

---

# Lecture Recording

---

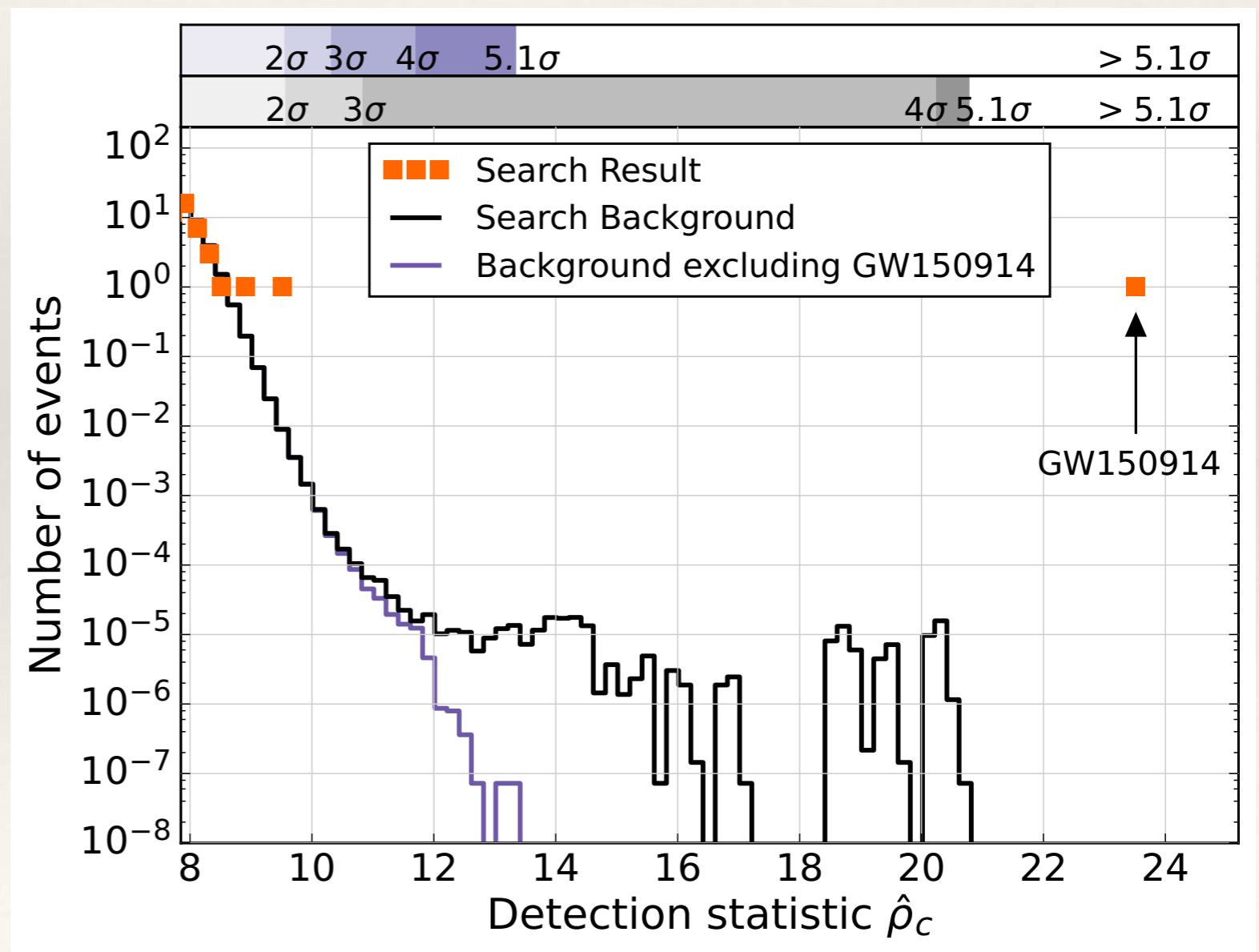
- ❖ **Note: These lectures will be recorded and posted onto the IMPRS website**
- ❖ Dear participants,
- ❖ We will record all lectures on “*Making sense of data: introduction to statistics for gravitational wave astronomy*”, including possible Q&A after the presentation, and we will make the recordings publicly available on the IMPRS lecture website at:
  - <https://imprs-gw-lectures.aei.mpg.de/2021-making-sense-of-data/>
- ❖ By participating in this Zoom meeting, you are giving your explicit consent to the recording of the lecture and the publication of the recording on the course website.

# Making sense of data: introduction to statistics for gravitational wave astronomy

## Lecture 4: Examples of frequentist statistics in GW data analysis

AEI IMPRS Lecture Course

Jonathan Gair [jgair@aei.mpg.de](mailto:jgair@aei.mpg.de)



---

# Likelihood

---

- ❖ The usual assumed model for the output of a GW detector is

$$s(t) = n(t) + h(t; \vec{\lambda})$$

- ❖ and for stationary noise we have (the meaning of  $S_n(f)$  will be discussed further in lecture 9)

$$\langle \tilde{n}^*(f) \tilde{n}(f') \rangle = S_n(f) \delta(f - f')$$

- ❖ If we additionally assume the noise is Gaussian then we can write down a probability distribution for  $s(t)$

$$p(s|\vec{\lambda}) = p(n(t) = s(t) - h(t; \vec{\lambda})) \propto \exp \left[ -\frac{1}{2} (s - h(\vec{\lambda}) | s - h(\vec{\lambda})) \right]$$

- ❖ where

$$(a|b) = \int_{-\infty}^{\infty} \frac{\tilde{a}^*(f) \tilde{b}(f) + \tilde{a}(f) \tilde{b}^*(f)}{S_n(f)} df$$

---

# Fisher Matrix Estimates of Precision

---

- ❖ Recall the Cramer-Rao bound on the variance of unbiased estimators

$$\text{cov}(\hat{\lambda}_i, \hat{\lambda}_j) \geq [\mathbf{\Gamma}_{\theta}^{-1}]_{ij}$$

- ❖ where  $\hat{\lambda}$  is some estimator of the parameter values and

$$\Gamma_{ij} = \mathbb{E} \left[ \frac{\partial l}{\partial \lambda_i} \frac{\partial l}{\partial \lambda_j} \right]$$

- ❖ is the Fisher Information Matrix.
- ❖ For the gravitational wave likelihood

$$\Gamma_{ij} = \left( \frac{\partial h}{\partial \lambda_i} \middle| \frac{\partial h}{\partial \lambda_j} \right)$$

---

# Linear Signal Approximation

---

- ❖ The Fisher matrix provides a lower bound on the variance, or uncertainty, of an estimator.
- ❖ In general, the Fisher matrix provides a good guide to how well parameters can be measured, particularly in the limit of high signal-to-noise ratio. This can be seen in the **linear signal approximation**. If we write

$$s(t) = n(t) + h(t; \vec{\lambda}_0)$$

- ❖ and expand

$$\vec{\lambda} = \vec{\lambda}_0 + \Delta\vec{\lambda} \qquad h(t; \vec{\lambda}) = h(t; \vec{\lambda}_0) + \partial_i h(t; \vec{\lambda}_0) \Delta\lambda^i$$

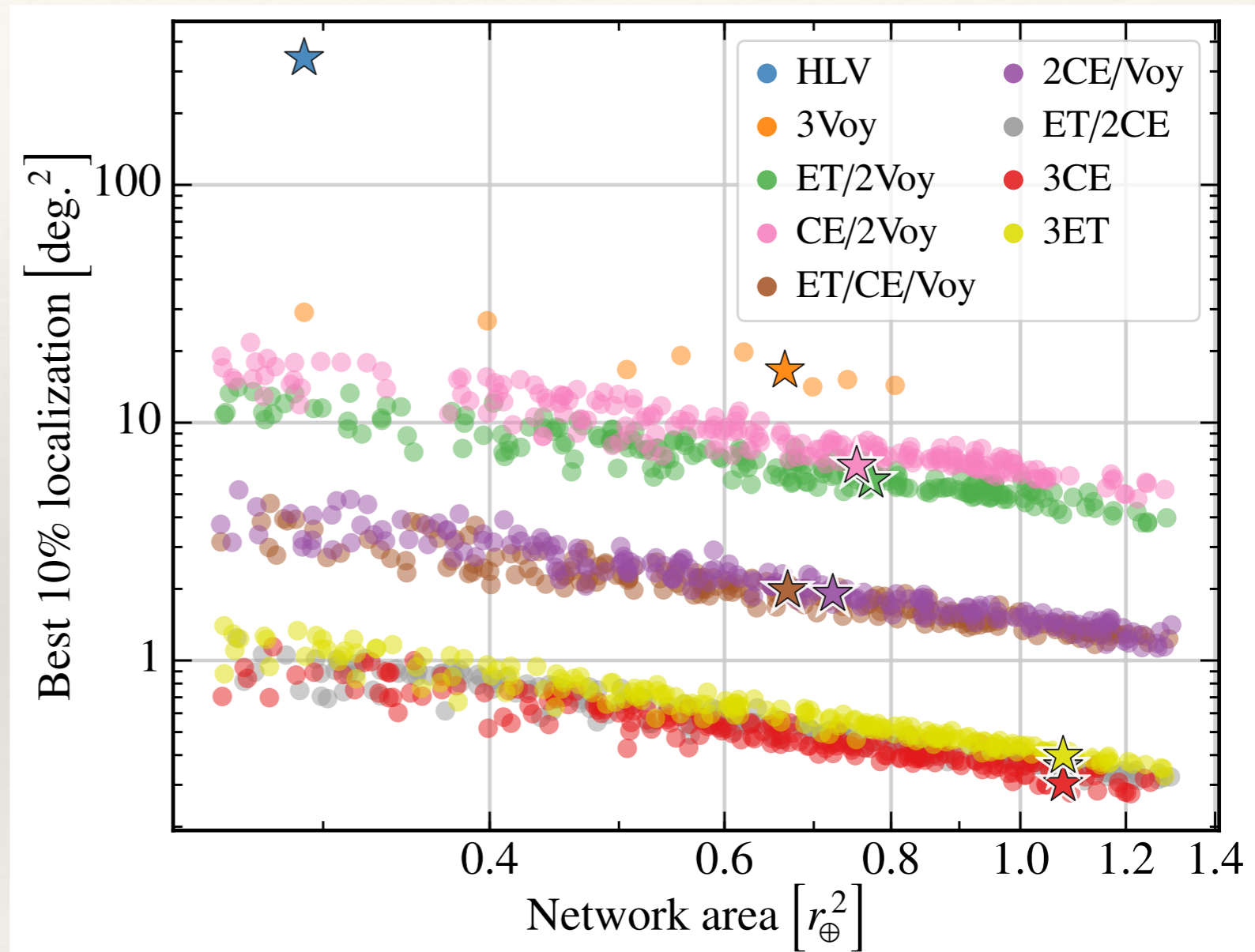
- ❖ we find

$$p(s|\lambda) \propto \exp \left[ -\frac{1}{2} (\Delta\lambda^i - (\Gamma^{-1})_{ik}(n|\partial_k h(t; \lambda_0))) \Gamma_{ij} (\Delta\lambda^j - (\Gamma^{-1})_{jl}(n|\partial_l h(t; \lambda_0))) \right]$$

