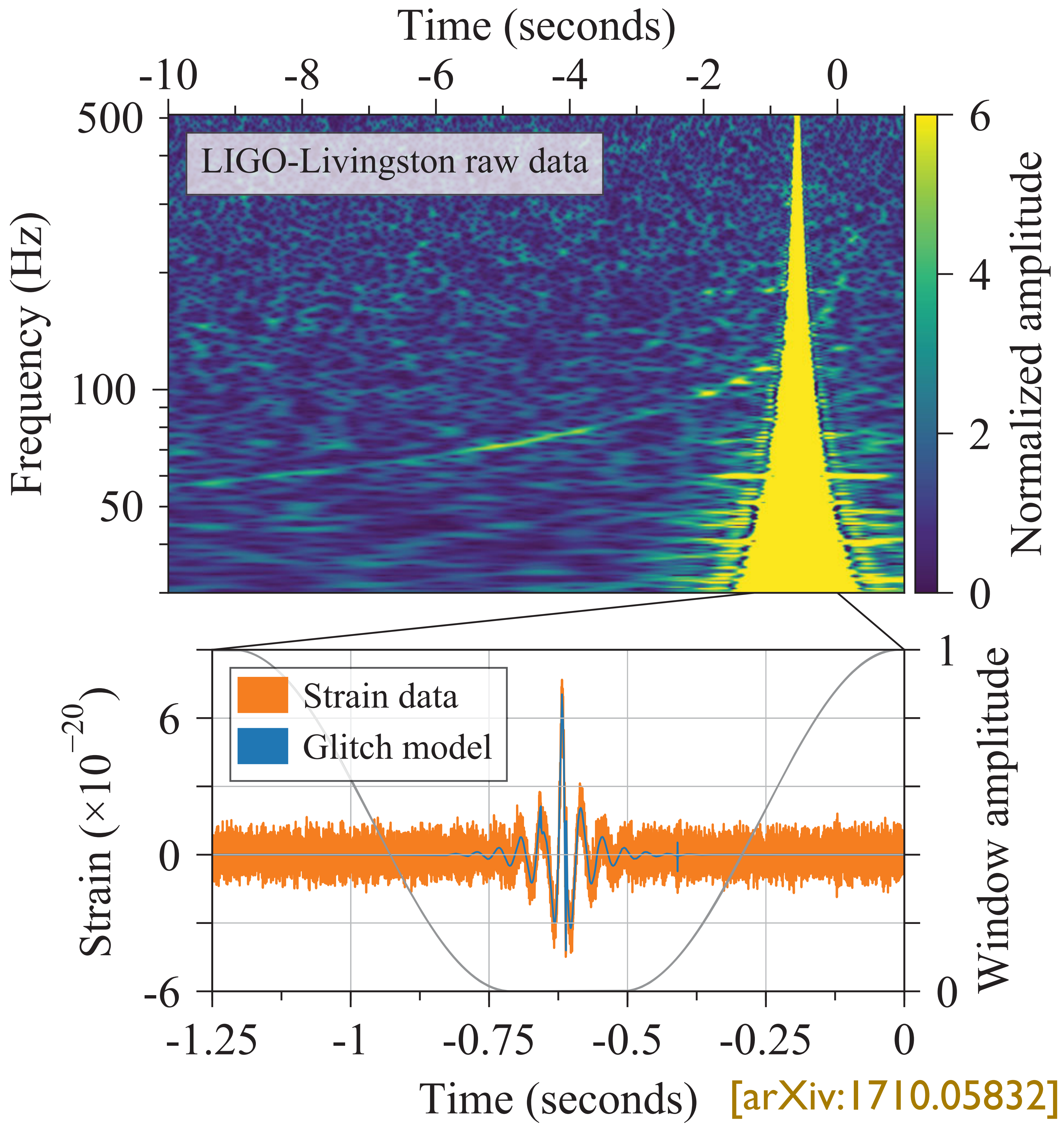


[arXiv:1908.11170]

Real data and artefacts: glitches, non-stationarity



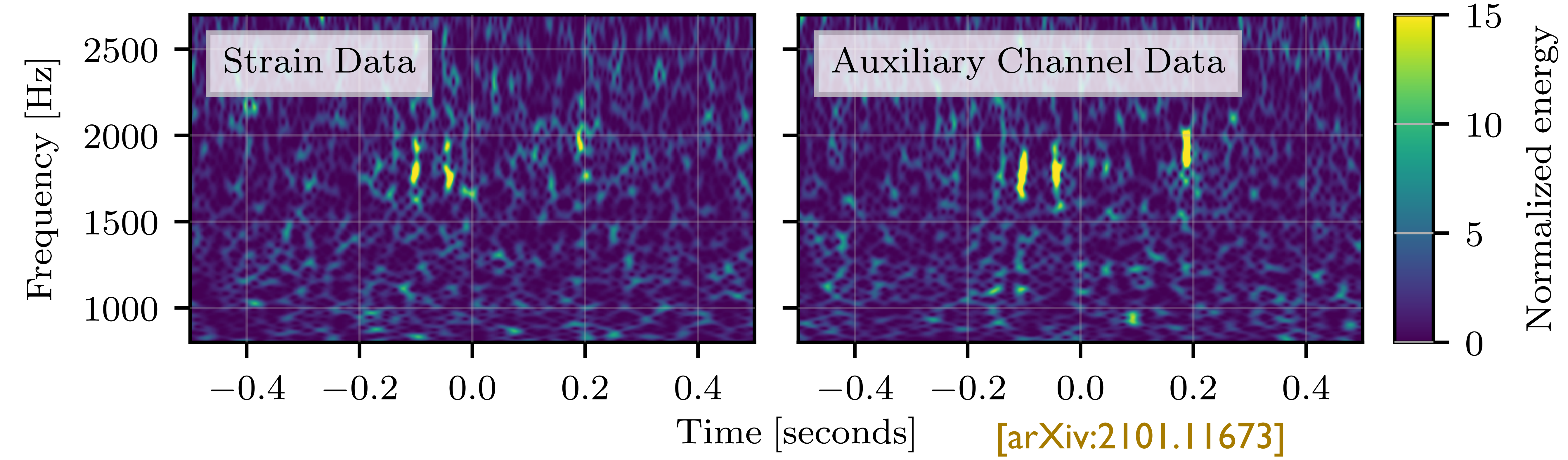
[arXiv:2101.11673]



[arXiv:1710.05832]

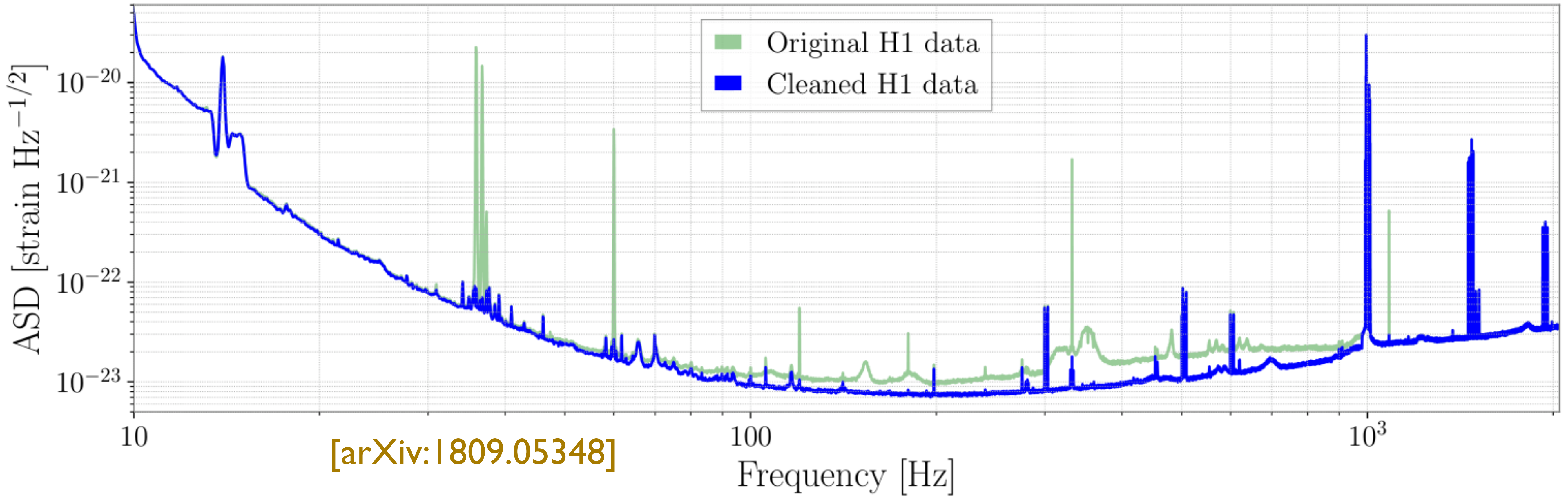
- Glitches: strong non-stationary, non-Gaussian events
- SNR alone would be dominated by glitches
- Need more robust significance metric

Real data: data quality, data cleaning



[arXiv:2101.11673]

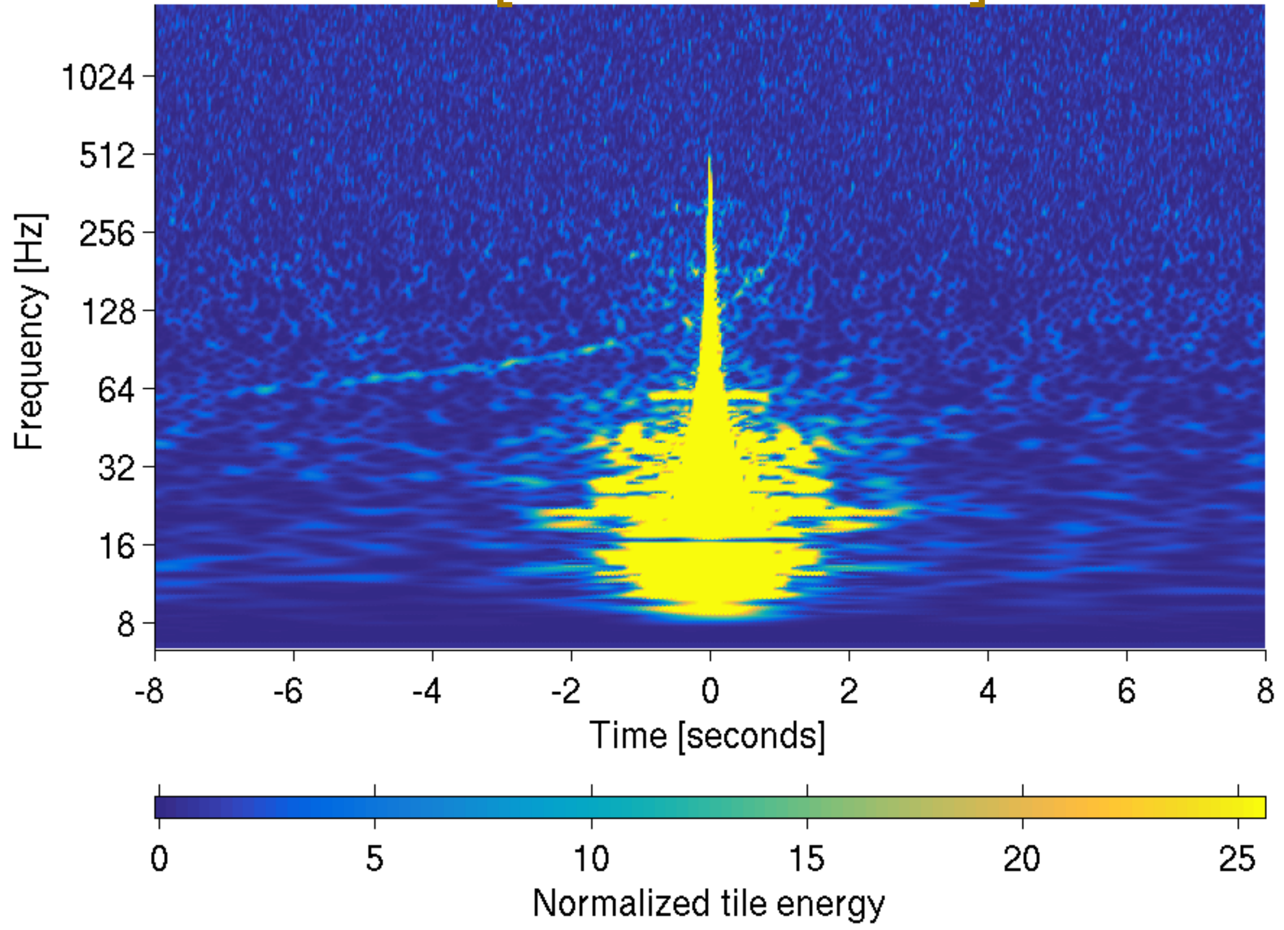
- Data quality: exploit auxiliary channels, issue vetoes