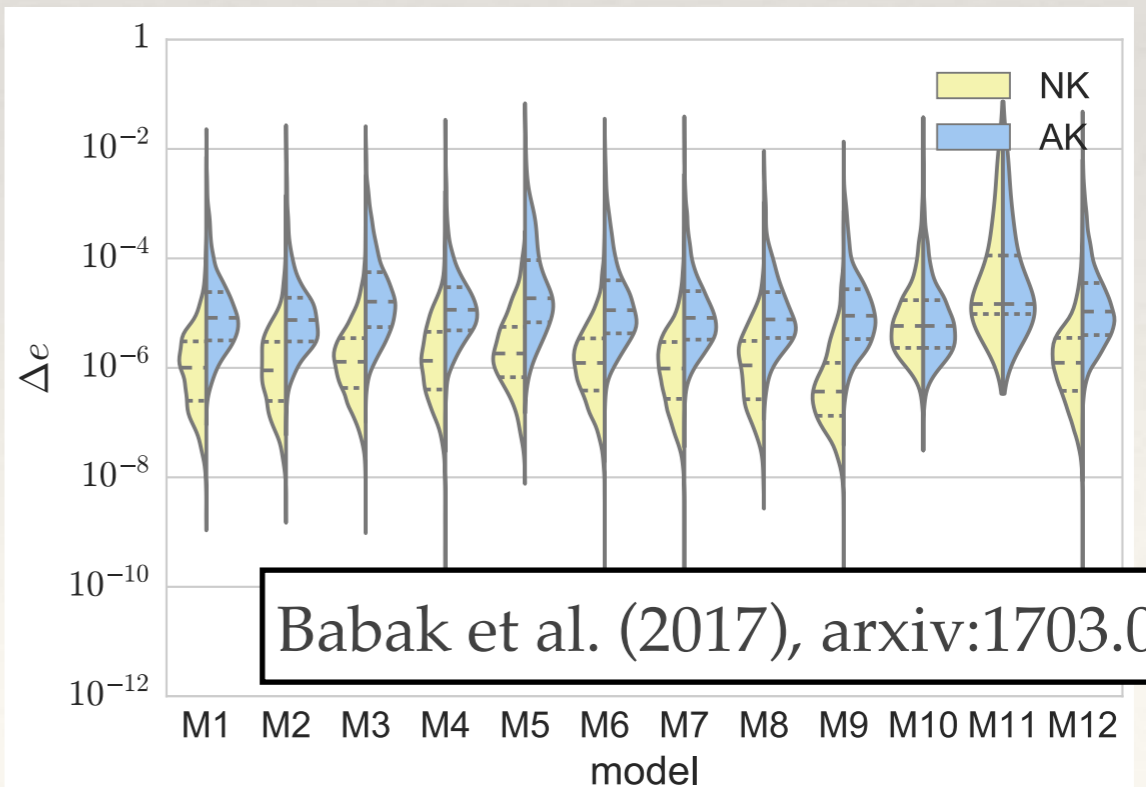
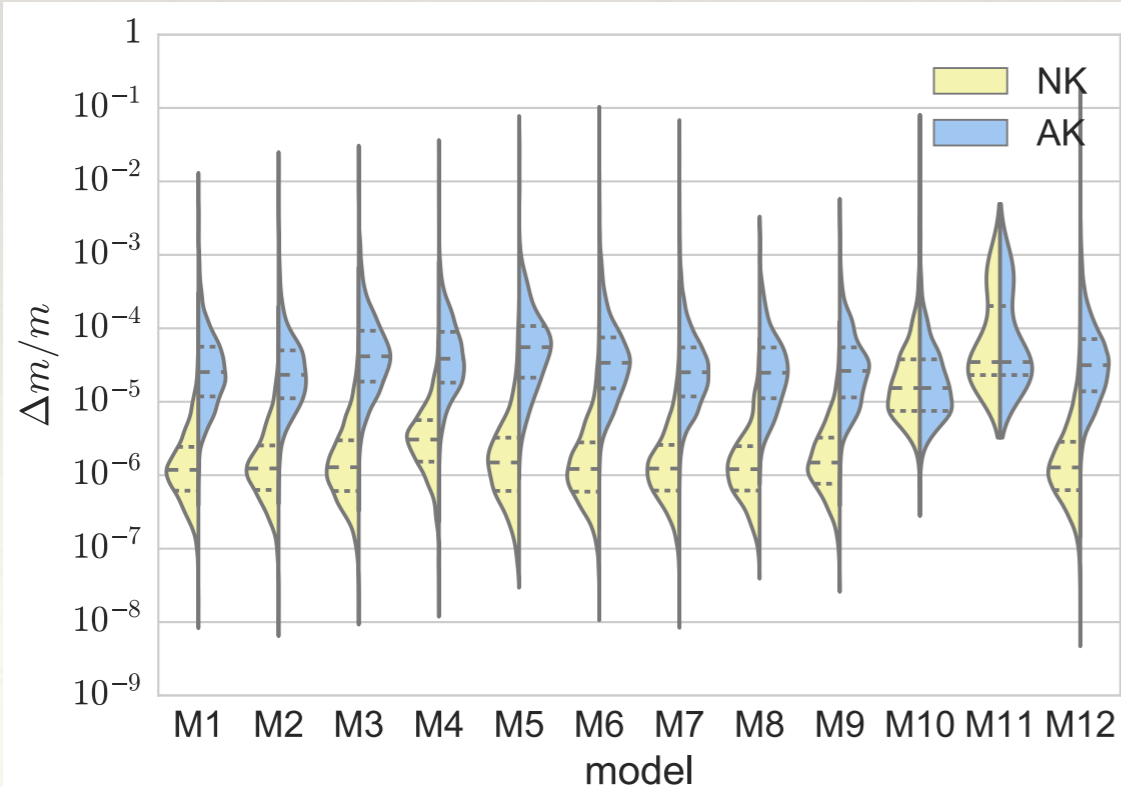
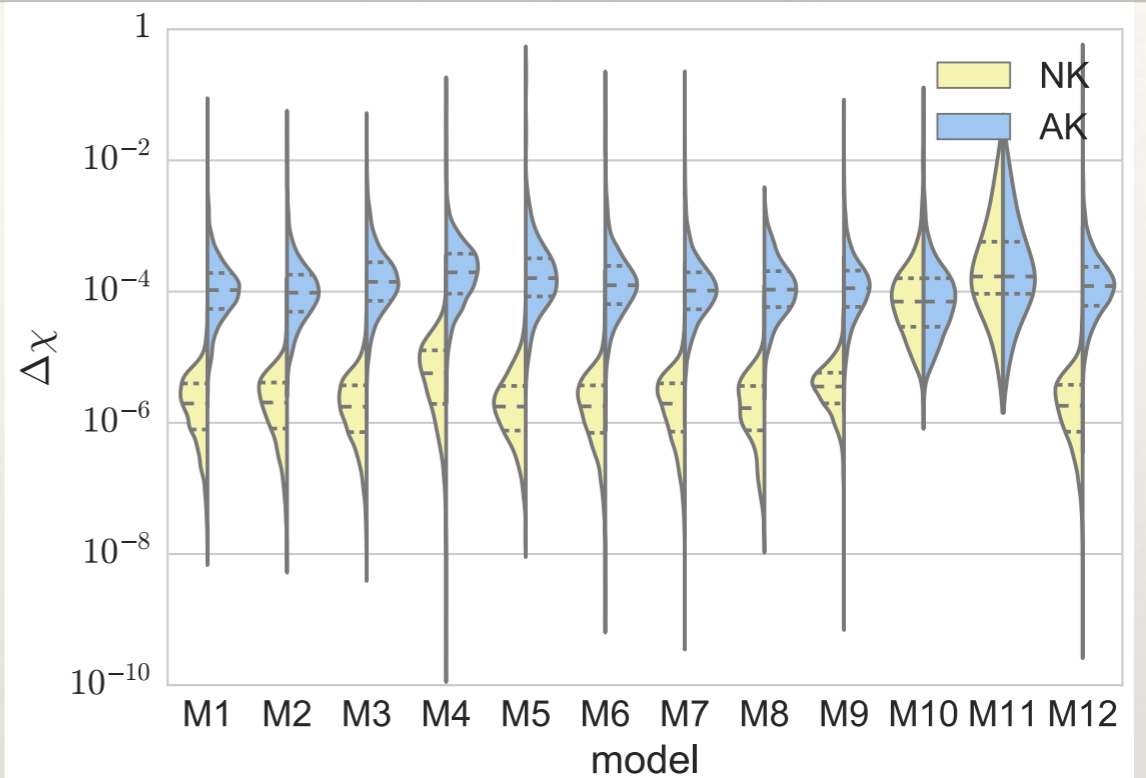
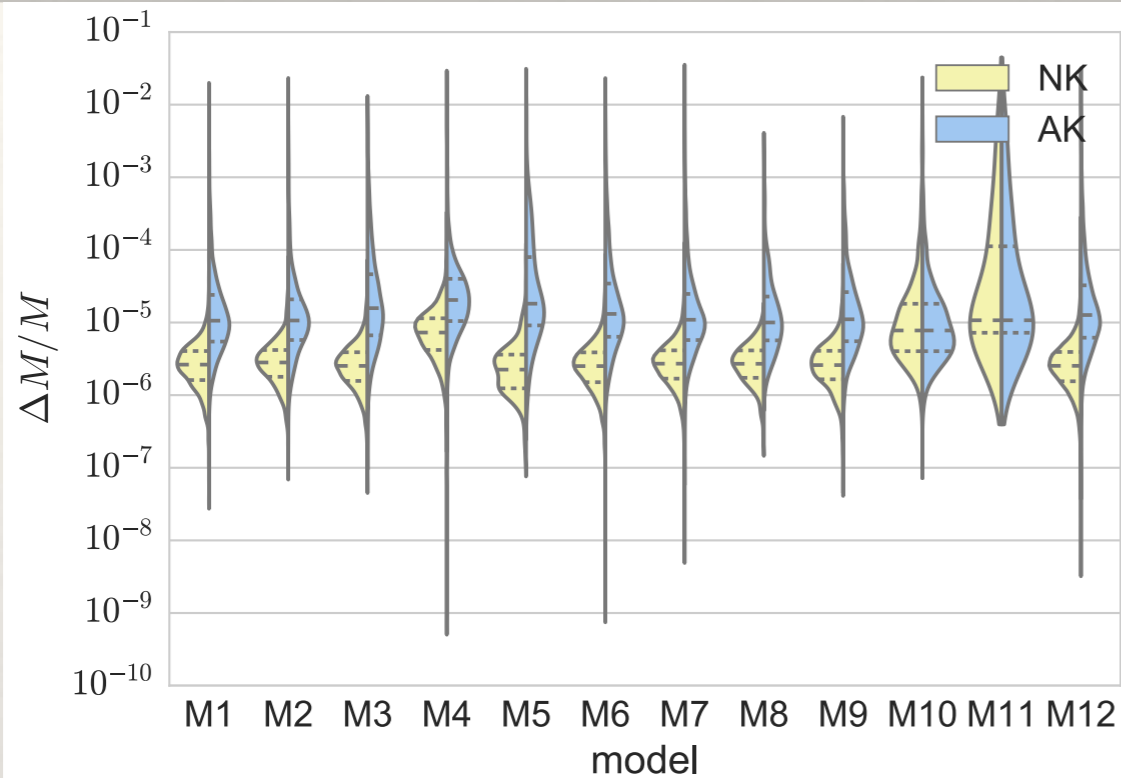
- ❖ where  $\Gamma_{ij}$  is the Fisher Matrix.

# Fisher Matrix Estimates of Precision

- ❖ The Fisher matrix is widely used in a gravitational wave context to characterise our ability to measure parameters.
- ❖ While only an approximation, it is much cheaper to compute than the posterior or MLE distribution and so can be used to survey parameter space much more efficiently.
- ❖ Widely used to scope out science for and optimise design of new facilities.



# Fisher Matrix Estimates of Precision



Babak et al. (2017), arxiv:1703.09722

# Matched filtering/template bank searches



---

# Matched Filtering

---

- ❖ The output of a *filter*, the overlap of the data with a template, is a statistic. The *optimal filter* for finding a known signal is one that matches the signal in the Fourier domain, weighted by the noise PSD (see lecture 9)

$$\tilde{K}(f) = \frac{\tilde{h}(f)}{S_n(f)}$$

- ❖ Considering the log-likelihood

$$\begin{aligned} l(\lambda) &= -\frac{1}{2}(\mathbf{s} - A\hat{\mathbf{h}}(\lambda)|\mathbf{s} - A\hat{\mathbf{h}}(\lambda)) = -\frac{1}{2} \left[ (\mathbf{s}|\mathbf{s}) - 2A(\mathbf{s}|\hat{\mathbf{h}}) + A^2 \right] \\ &= -\frac{1}{2} \left[ (\mathbf{s}|\mathbf{s}) + (A - (\mathbf{s}|\hat{\mathbf{h}}))^2 - (\mathbf{s}|\hat{\mathbf{h}})^2 \right]. \end{aligned}$$

- ❖ We deduce that the maximum of the optimal filter over the parameter space is the **maximum likelihood estimator**.

---

# Matched Filtering

---

- ❖ The optimal filter can also be interpreted in terms of hypothesis testing. Assuming all parameters except the amplitude are known then the question “Is there a gravitational wave in the data” is equivalent to the hypothesis test

$$H_0 : A = 0, \quad \text{vs.} \quad H_1 : A > 0$$

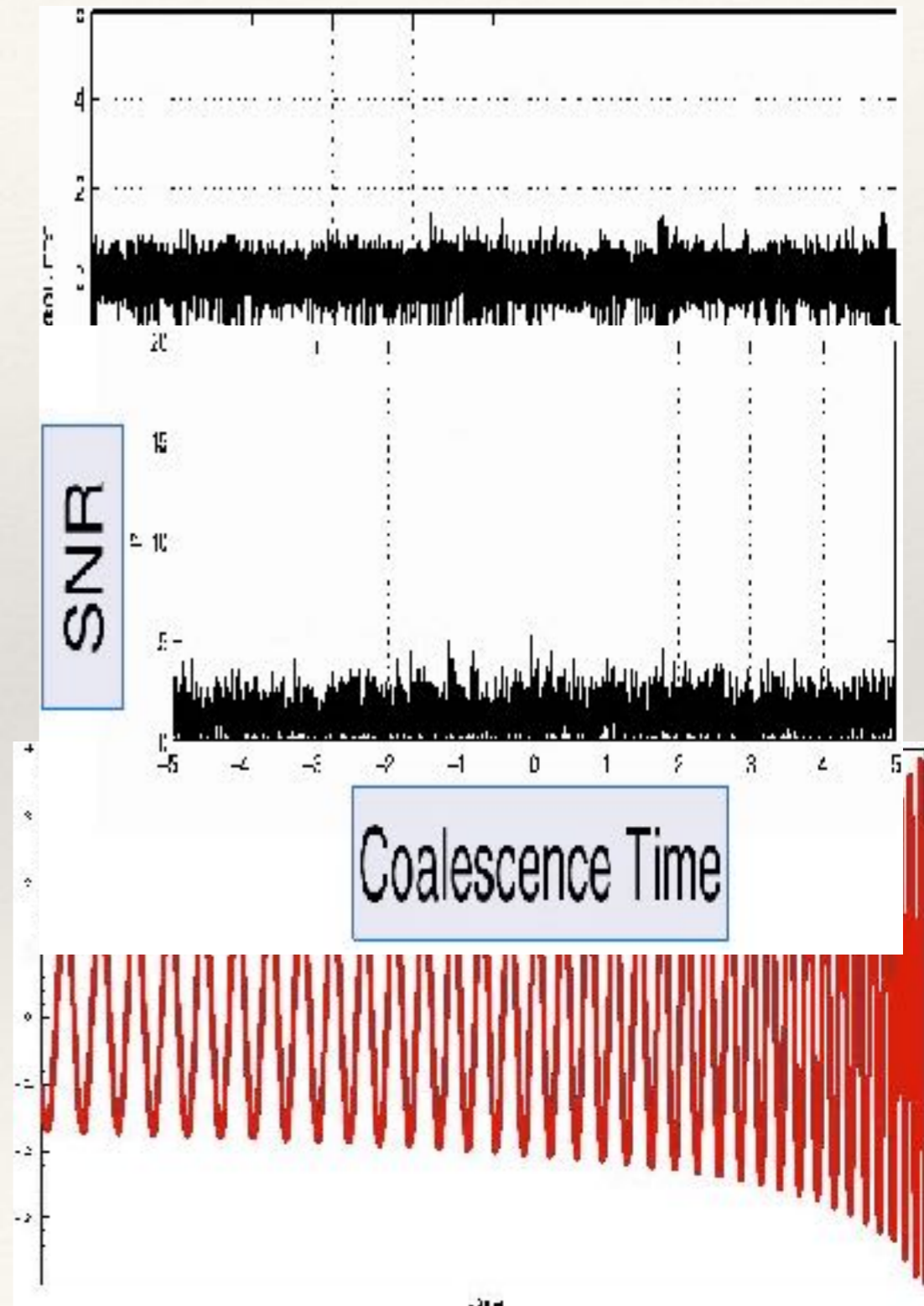
- ❖ From the Neyman-Pearson Lemma, the most powerful statistic for testing  $A=0$  versus  $A=A_1 > 0$  is the likelihood ratio

$$\exp \left[ A_1 (\mathbf{s} | \hat{\mathbf{h}}(\lambda)) - \frac{1}{2} A_1^2 \right]$$

- ❖ We deduce that the optimal filter is also the most powerful test statistic for this hypothesis. As the statistic does not depend on  $A$  it is also uniformly most powerful for the composite hypothesis  $A > 0$ .

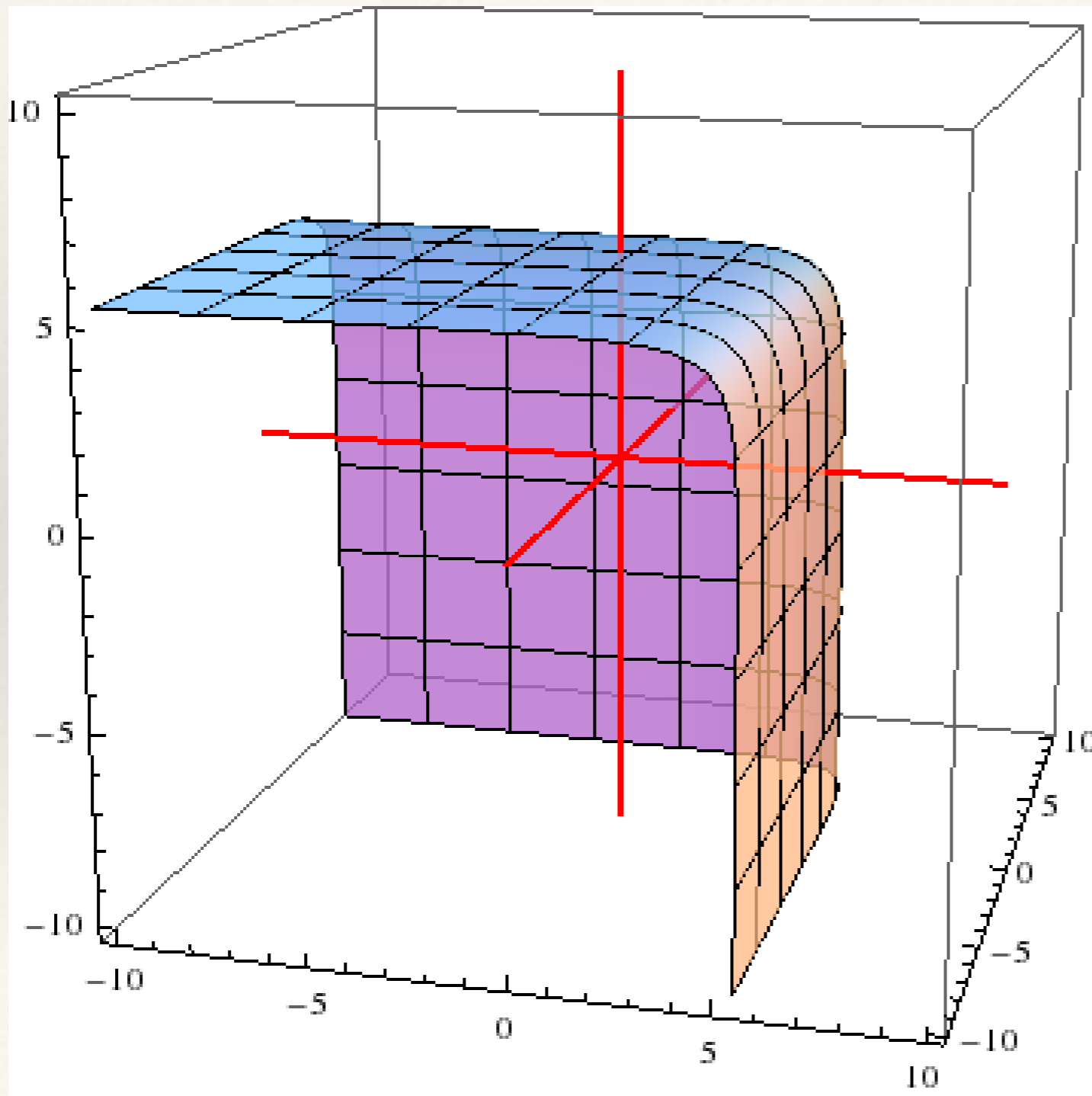
# Matched Filtering

- ❖ Many searches are based on matched filtering.
- ❖ Since the signal is not known, we employ a *template bank* of possible waveforms.
- ❖ If a template in the bank matches a signal in the data, we can pull it out of the noise.
- ❖ The detection statistic is the maximum overlap in the template bank, which is (approximately) maximum likelihood estimator, but it is no longer the most powerful hypothesis test statistic.



# Matched Filtering

- ❖ While it is not the most powerful test statistic, it is a reasonable approximation.



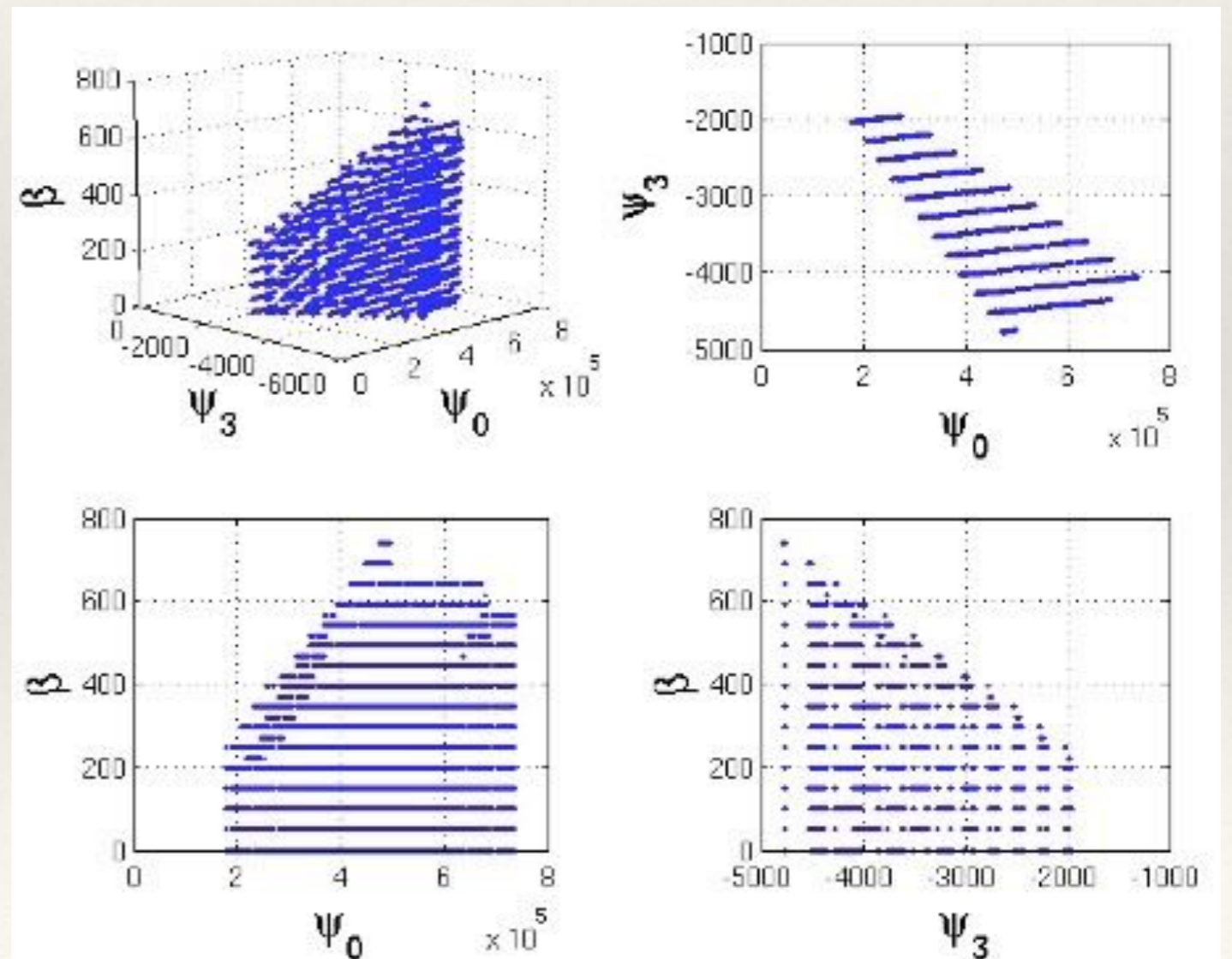
AJK Chua & J Gair (2015)

# Template Bank Construction

- ❖ The Fisher matrix can also be used as a metric to construct a template bank satisfying a *minimal match* criterion

$$\min_{h_{\text{true}}} \max_{h_{\text{temp}}} (h_{\text{true}} | h_{\text{temp}}) \gtrsim 1 - \text{MM}$$

- ❖ Fisher Matrix metric not easy to use in higher dimensional parameter spaces. Now common to use *stochastic banks*.
- ❖ Can also do *stochastic searches* (MCMC) that generate templates on the fly.



---

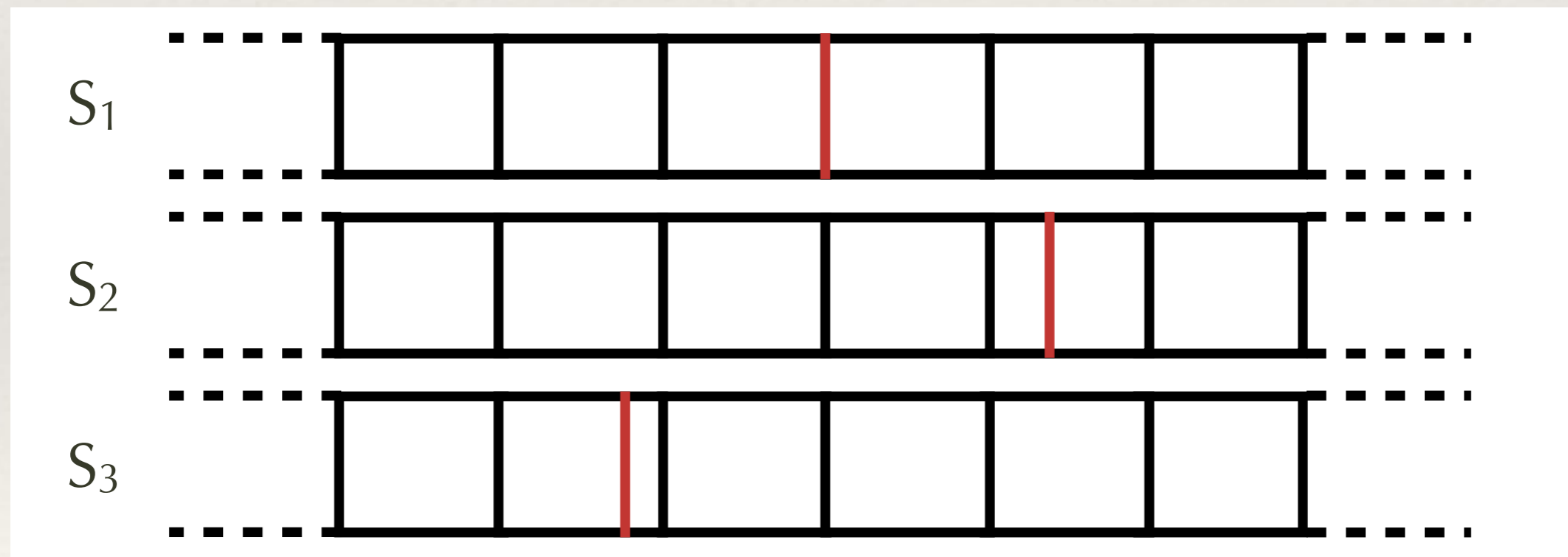
# LIGO searches

---

- ❖ Whether using the optimal filter directly, or refined versions of it, LIGO searches are essentially frequentist. There is a defined **detection statistic** and if the detection statistic is exceeded, the data is flagged as interesting.
- ❖ The **distribution of the detection statistic** is computed in the absence of signals, and the threshold determined based on the probability of a false detection.
- ❖ The **sensitivity** of the search is assessed by injection and recovery of sources into the detector data.
- ❖ The **background distribution** is typically evaluated using **time slides**. This approach relies on the fact that LIGO searches normally use data from more than one detector.