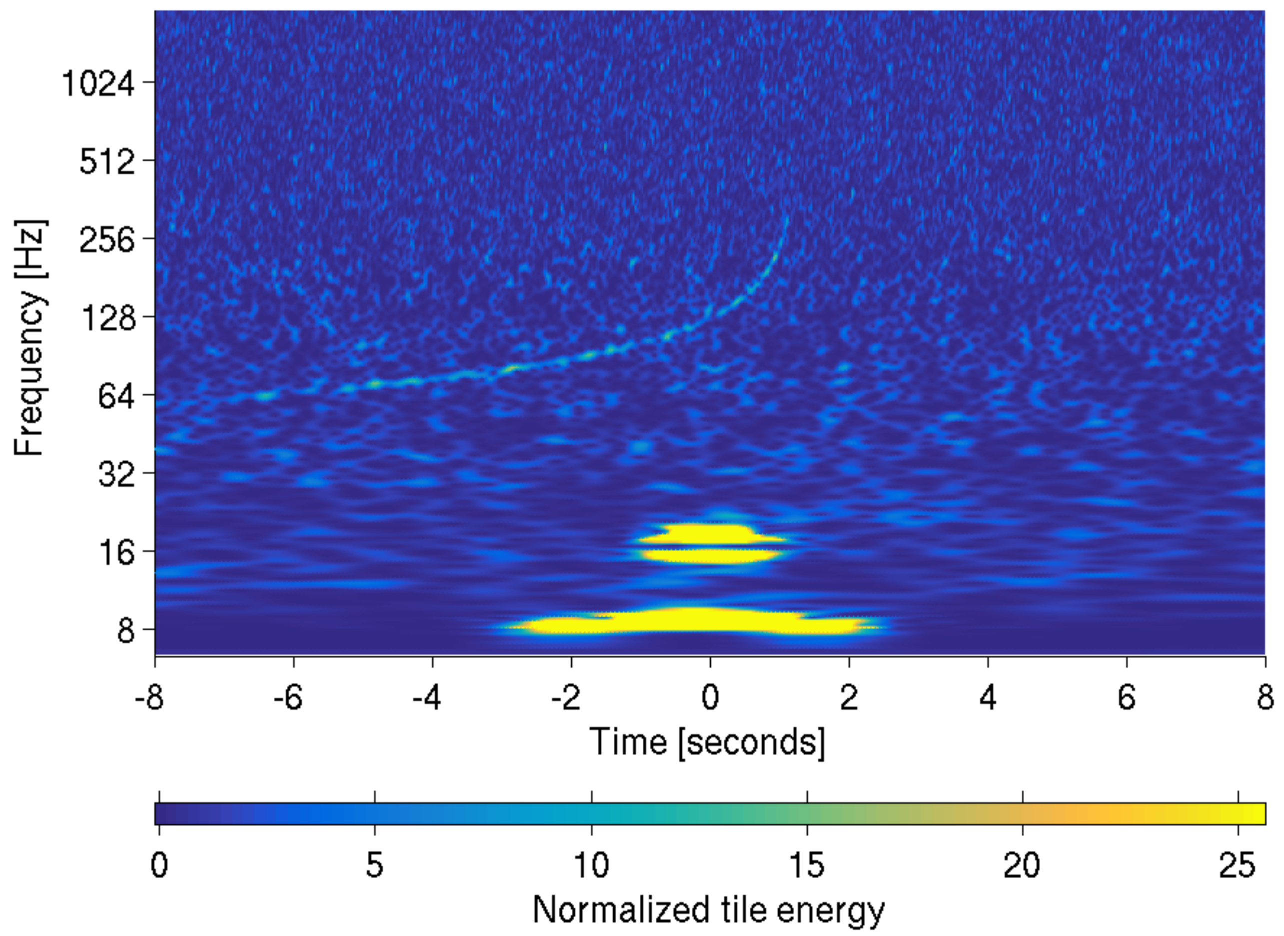


[arXiv:1809.05348]

- Data cleaning (removal of noise lines)

[arXiv:1808.03619]





- Glitch gating or removal (BayesWave)

Signal consistency and ranking statistic

PyCBC

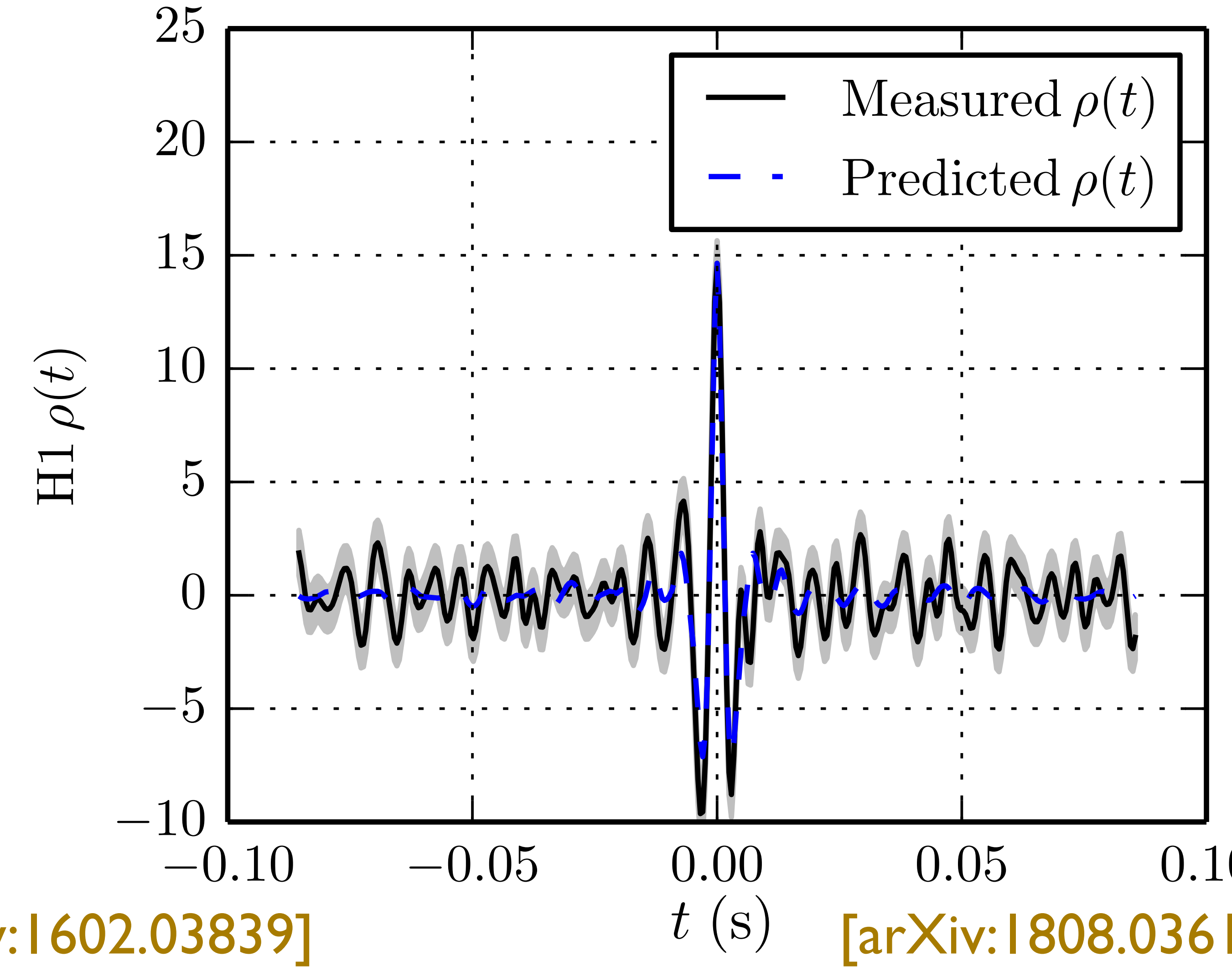
Penalization of large residuals:
computed on n freq. bands, χ^2 with $\nu = 2n - 2$ d.o.f.

$$\hat{\rho} = \rho \times \begin{cases} 1 & \chi^2 \leq \nu \\ \left[\frac{1}{2} + \frac{1}{2}(\chi^2/\nu)^3\right]^{-1/6} & \chi^2 > \nu \end{cases}$$