# Background Estimation

- ❖ Noise is not stationary or Gaussian and contains glitches, lines etc.
- ❖ Use frequentist techniques to characterise noise background properties
  - process data in a way that eliminates signal but not noise
  - for LIGO - time slide data from different detectors



- noise + signal coincidences are not background
- significance of events in tail, i.e., sources, is hard to estimate

# False Alarm Rate (FAR)

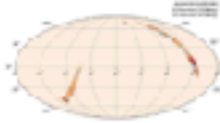
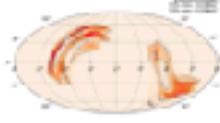
- ❖ In the hypothesis testing part of this course we placed a lot of emphasis on the **significance** or **p-value** of a test. These make sense when we are referring to finite data sets. However, in a gravitational wave context, data is continuously being collected. It therefore makes sense to quote a **false alarm rate (FAR)**.
- ❖ The FAR is the frequency with which triggers of the observed threshold value or higher occur in continuous observation with the detector in the absence of signals.

## LIGO/Virgo O3 Public Alerts

Detection candidates: 46

SORT: EVENT ID (A-Z) ▼



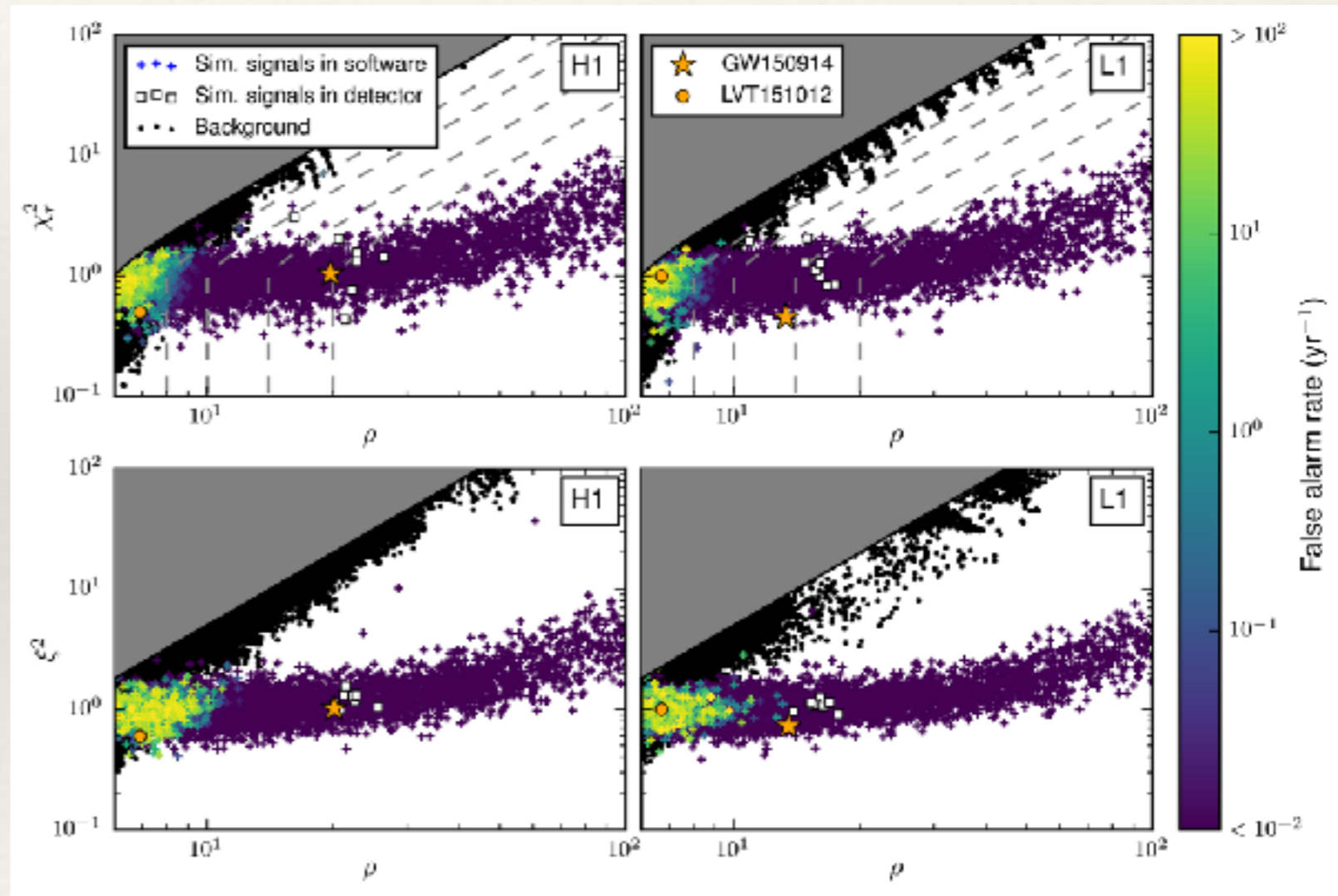
Event ID	Possible Source (Probability)	UTC	GCN	Location	FAR	Comments
<a href="#">S200115j</a>	MassGap (>99%)	Jan. 15, 2020 04:23:09 UTC	<a href="#">GCN Circulars</a> <a href="#">Notices</a>   <a href="#">VOE</a>		1 per 1513 years	
<a href="#">S200114f</a>		Jan. 14, 2020 02:08:18 UTC	<a href="#">GCN Circulars</a> <a href="#">Notices</a>   <a href="#">VOE</a>	No public skymap image found.	1 per 25.838 years	<a href="https://gracedb.ligo.org">https://gracedb.ligo.org</a>
<a href="#">S200112r</a>	BBH (>99%)	Jan. 12, 2020 15:58:38 UTC	<a href="#">GCN Circulars</a> <a href="#">Notices</a>   <a href="#">VOE</a>		1 per 2469.9 years	



# Search refinements: waveform consistency

- ❖ If we subtract the correct template the residual at each frequency should be Normally distributed.
- ❖ Hence the quantity
 
$$\chi^2 = \sum_{k=1}^N \frac{|\hat{s}_k - \hat{h}_k|^2}{S_n(f_k)}$$
- ❖ follows a chi-squared distribution.
- ❖ Construct an *effective SNR* that penalises lack of fit

$$\hat{\rho} = \frac{\rho}{(1 + (\chi^2/N)^3)^{\frac{1}{6}}}$$



LVC, *Phys. Rev. D* 93, 122003 (2016)

---

# Search refinements: phase and Time Parameters

---

- ❖ Certain parameters can be maximised over cheaply, e.g., unknown phase

$$h(t; A, f_0, t_c, \phi_0) = A \cos(2\pi f_0(t - t_c) + \phi_0)$$

$$\max_{\phi_0} (s|h)^2 = A^2 \left( (s|h(t; A, f_0, t_c, 0))^2 + (s|h(t; A, f_0, t_c, -\pi/2))^2 \right)$$

- ❖ and unknown coalescence time

$$\tilde{h}(f; A, f_0, t_c, \phi_0) = \tilde{h}(f; A, f_0, 0, \phi_0) \exp(-2\pi i f t_c)$$

$$(s|h(t; A, f_0, t_c, \phi_0)) = 2\text{Re} \int_{-\infty}^{\infty} \frac{\tilde{s}^*(f) \tilde{h}(f; A, f_0, 0, \phi_0)}{S_n(f)} \exp(-2\pi i f t_c) df$$

- ❖ This is the inverse Fourier transform of  $\tilde{s}^*(f) \tilde{h}(f; A, f_0, 0, \phi_0) / S_n(f)$ . Obtain overlap for all time offsets cheaply using a Fast Fourier Transform. **Maximisation** is a frequentist approach. The Bayesian equivalent would be **marginalisation**.

---

# The F-statistic

---

- ❖ The **F-statistic** extends the idea of phase maximisation to more of the extrinsic parameters. We write the waveform as a sum over four basis functions

$$h(t) = \sum_{i=1}^4 a_i (\iota, \psi, D_L, \varphi_c) A^i (t; M_c, \mu, t_c, \theta, \phi)$$

- ❖ where

$$a_1 = \Lambda \left[ (1 + \cos^2 \iota) \cos 2\psi \cos \varphi_c - 2 \cos \iota \sin 2\psi \sin \varphi_c \right]$$

$$a_2 = -\Lambda \left[ (1 + \cos^2 \iota) \sin 2\psi \cos \varphi_c + 2 \cos \iota \cos 2\psi \sin \varphi_c \right]$$

$$a_3 = \Lambda \left[ (1 + \cos^2 \iota) \cos 2\psi \sin \varphi_c + 2 \cos \iota \sin 2\psi \cos \varphi_c \right]$$

$$a_4 = -\Lambda \left[ (1 + \cos^2 \iota) \sin 2\psi \sin \varphi_c - 2 \cos \iota \cos 2\psi \cos \varphi_c \right]$$

- ❖ and

$$A^1 = m_o \eta x(t) D^+ \cos(\varphi)$$

$$A^2 = m_o \eta x(t) D^\times \cos(\varphi)$$

$$A^3 = m_o \eta x(t) D^+ \sin(\varphi)$$

$$A^4 = m_o \eta x(t) D^\times \sin(\varphi)$$

---

# The F-statistic

---

- ❖ Treating the amplitudes of the components of the waveform as independent, we can find the best-fit values via

$$a_i = M_{ij} N^j \quad N^i = \langle s | A^i \rangle \quad M_{ij} = (M^{ij})^{-1} = \langle A^i | A^j \rangle^{-1}$$

- ❖ Substituting this into the likelihood

$$\ln \mathcal{L}(\vec{x}) = \langle s | h(\vec{x}) \rangle - \frac{1}{2} \langle h(\vec{x}) | h(\vec{x}) \rangle$$

- ❖ gives

$$\mathcal{F} = \frac{1}{2} M_{ij} N^i N^j$$

- ❖ This is the F-statistic. It depends only on intrinsic parameters, but can be used directly in a template bank instead of including templates in the extrinsic parameters as well.
- ❖ An estimate for the extrinsic parameters can be obtained from the maximized values of the  $a_i$  coefficients.

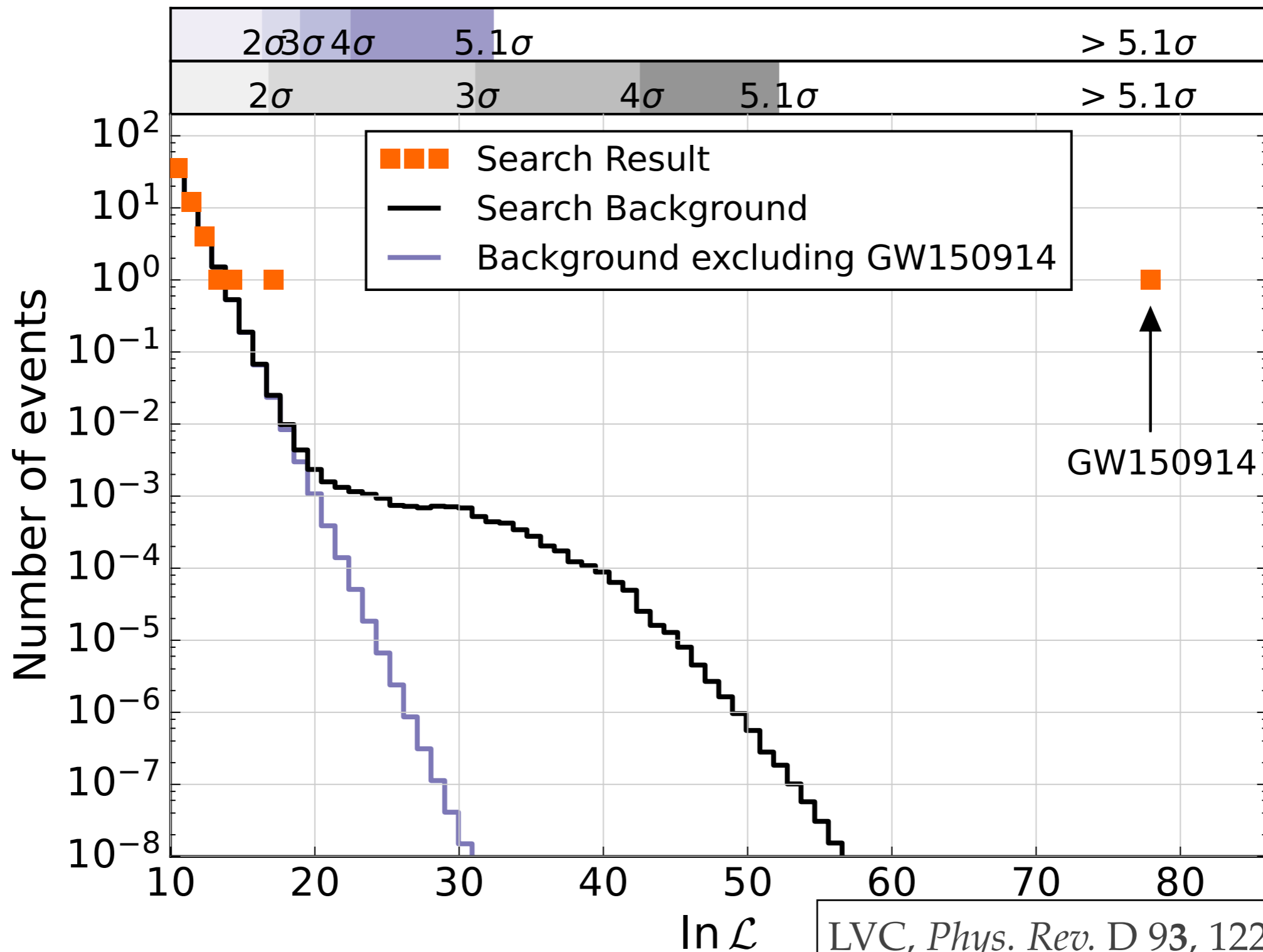
---

# LIGO Pipelines

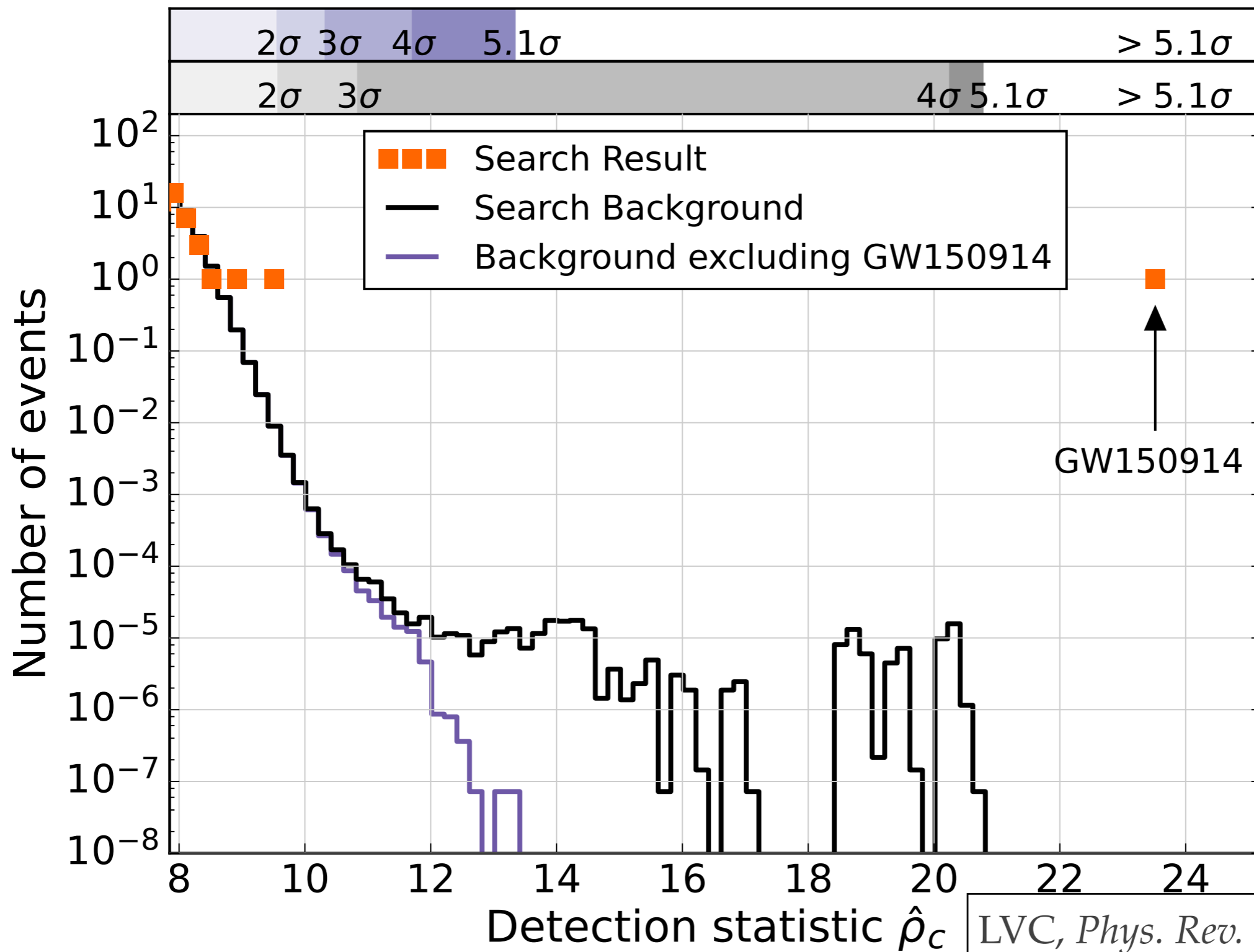
---

- ❖ Two main matched filtering pipelines used in LIGO for compact binary coalescence searches.
  - ❖ *pycbc*: constructs template bank of waveforms; computes chi-squared test for fit; uses effective SNR as a ranking statistic; background computed using time slides.
  - ❖ *gstLAL*: constructs template bank of waveforms, then does SVD decomposition to form a signal basis; detection statistic is likelihood ratio for signal versus noise; waveform consistency assessed by comparing SNR time series to theory; time slides again used to assess background.

# LIGO Pipelines



# LIGO Pipelines



---

# PSD Estimation

---

- ❖ Matched filter is noise-weighted. OK if you know the noise PSD, but in general we will not. For LIGO, estimate this using off-source data.

$s_{-3} = n_{-3}$	$s_{-2} = n_{-2}$	$s_{-1} = n_{-1}$	$s_0 = h + n_0$	$s_1 = n_1$	$s_2 = n_2$	$s_3 = n_3$
-------------------	-------------------	-------------------	-----------------	-------------	-------------	-------------

$$\sigma_0^2 \approx \frac{1}{2N} \sum_{k=1}^N (s_k^2 + s_{-k}^2)$$

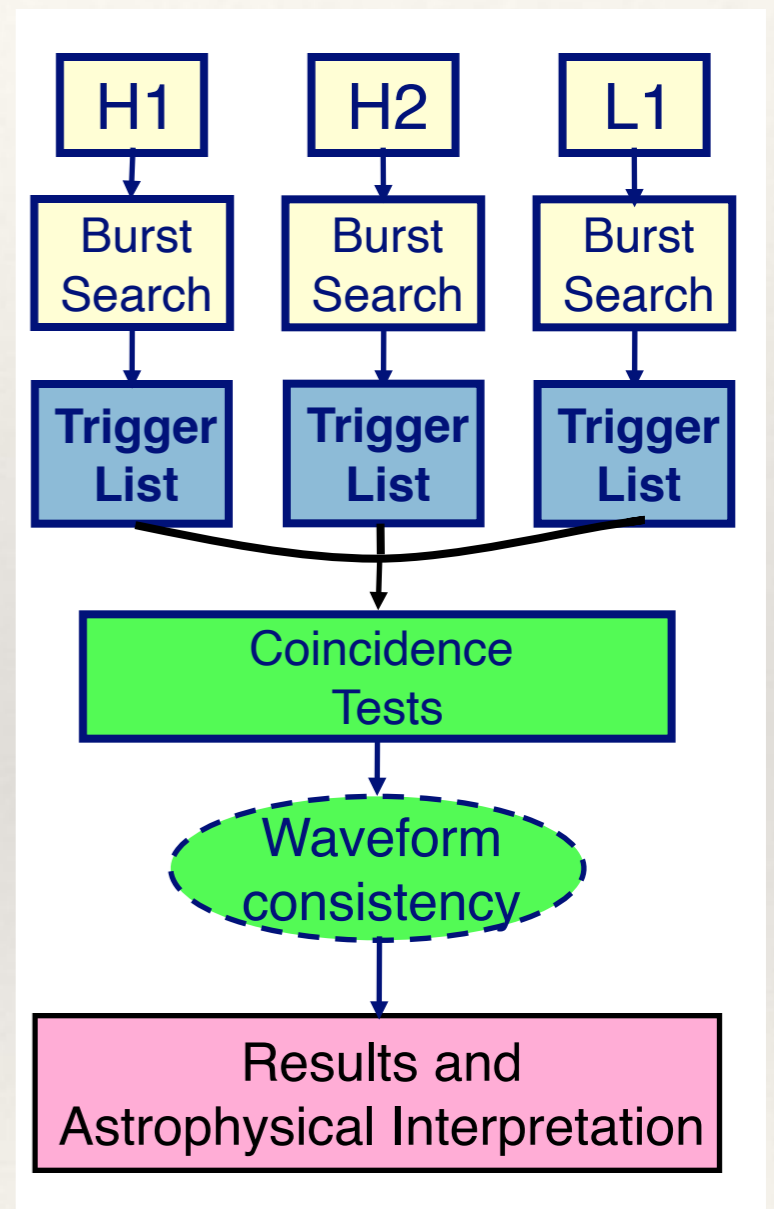
- ❖ In practice, use median of noise estimates, rather than the average. This is less sensitive to non-stationarities.
- ❖ No off-source data for LISA. Make progress by fitting noise and signal properties simultaneously - need reasonable noise model.



# Unmodelled/Excess power searches

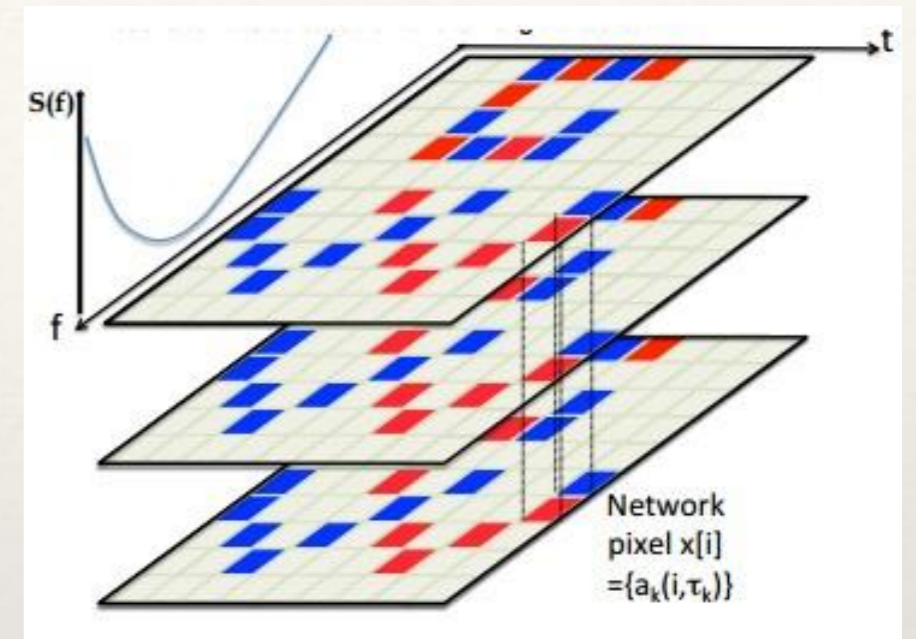
# Unmodelled Searches

- ❖ Detection of gravitational wave bursts relies on two techniques
  - *Coincidence analysis*. As for stochastic background, combine data from multiple detectors. Likelihood of an instrumental artefact in two detectors simultaneously is small.
  - *Time-frequency analysis*. Look for changes in spectral properties over time, e.g., excess power in a set of connected pixels.
- ❖ Basic idea: construct *time-frequency* spectrograms of the data, i.e., estimate power at each frequency and time. Use spectrograms at multiple resolutions to give sensitivity to different burst morphologies.
- ❖ Look for clusters of pixels coincident between instruments.