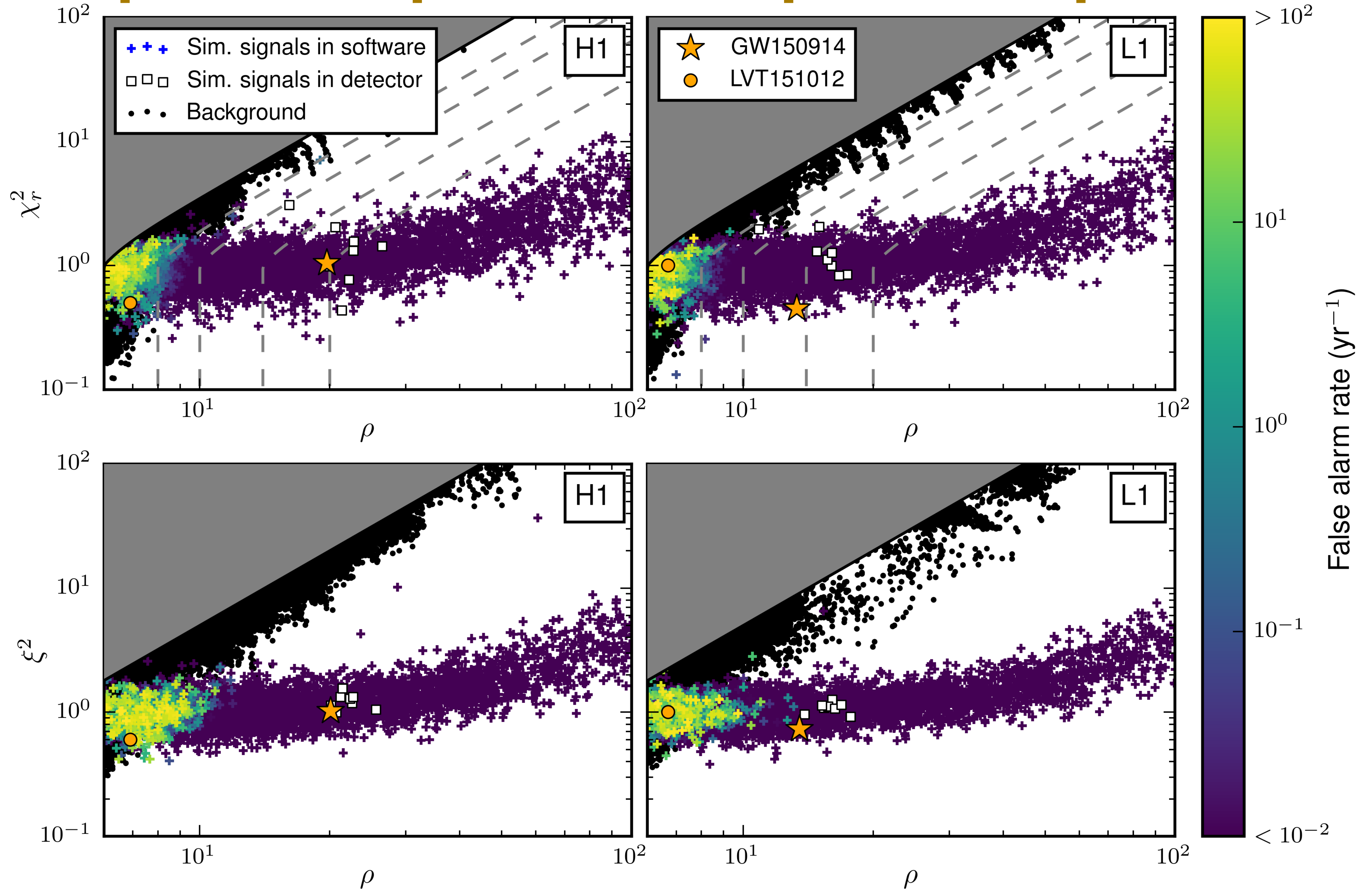
GstLAL

Consistency between time series and autocorrelation of template

- Idea: use a ranking statistic for all foreground/ background events
- Tradeoff between false alarm and false dismissals
- Can use different ranking statistics !

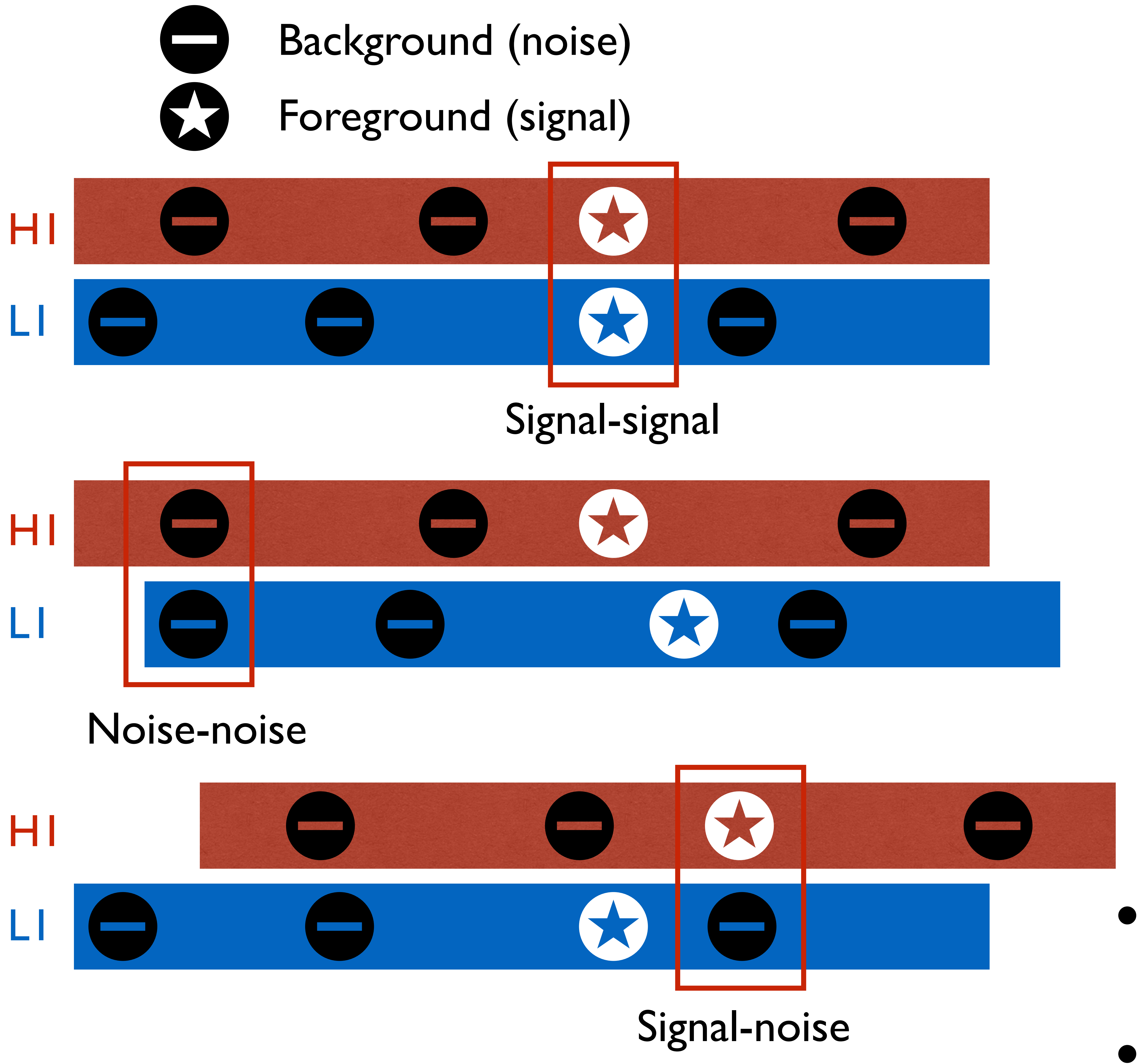


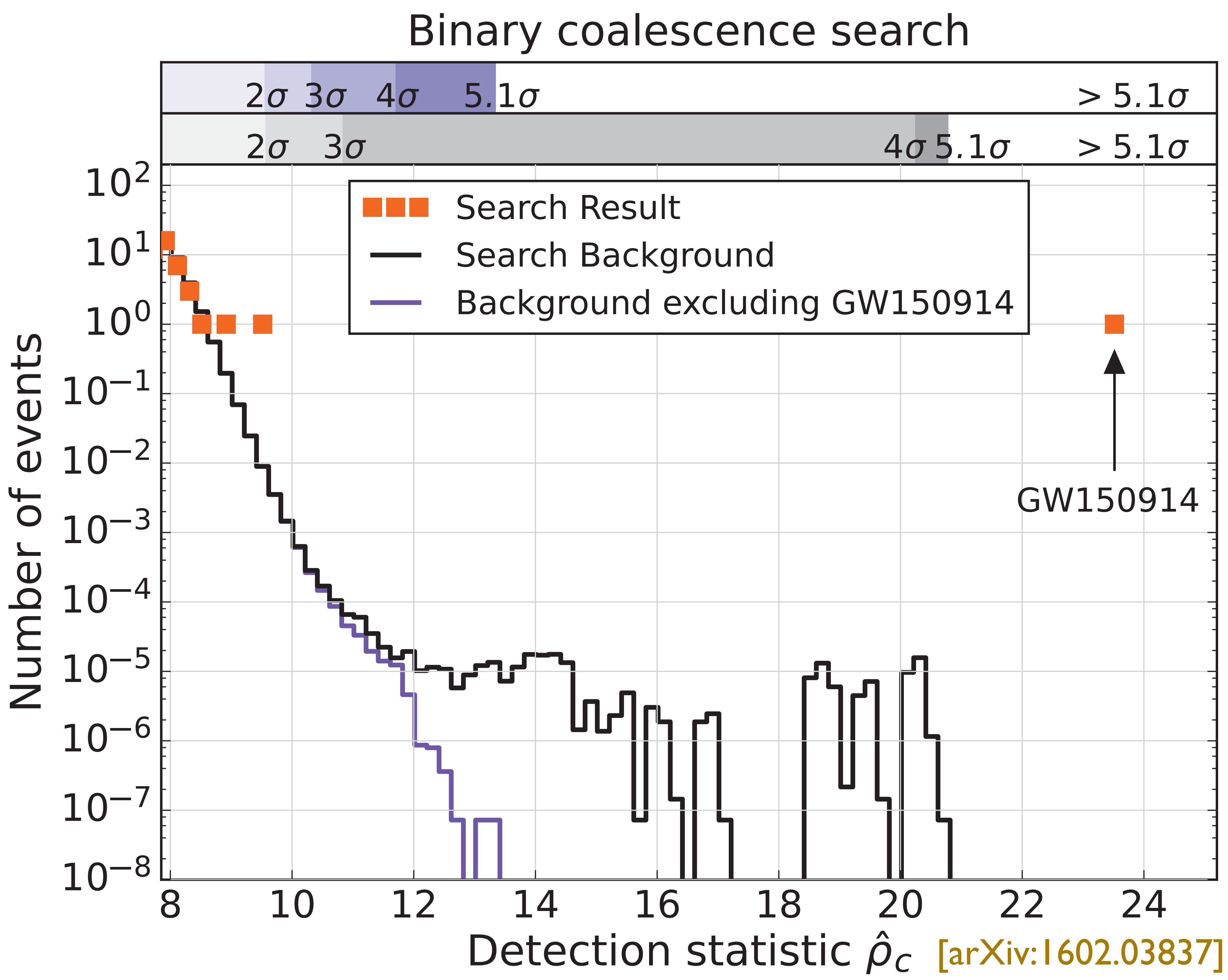
[arXiv:1602.03839]



[arXiv:1808.03619]