# Unmodelled Searches: Coherent Wave Burst

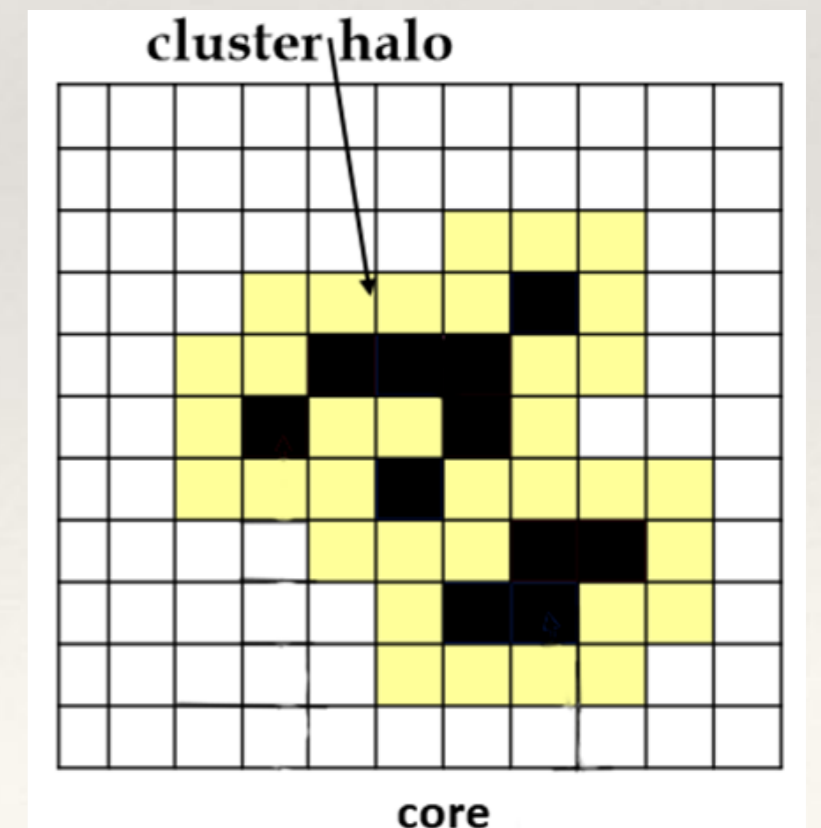
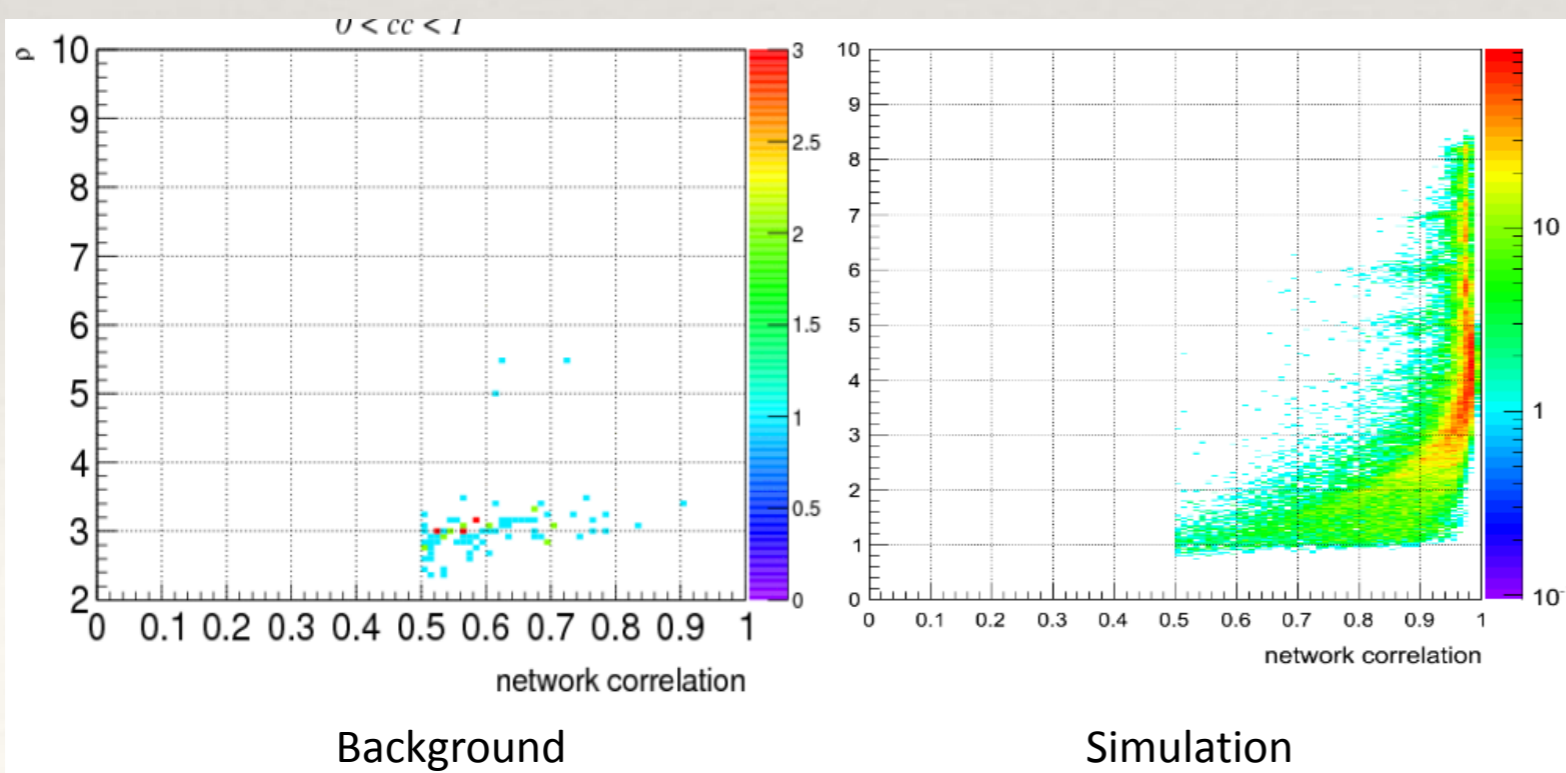
- ❖ Combines spectrograms at multiple resolutions. Identifies pixel clusters.
- ❖ Uses various derived quantities to distinguish signals from noise artefacts, e.g., coherent and residual noise energies.



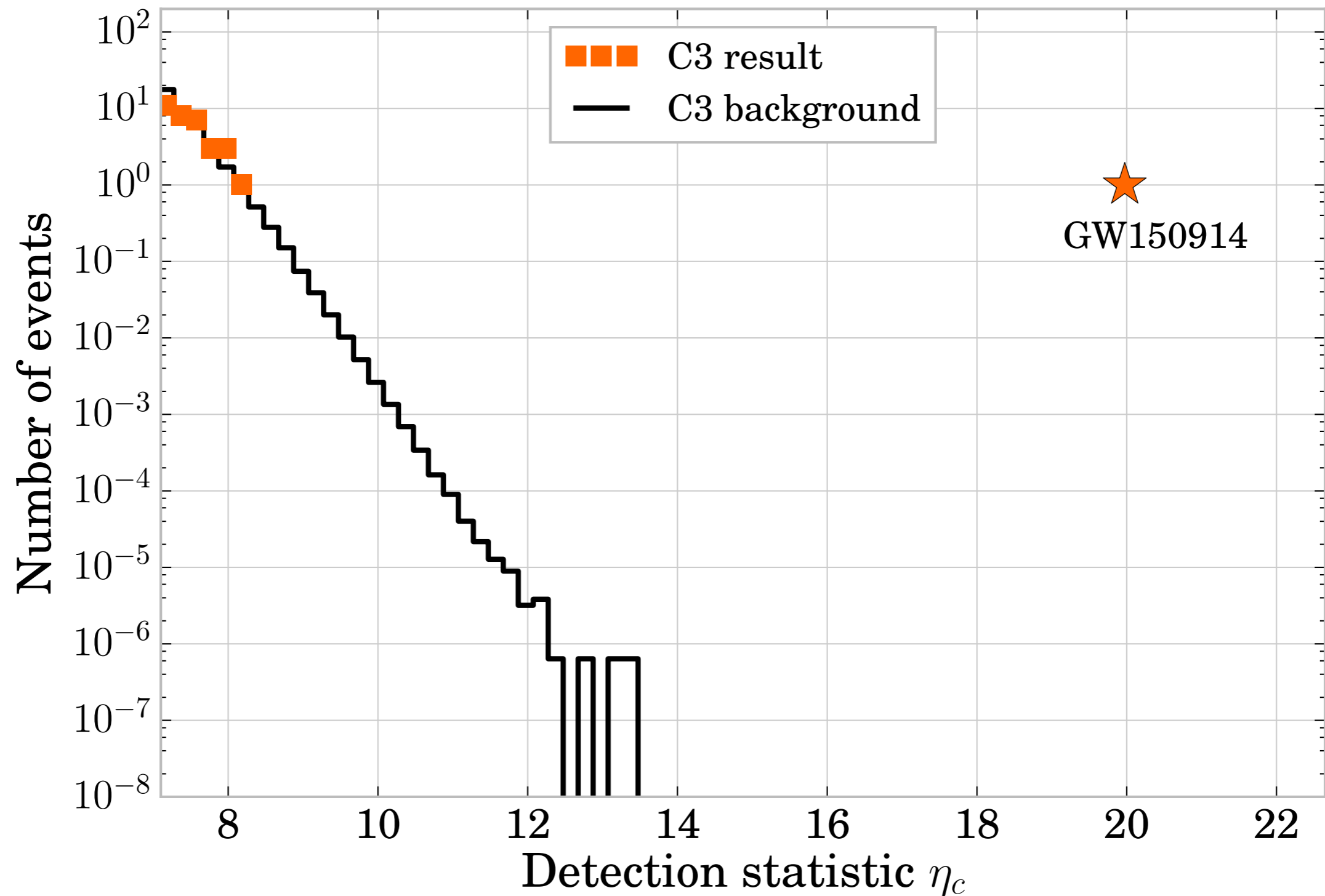
$$E_c = \sum_{m \neq n} L_{mn}$$

$$N = |X - \xi_\sigma|^2$$

$$cc = \frac{E_c}{N + E_c}$$

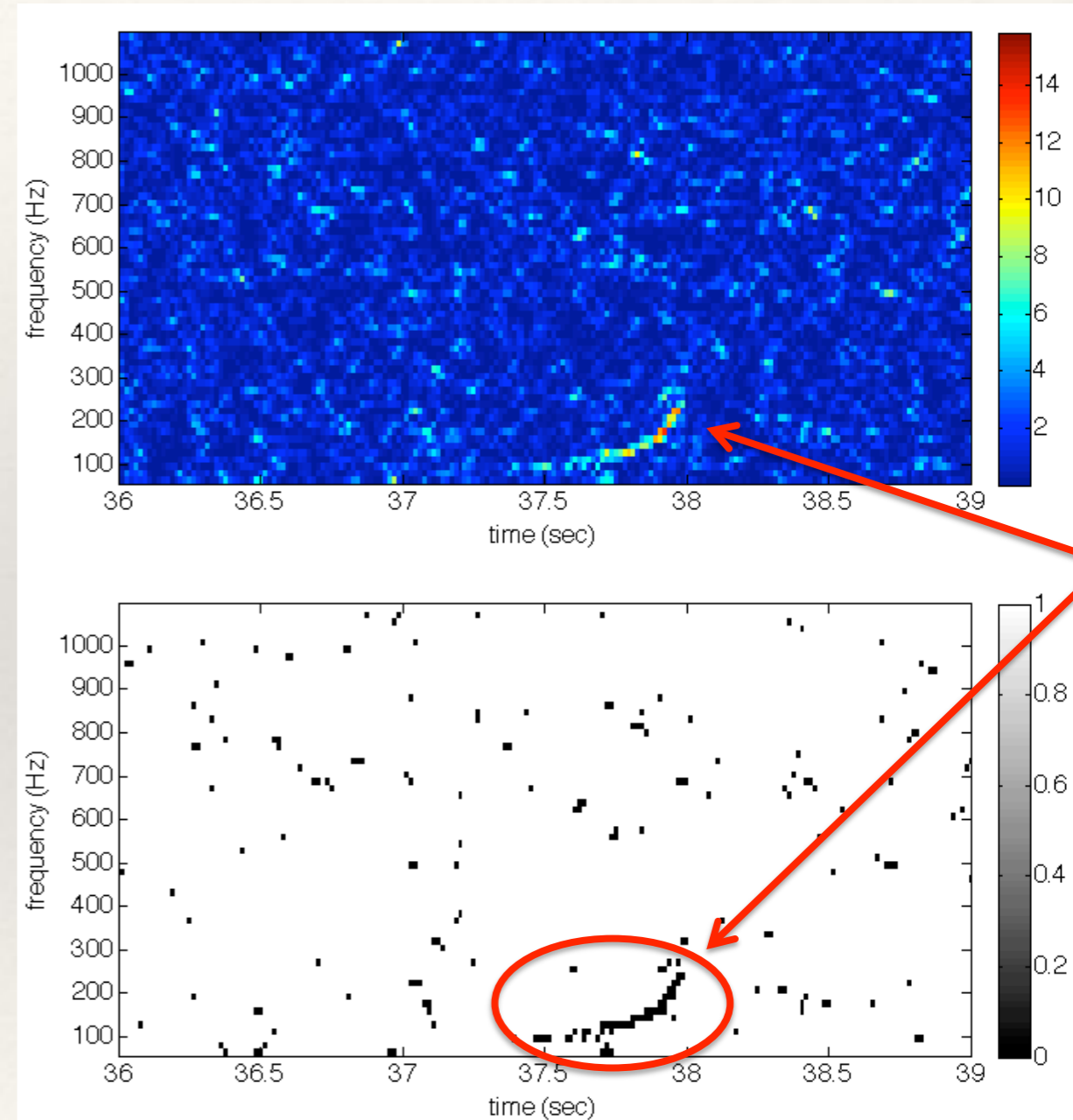


# Unmodelled Searches: Coherent Wave Burst



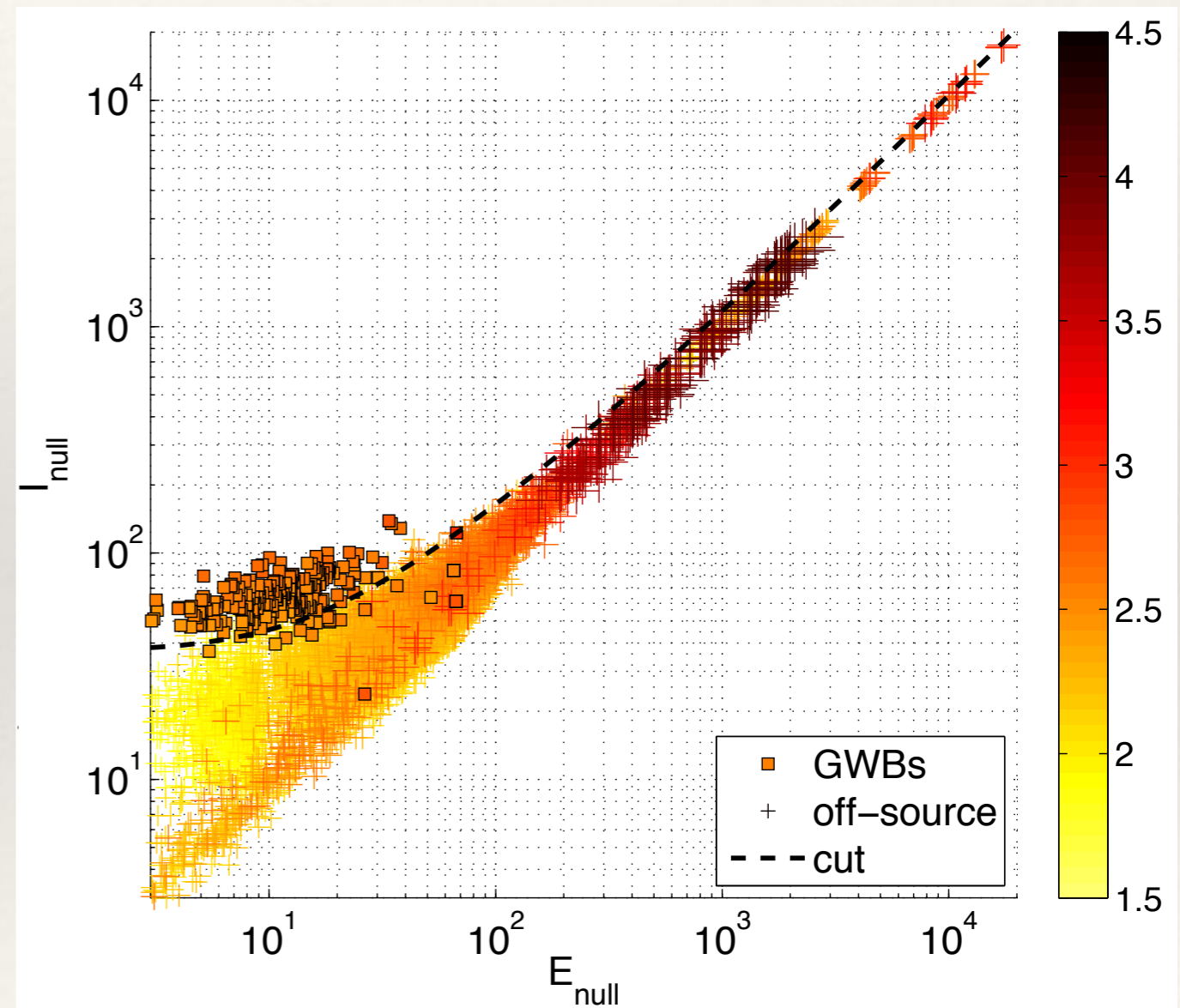
# Unmodelled Searches: X-pipeline

- ❖ *X-pipeline* uses similar methods to CWB, but different implementation.
- ❖ Data is whitened and FFT'd at multiple resolutions. Data from different detectors is then summed to construct various energy measures.
- ❖ High energy (black) pixels are identified and clustered.
- ❖ Each event characterised by certain parameters - start time, peak time, duration, start frequency, peak frequency, bandwidth, number of pixels, energy measures, sky position, FFT length etc.
- ❖ Loop over sky positions and length of FFTs.

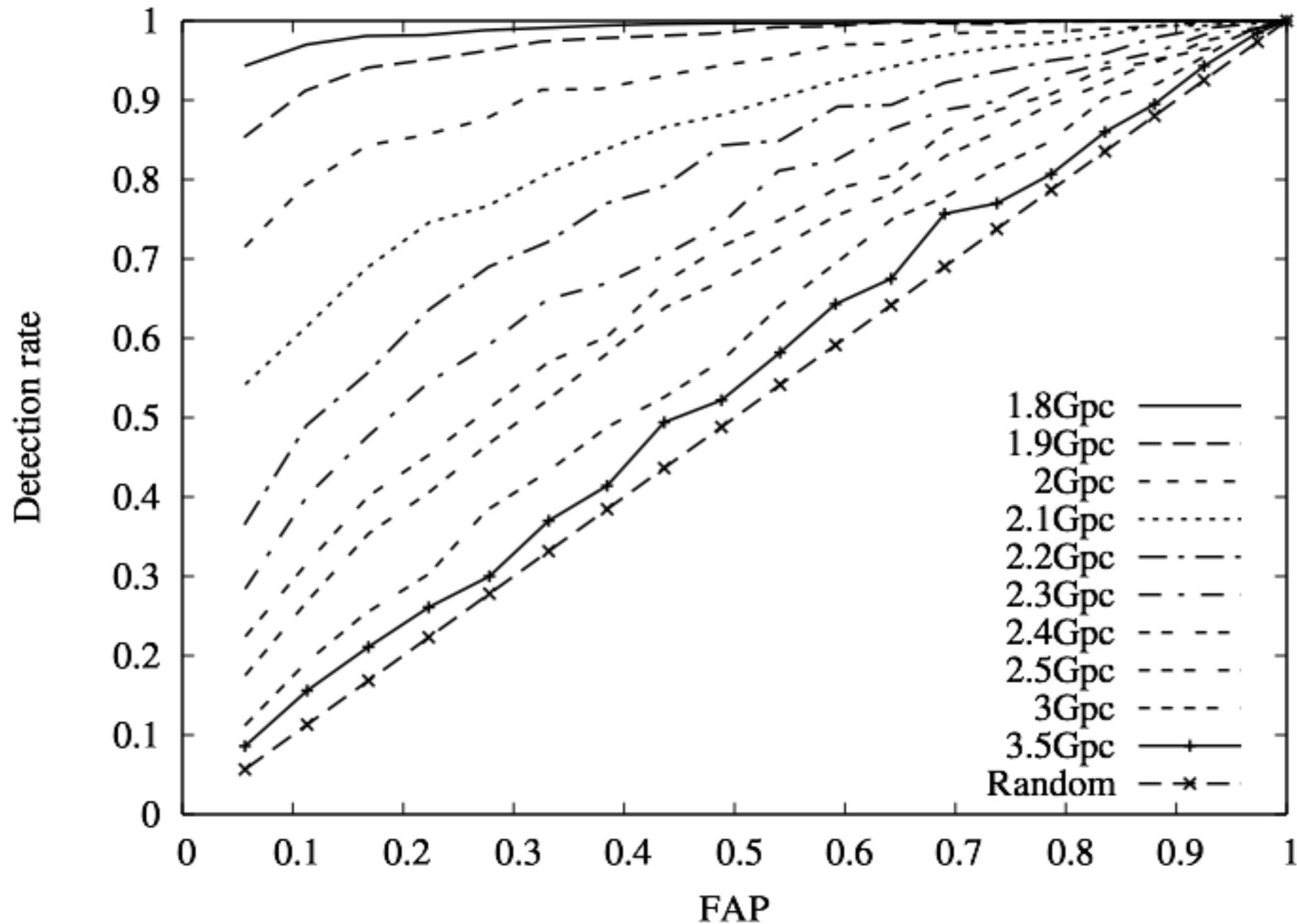


# Unmodelled Searches: X-pipeline

- ❖ Analysis is in two stages. Trigger generation, as described above, then post processing.
- ❖ Post processing involves rejecting background events based on event properties, and assessment of search efficiency.
- ❖ Rejection uses different combinations of energy measures, based on randomly selected training set of injections and time slides.



# Unmodelled Searches: LISA



# Semi-coherent search methods



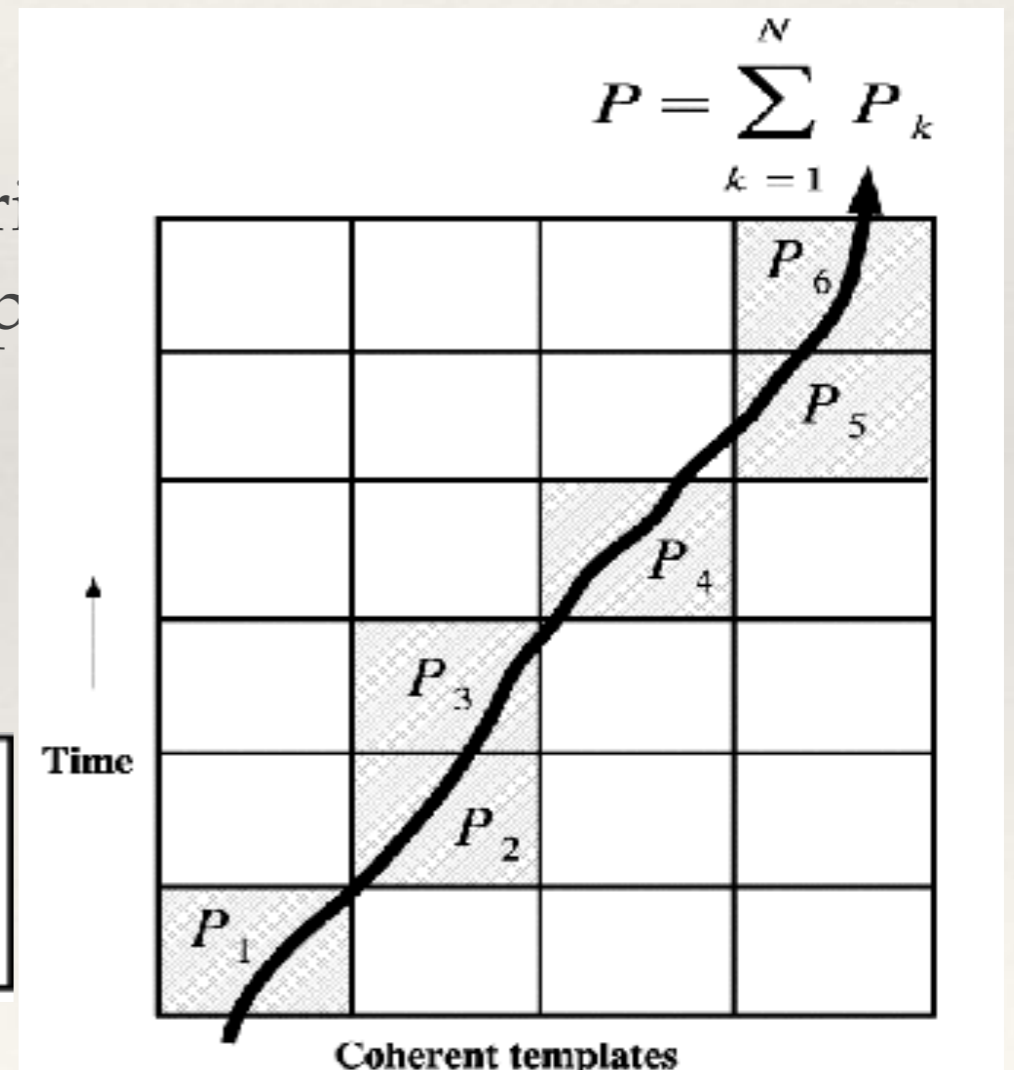
# Semi-coherent searches: EMRIs

- ❖ First stage is coherent matched filtering of shorter (~2 week) waveform segments. Segment length set by computational limits.
- ❖ Second stage involves incoherent summation of maximized power along trajectories through the segments.