Significance of coincident triggers: time slides



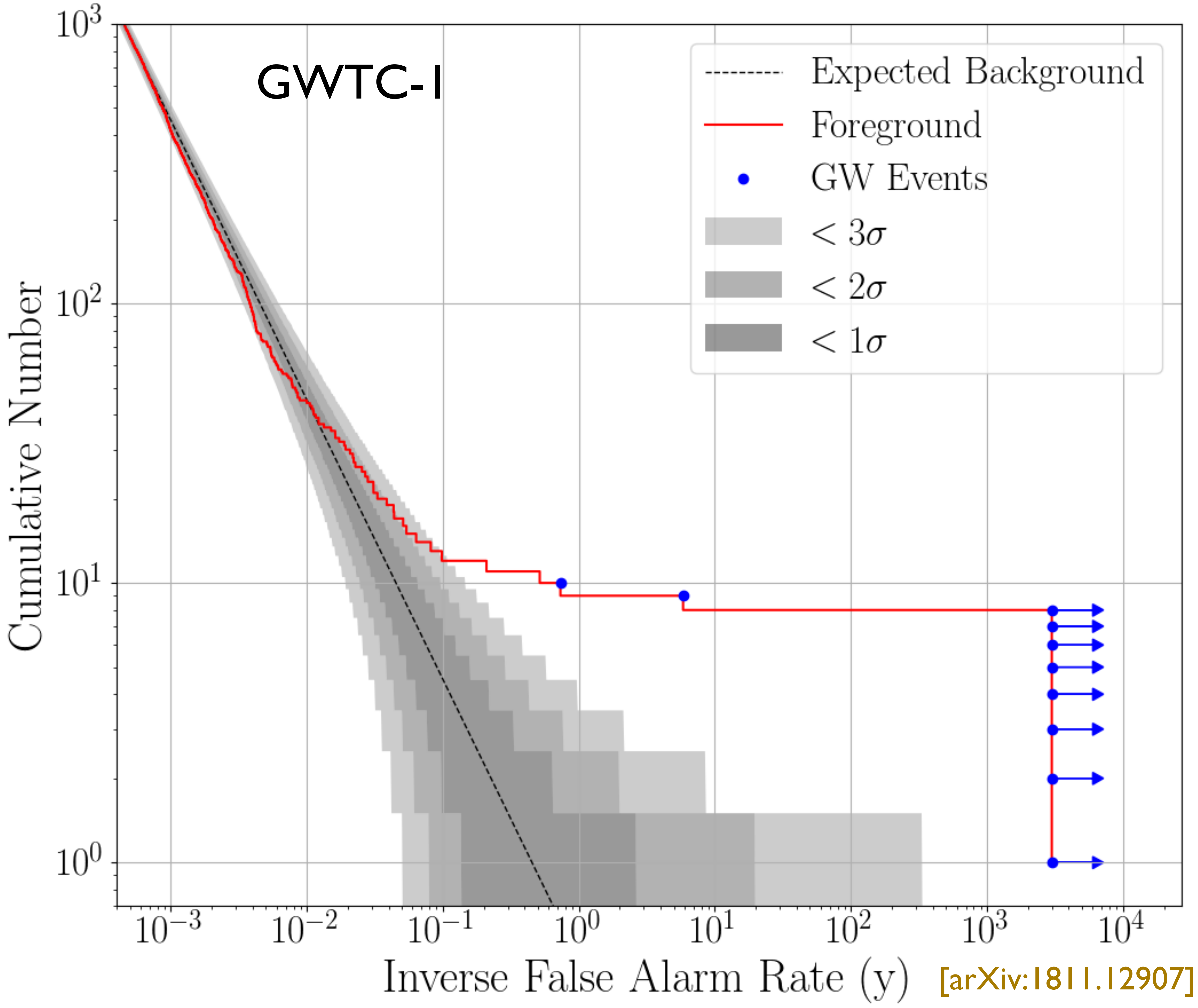


[arXiv:1602.03837]

- Generate large background of coincidences by sliding time series
- False Alarm Rate: with and without signal

Inverse False Alarm Rate (IFAR)

- Rank all triggers with ranking statistic of choice
- From rank of trigger False Alarm Rate (over time of extended data)
- Cumulative distribution of IFAR: N=T/IFAR
- Consistency, does not say how sensitive the search is



[arXiv:1811.12907]

Sensitivity and p_astro

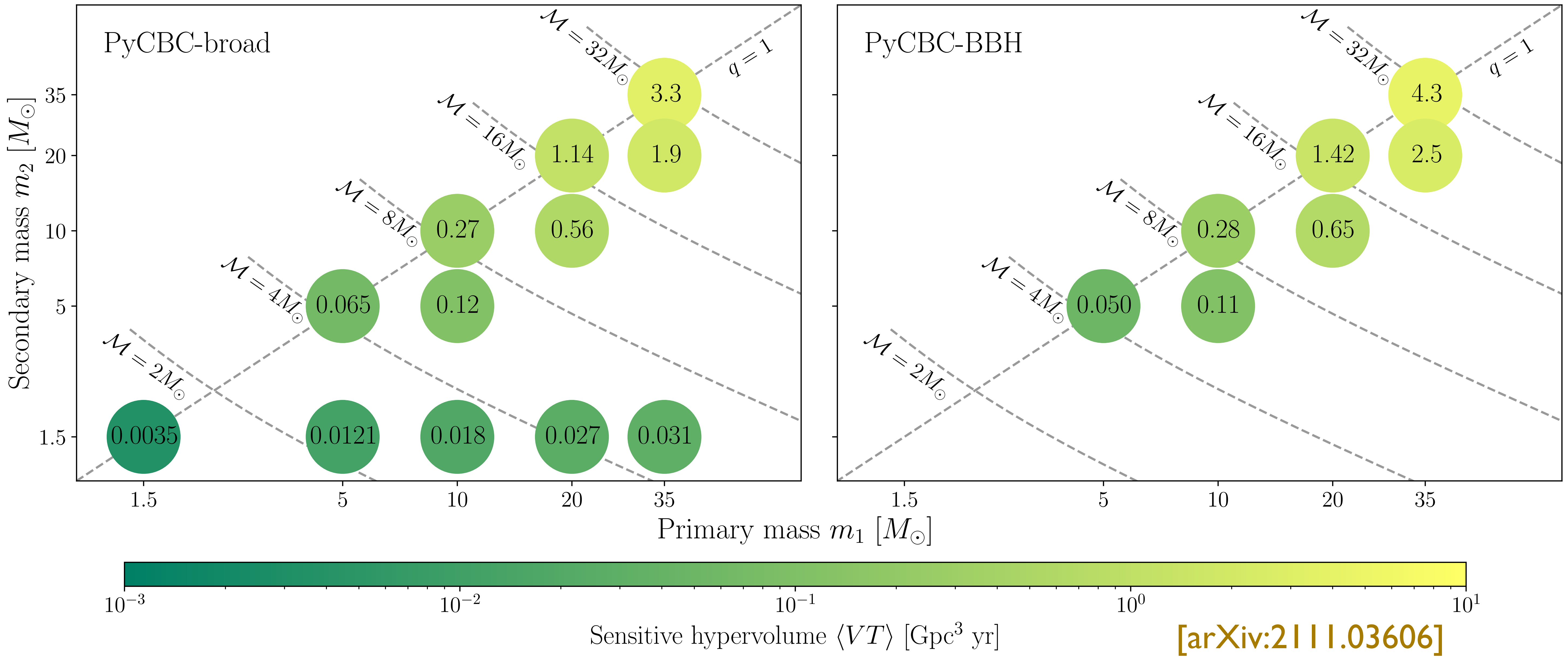
- Injection campaigns (simulated signals) to estimate sensitivity

$$N = \langle VT \rangle R$$

expected count

sensitive volume

astrophysical rate



[arXiv:2111.03606]

- Probability of astrophysical origin: p_astro

$$\frac{\text{foreground}}{\text{foreground} + \text{background}}$$

+ marginalisation over fg-bg counts

p_astro

likelihood for counts

Λ_1, Λ_0 counts for fg, bg

$$P_1(x \mid \vec{x}) = \int_0^\infty p(\Lambda_0, \Lambda_1 \mid \vec{x}) \frac{\Lambda_1 f(x)}{\Lambda_0 b(x) + \Lambda_1 f(x)} d\Lambda_0 d\Lambda_1$$

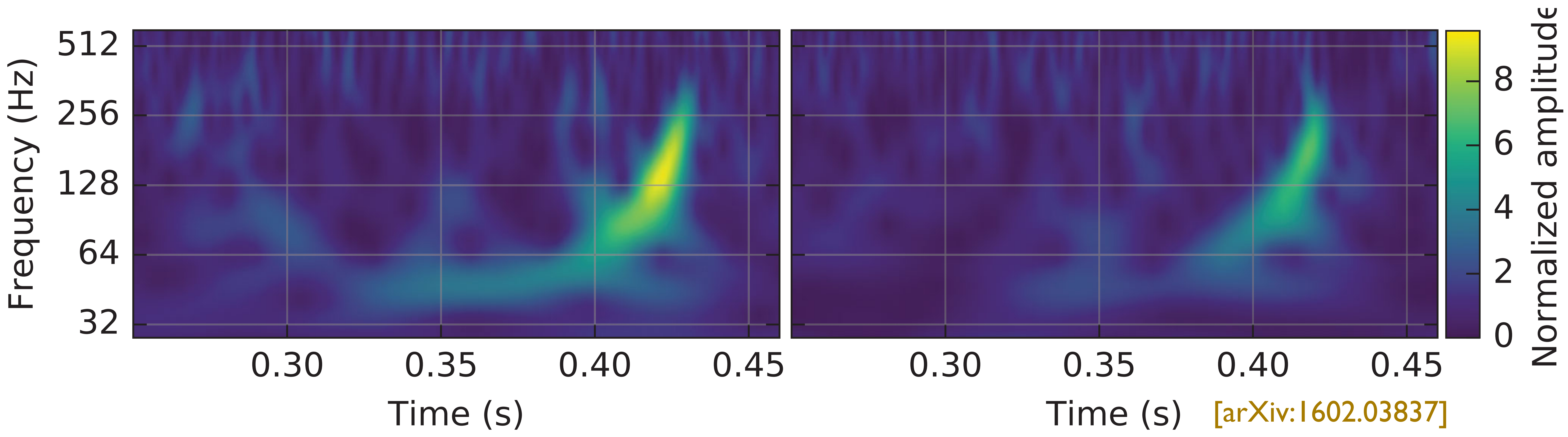
$x, \vec{x}$ ranking stat. event, data

$f(x), b(x)$ ranking stat. distrib, fg and bg

[arXiv:1302.5341]

[arXiv:1903.06881]

Unmodeled search for bursts



- Looking for generic transients
- Crucial for SN, robustness of CBC

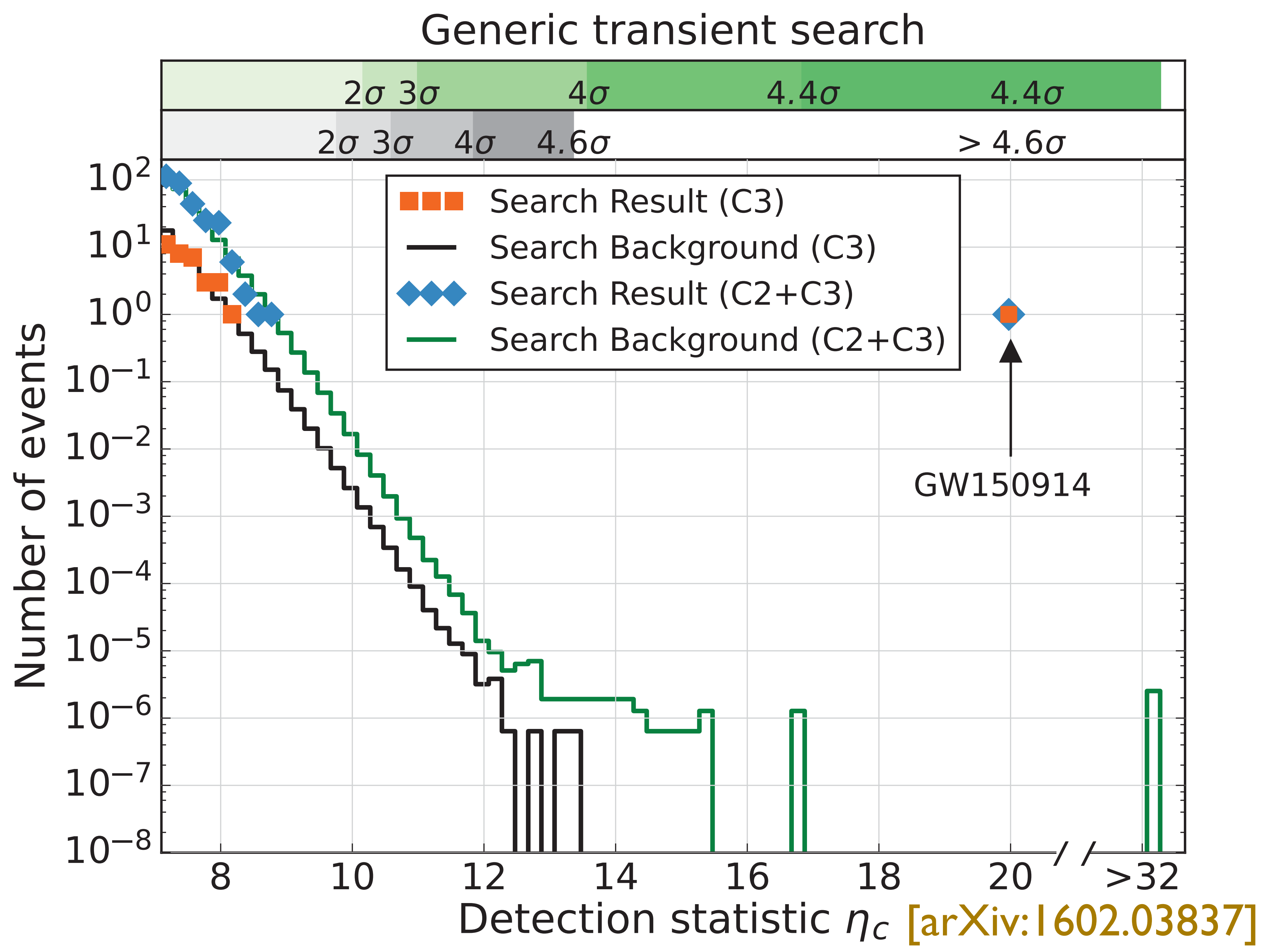
- Time-frequency domain: wavelets (cWB, BayesWave)
- Exploit direction-dependent detector response: signal reconstruction
- Background estimation challenging, introduce chirp morphology, vetoes