- ❖ Performance analysed theoretically to derive rates. Computational cost has prevented practical

$$\rho^2 = \sum_{\alpha=1}^I \sum_{i=1}^5 \langle h_i(\lambda_I), s_\alpha \rangle^2$$

where  $\langle a, b \rangle = 4 \Re \left[ \int_0^\infty \frac{\tilde{a}^*(f) \tilde{b}(f)}{S_b(f)} df \right]$



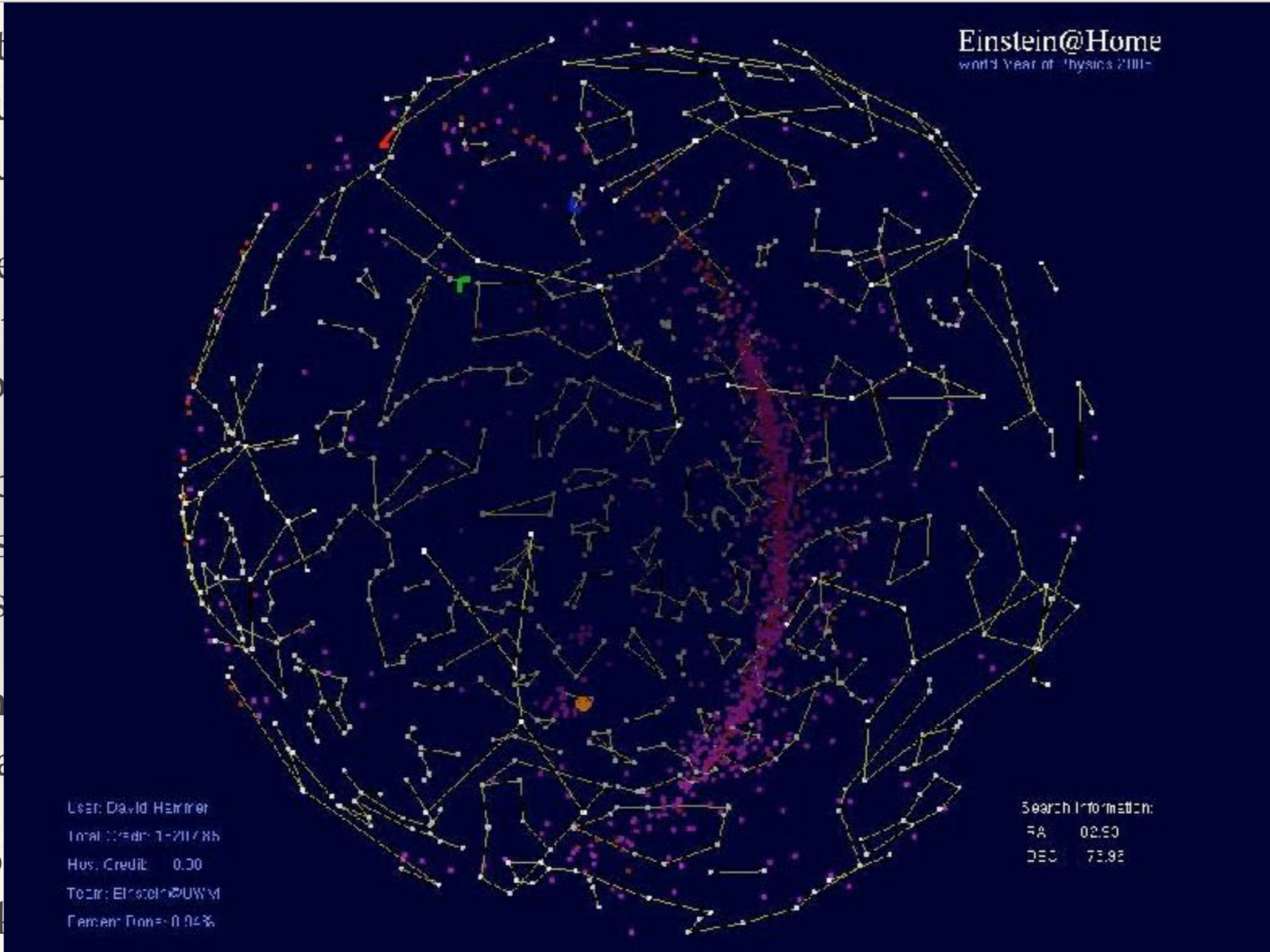
---

# Semi-coherent searches: pulsars

---

- ❖ LIGO unknown pulsar search also uses semi-coherent techniques.
- ❖ *Stack-Slide* algorithm is very similar to EMRI algorithm described above.
- ❖ *Hough Transform* applies the Hough Transform, a well-established technique for detecting simple shapes (edges) in an image, to the output of the coherent stage of the search.
- ❖ Requires a huge amount of computer power - Einstein@home.

- ❖ In the...  
hou...  
You...
- ❖ The...  
Ne...  
Wo...
- ❖ Eac...  
pos...  
ens...
- ❖ Ein...  
tota...
- ❖ No...  
unl...



pu  
arch.

e

r sky

nd a

# Searching for Backgrounds

---

# Stochastic Gravitational Wave Fore/Backgrounds

---

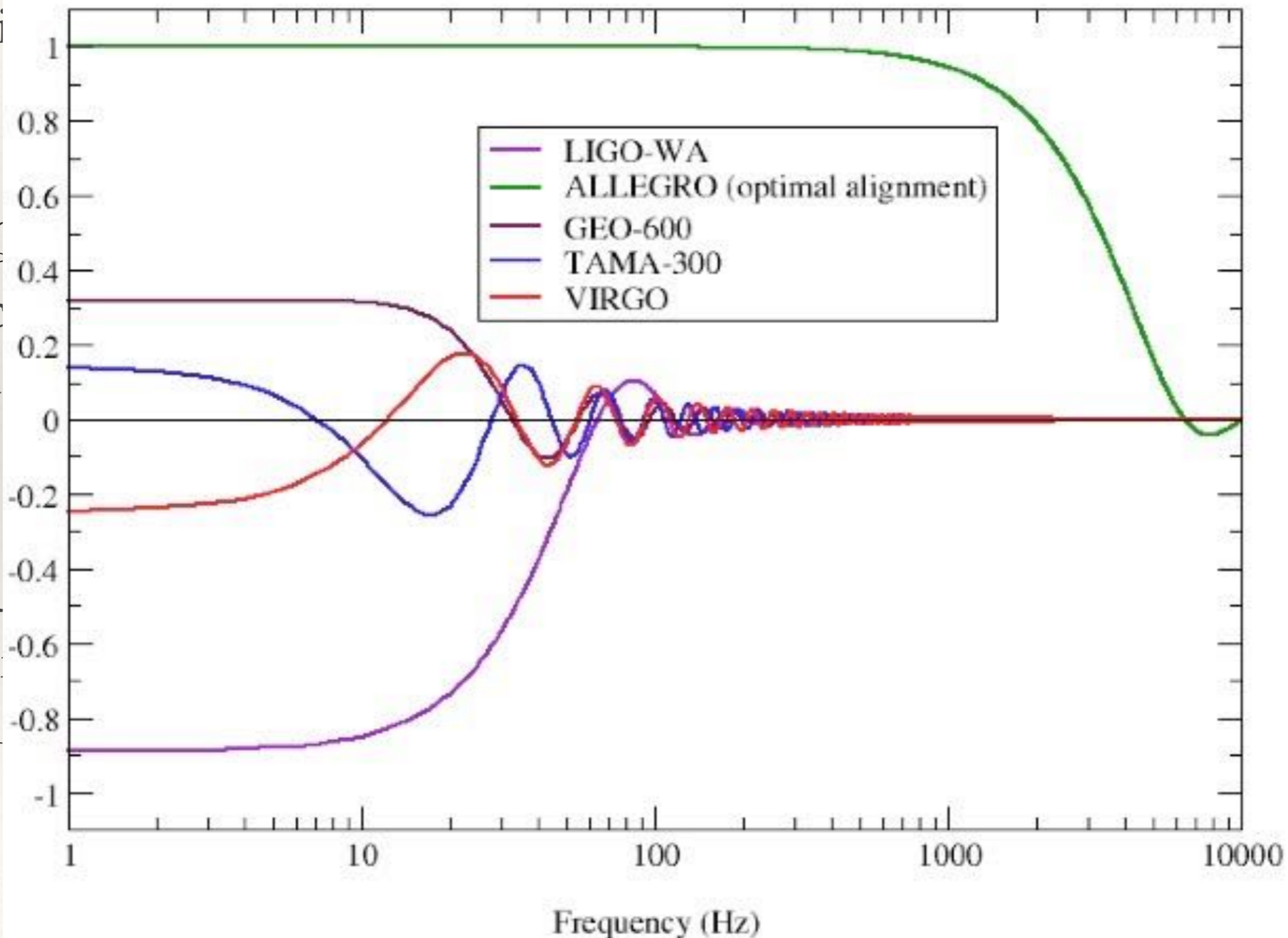
- ❖ Stochastic backgrounds are potentially present in all frequency bands, and could therefore be seen by any of our gravitational wave detectors.
- ❖ The Polarisation of the Cosmic Microwave Background is a direct probe of cosmological gravitational waves.
- ❖ In interferometers, search for an isotropic background using cross-correlation between multiple detectors to identify common noise.

$$\begin{aligned} Y_Q &= \int_0^T dt_1 \int_0^T dt_2 h_1(t_1) Q(t_1 - t_2) h_2(t_2) \\ &= \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \delta_T(f - f') \tilde{h}_1^*(f) Q(f') \tilde{h}_2(f') \end{aligned}$$

# Stochastic Gravitational Wave Fore/Backgrounds

## Overlap Reduction Function

(LIGO-LA and other detectors)



❖ In the  
the Di

❖ and  $C$   
is unc  
correl

❖ The fu  
loss o  
two d

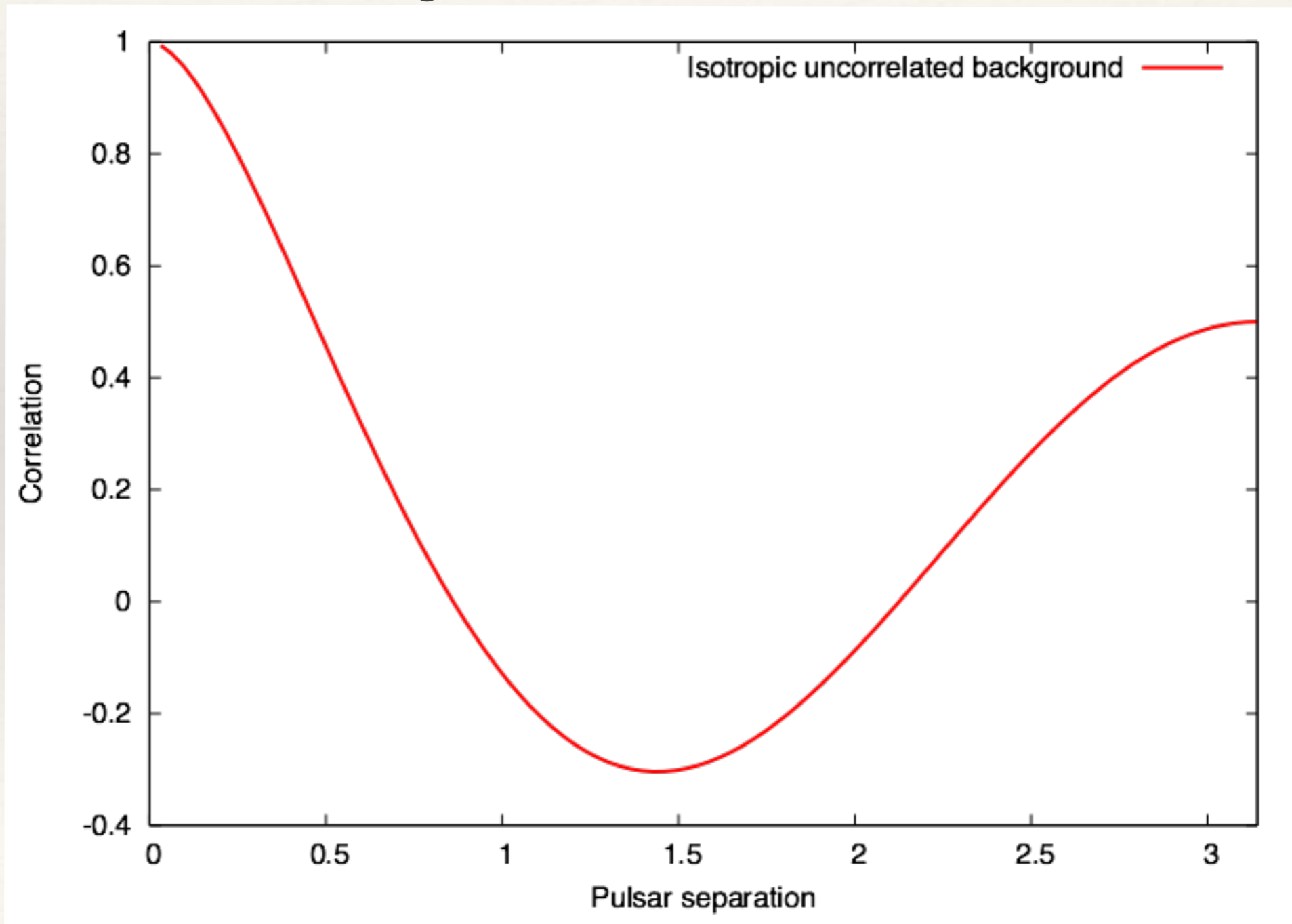
ation to

etectors  
cross-

ares the  
of the

# Stochastic Gravitational Wave Fore/Backgrounds

- ❖ For pulsar timing, the overlap reduction function for an isotropic background is the Hellings and Downs curve.



# Stochastic Gravitational Wave Fore/Backgrounds

- ❖ Uncorrelated anisotropic and correlated backgrounds have different correlation functions.