Detection statistic:

$$\eta_c = \sqrt{\frac{2E_c}{(1 + E_n/E_c)}}$$

E_c coherent signal power

E_n residual noise power



Overview of search pipelines

- cWB — Generic search
- GstLAL
- MBTA
- PyCBC
- SPIIR

Dedicated searches

Run offline

Run online (low-latency)

Online analysis:

- minimize latency
- limited data quality/calibration information
- send alerts based on FAR

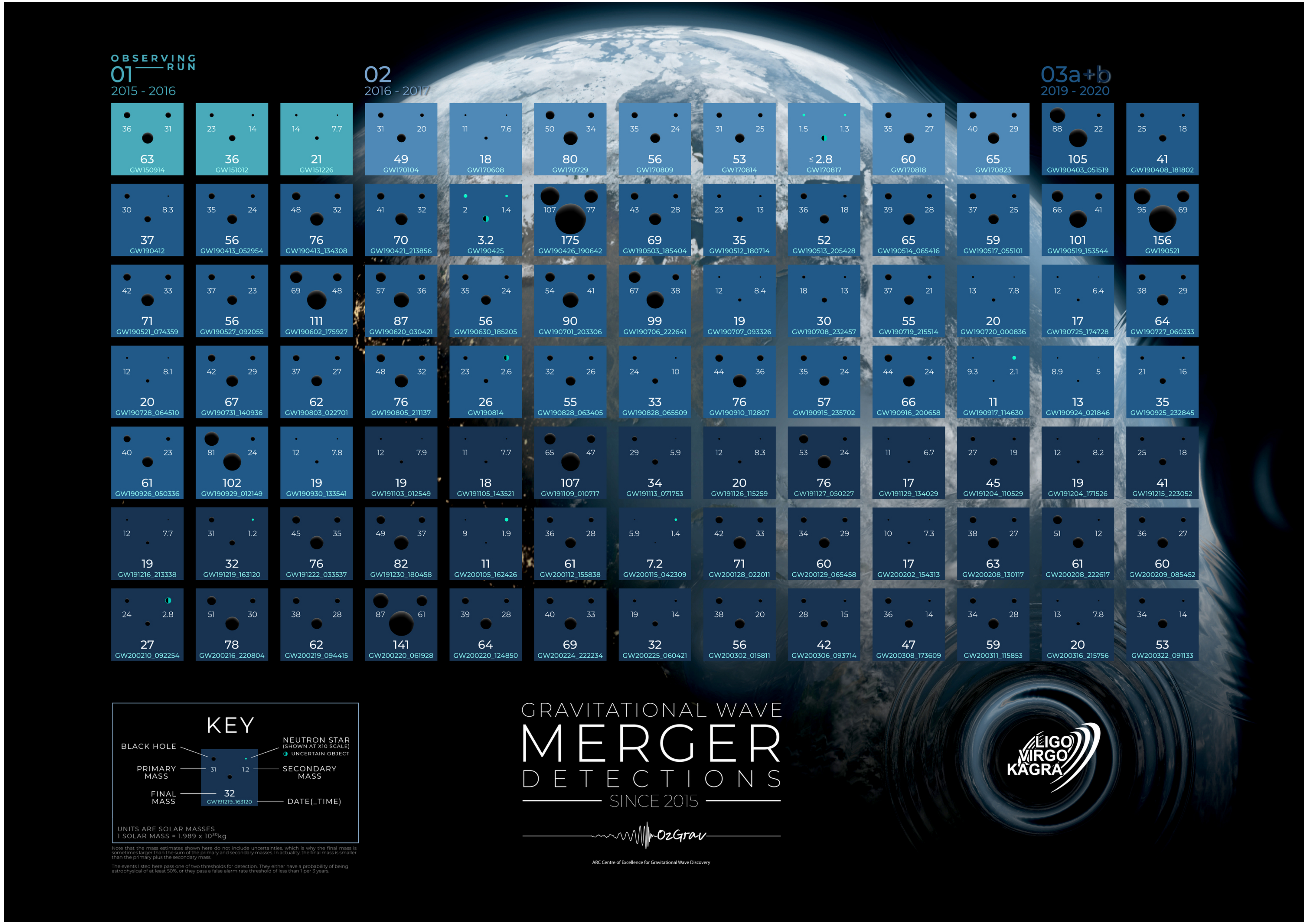
Offline analysis:

- run on ~1week chunks
- final data quality/calibration information
- use p_astro > 0.5 for catalogs

Outside groups:

- PyCBC
- Princeton

CBC detections



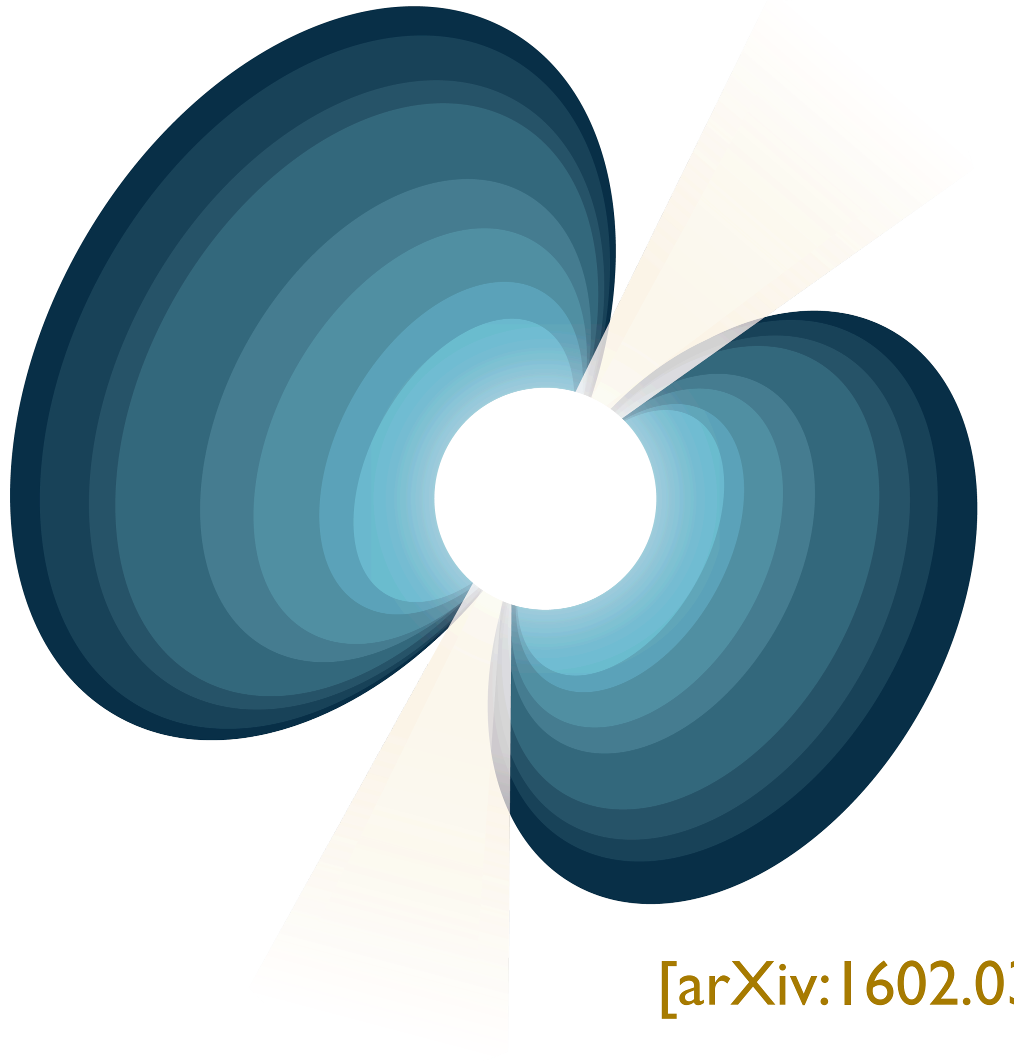
Outline

Part I

- GW signals: the basics
- Noise as a stochastic process
- Introducing matched filtering
- Towards real CBC searches
- **Other signals: continuous waves, stochastic backgrounds**

Continuous waves

Target: rotation of asymmetric neutron star (pulsar)



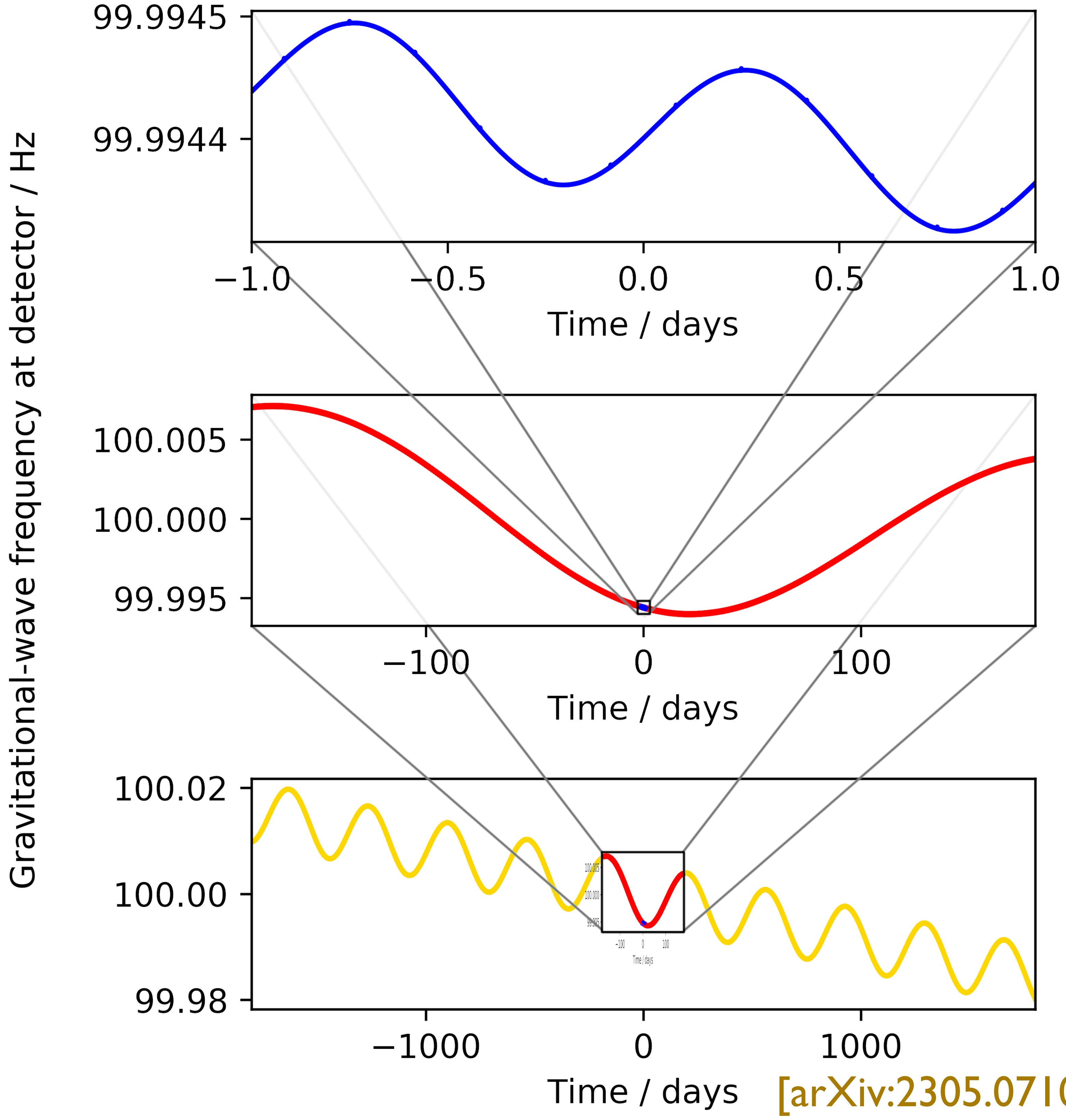
[arXiv:1602.03837]

- Targeted/directed search: known pulsars, galactic center
- All-sky coherent search untractable !
- Semi-coherent searches required

Long-lived quasi-monochromatic signals: modulated response

$$F_+,\, F_\times \to F_+(t),\, F_\times(t)$$

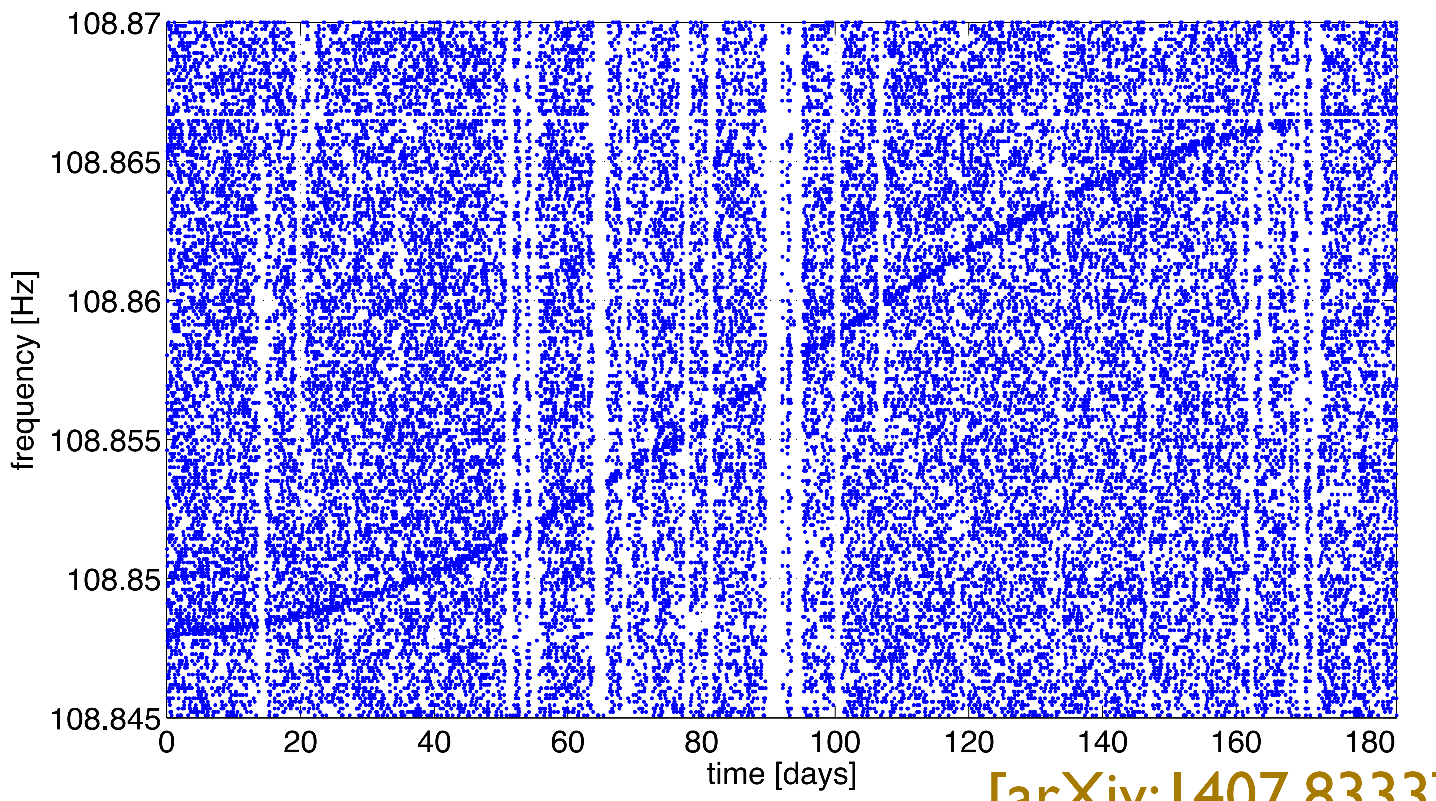
$$\tau(t) = t + \frac{\vec{r}(t)\cdot\vec{n}}{c} + \Delta_{E\odot} - \Delta_{S\odot}$$



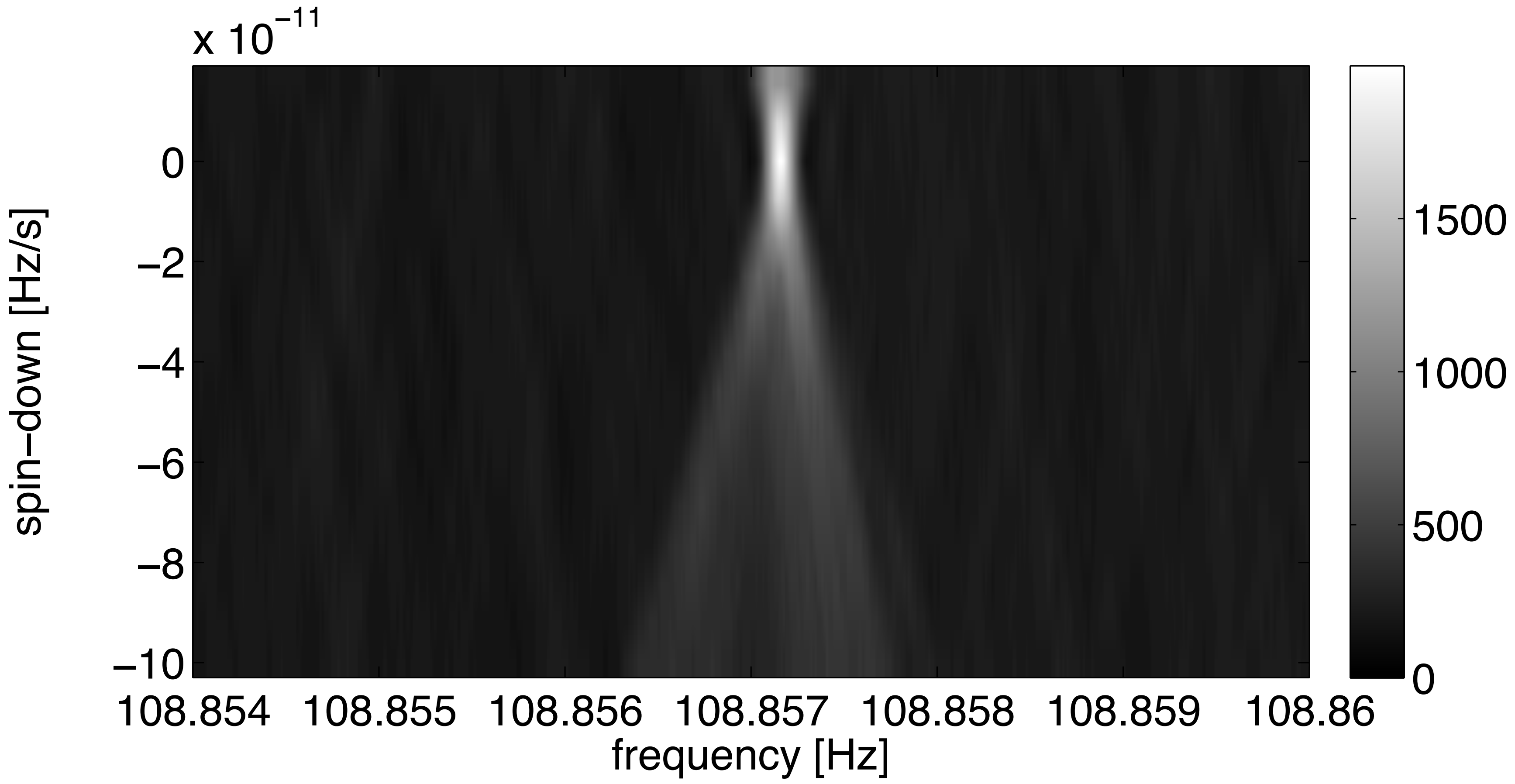
[arXiv:2305.07106]

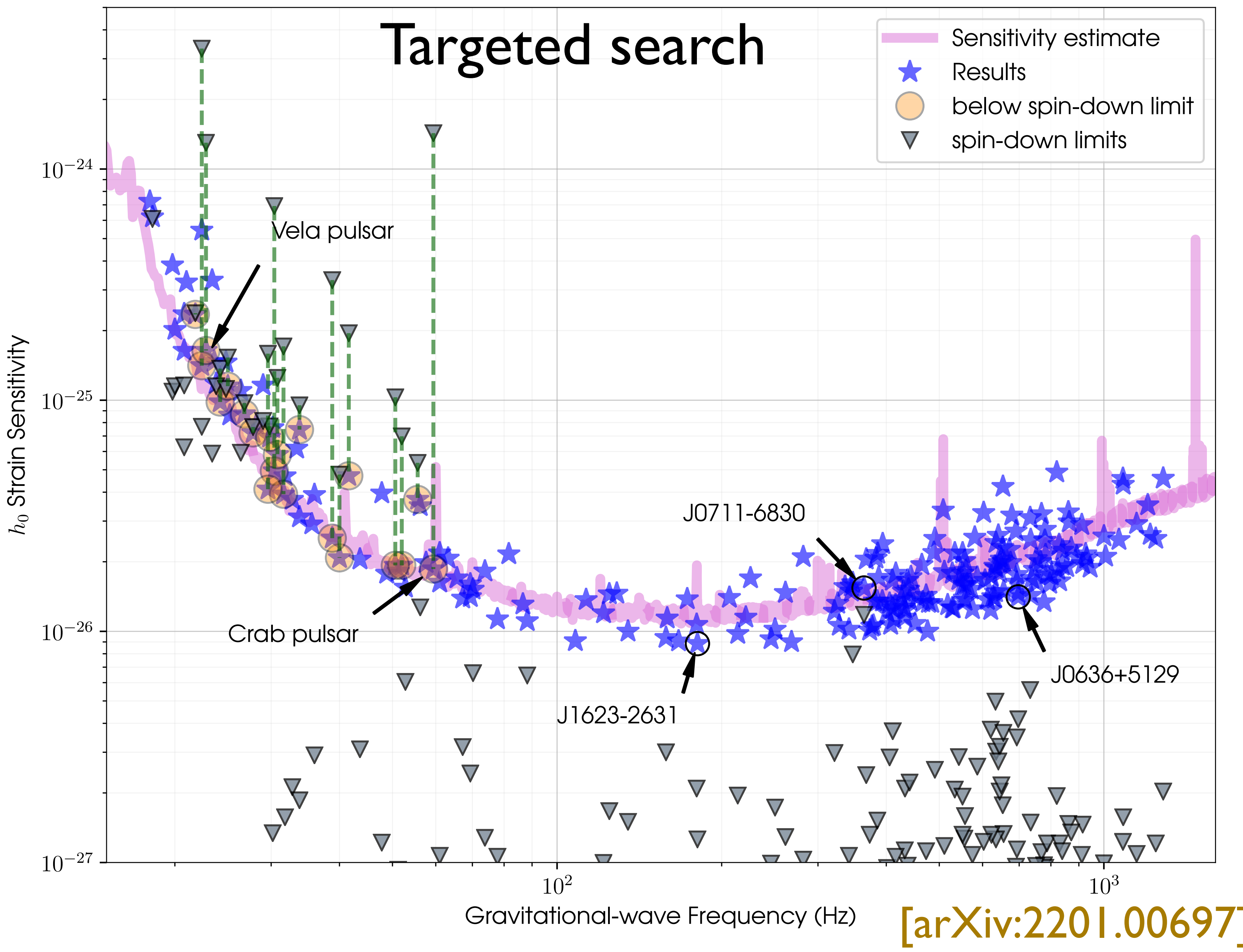
Continuous waves

Hough transform: threshold pixels in $t-f$ space, each is a line in $f,\dot{f}$

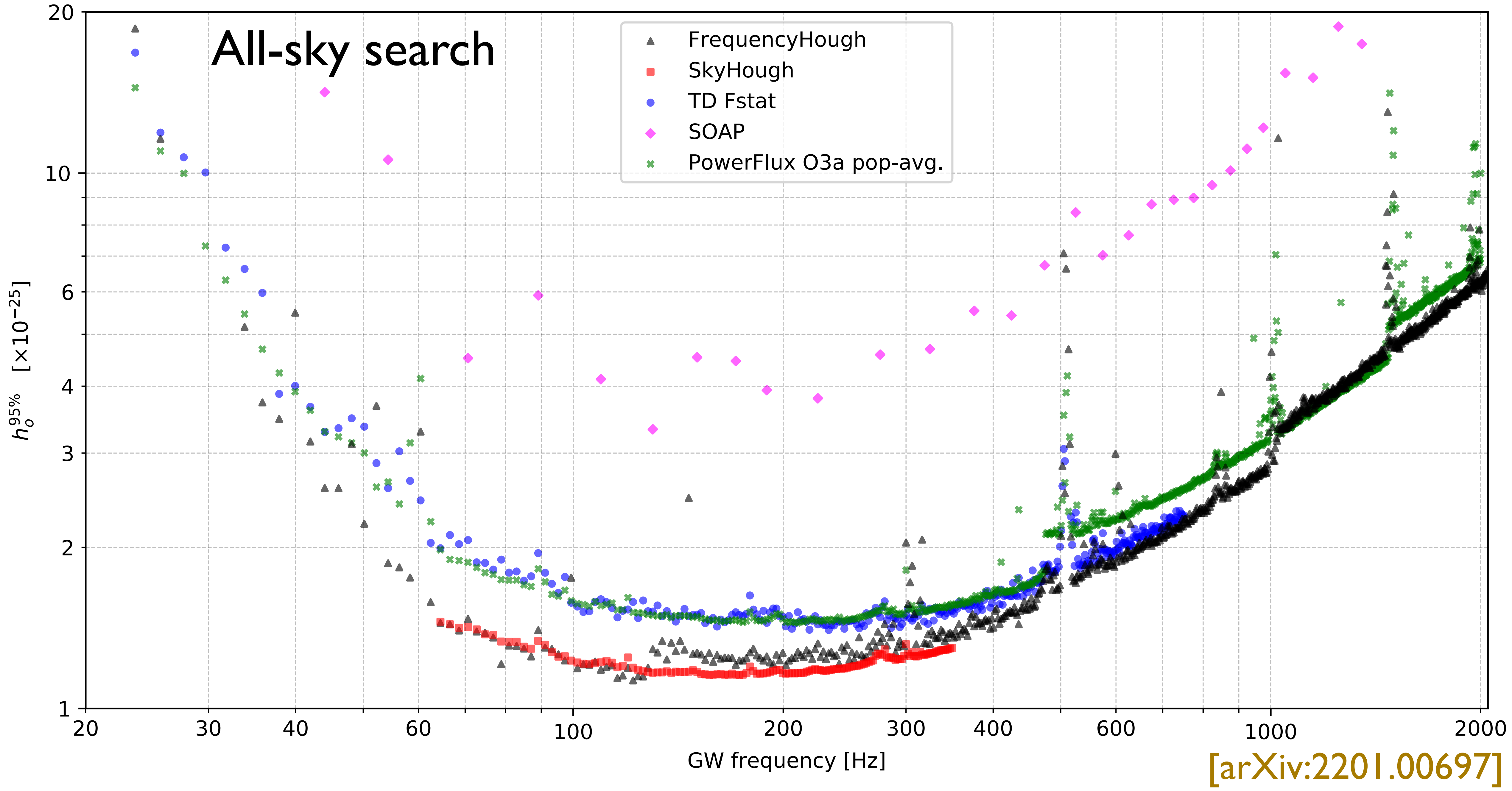


[arXiv:1407.8333]





[arXiv:2201.00697]



[arXiv:2201.00697]

Stochastic backgrounds

Superposition of signal(s) from all directions:

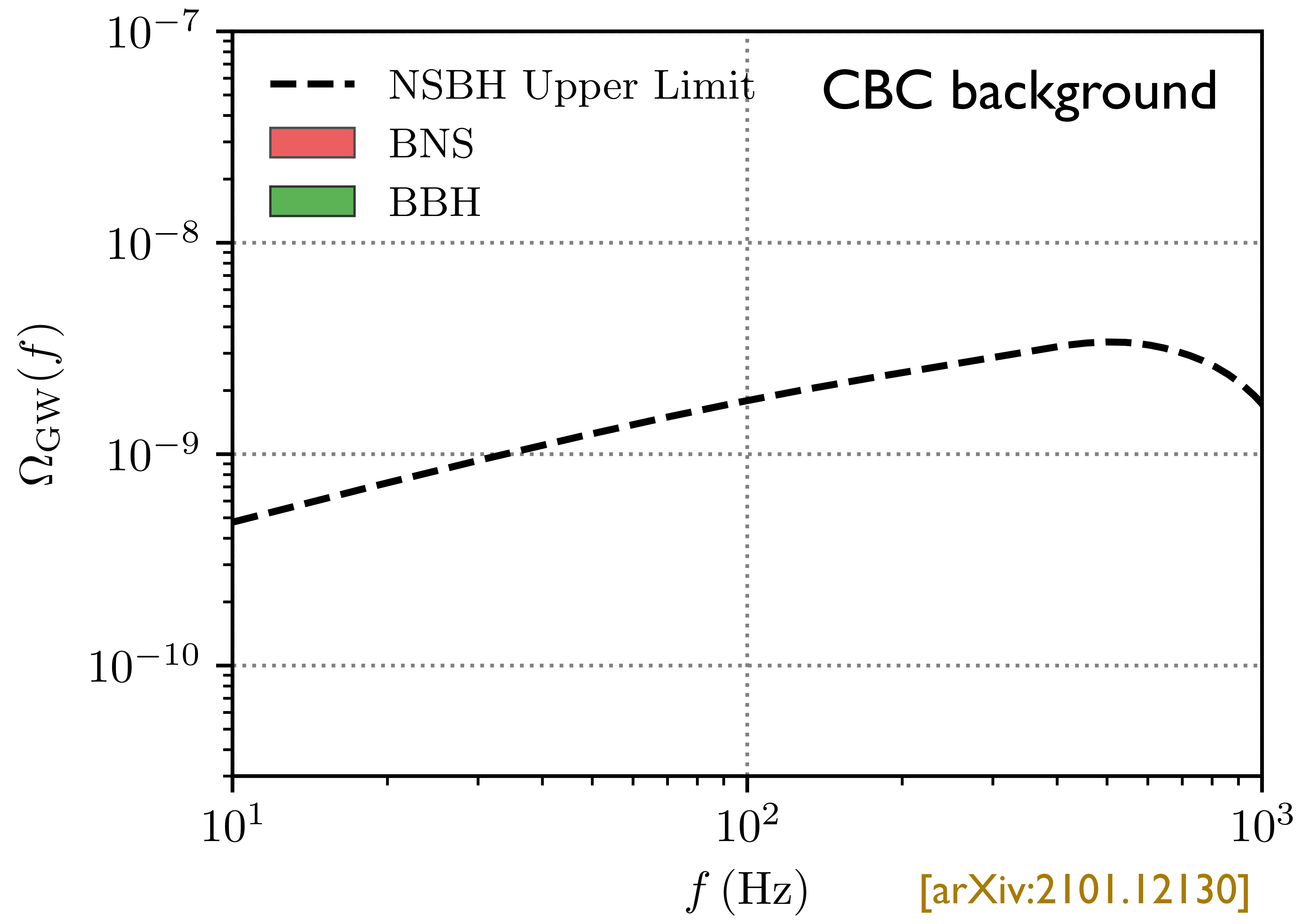
$$\mathbf{h}(t,\mathbf{x}) = \sum_{A=+,\times} \int df \int d^2\mathbf{n}\, \tilde{h}_A(f,\mathbf{n})\mathbf{e}_A(\mathbf{n}) e^{-2i\pi f(t-\mathbf{n}\cdot\mathbf{x}/c)}$$

- Isotropic, Stationary
- Gaussian ? 'Pop-corn' ?

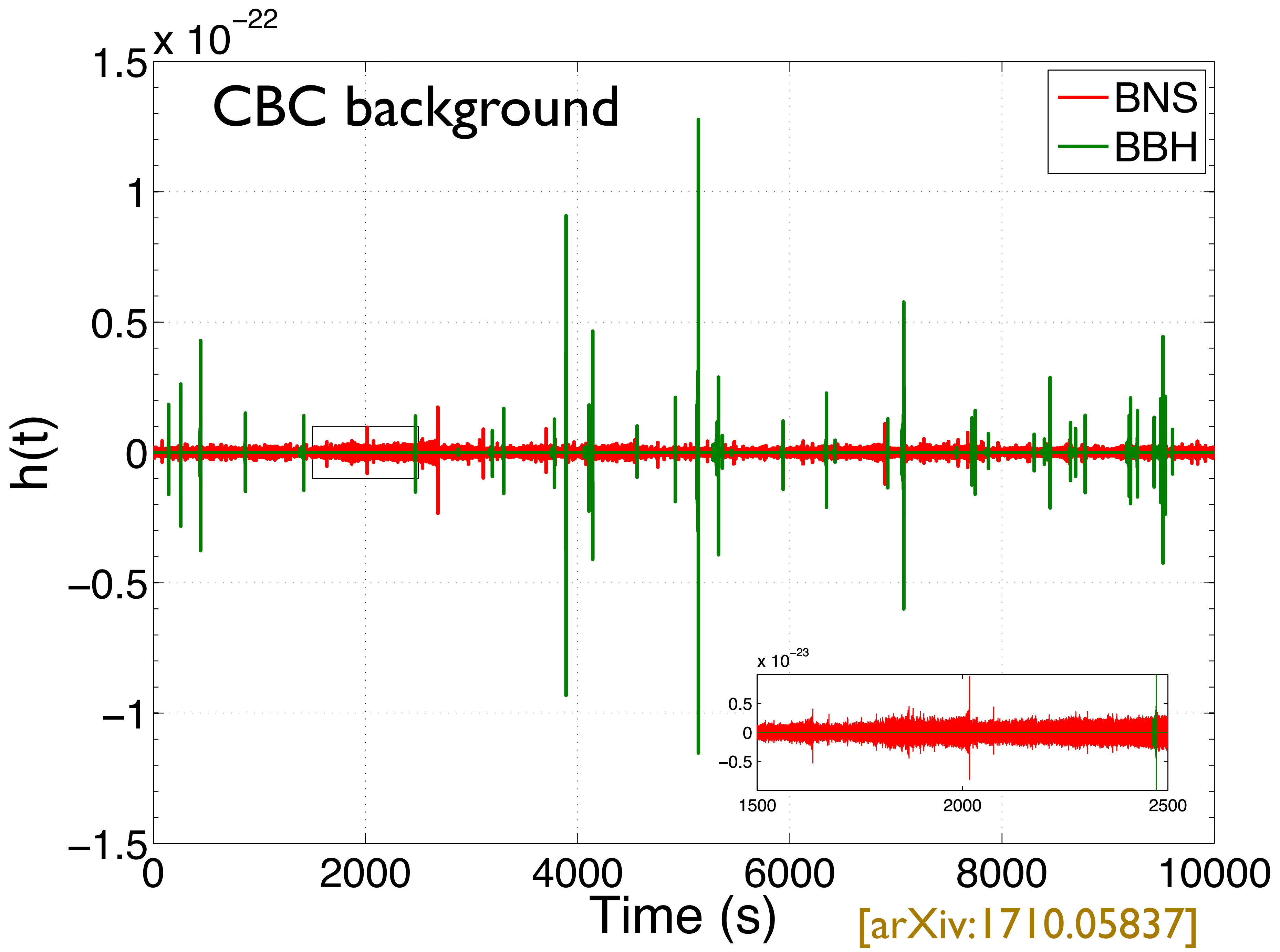
Energy density spectrum of GW background:

$$\Omega_{\rm GW}(f) = \frac{f}{\rho_c}\frac{\mathrm{d}\rho_{\rm GW}}{\mathrm{d}f}$$

- Main target: CBC background
- Backgrounds of cosmological origin



[arXiv:2101.12130]



[arXiv:1710.05837]

Stochastic backgrounds

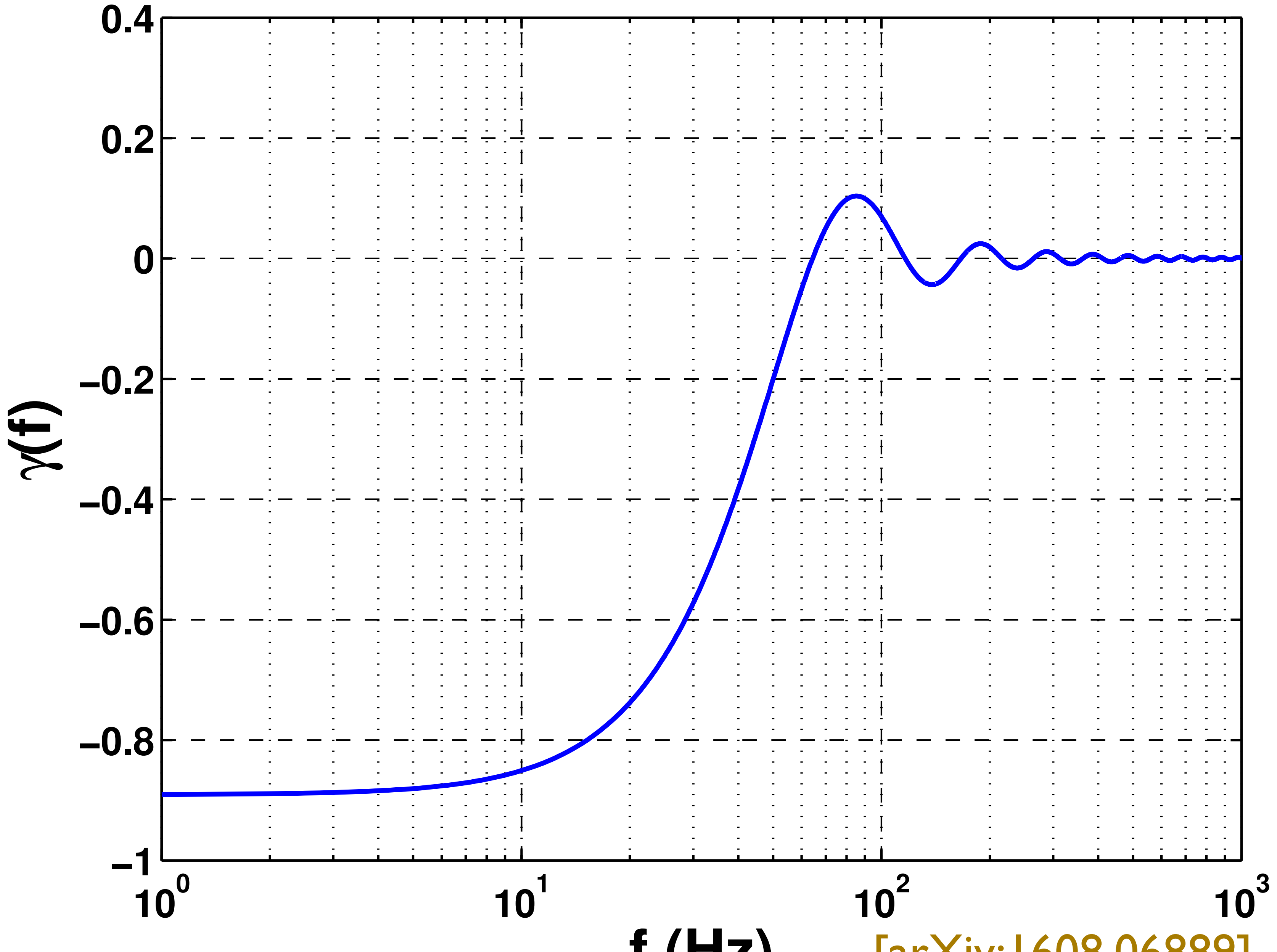
Cross-Correlation between detectors:

$$\hat{C}^{IJ}(f) = \frac{2}{T}\frac{\mathrm{Re}[\tilde{s}_I^\star(f)\tilde{s}_J(f)]}{\gamma_{IJ}(f)S_0(f)}$$

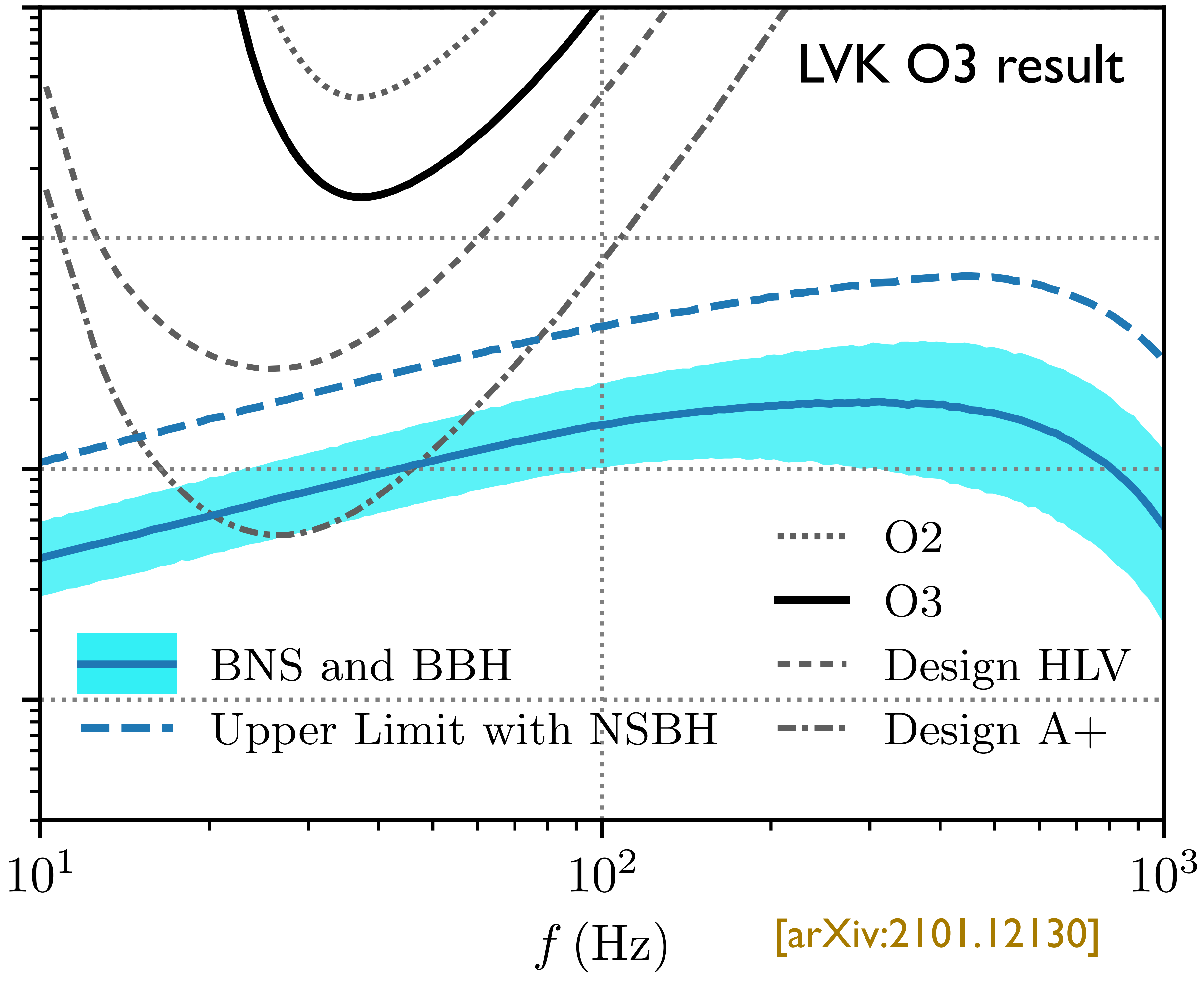
Variance:

$$\sigma^2_{IJ}(f) \approx \frac{1}{2T\Delta f}\frac{P_I(f)P_J(f)}{\gamma^2_{IJ}(f)S_0^2(f)}$$

Overlap reduction function: $\gamma(f)$
at high frequency, signals become incoherent between detectors



[arXiv:1608.06889]



[arXiv:2101.12130]