Gravitational-Wave Burst Detection: Sources & Methods

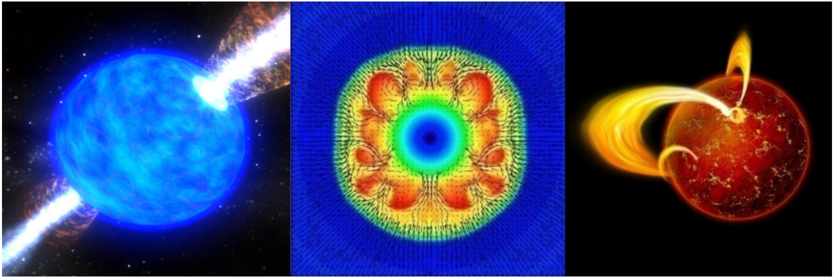
Patrick J. Sutton

Cardiff University

International School of Physics "Enrico Fermi"
Varenna, 2017/07/03-04

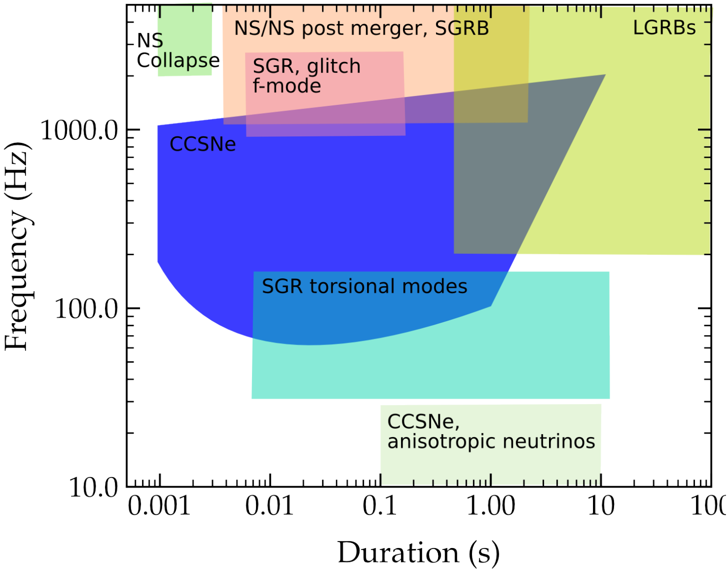
What are GW Bursts?

By convention, “bursts” are transients for which the GW waveform is not known or is too complicated to allow for a templated search.

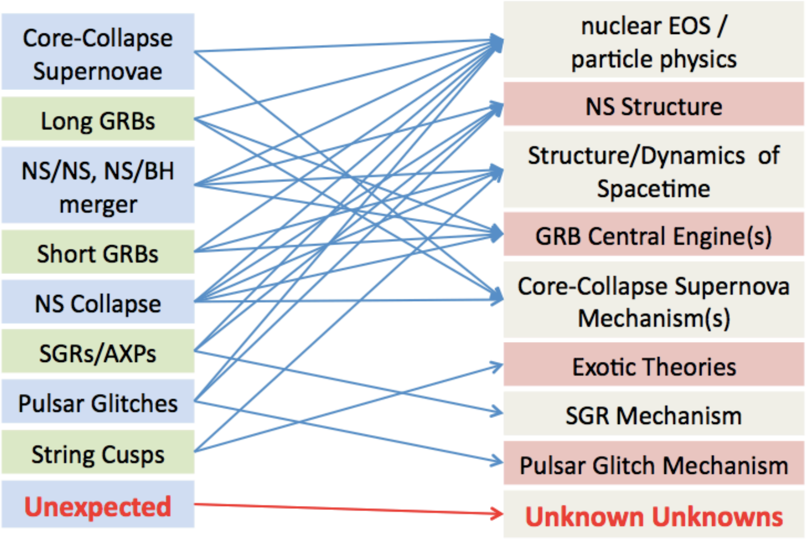


- Examples: supernovae, long gamma-ray bursts, post-merger binary neutron star systems, flaring magnetars, . . .
- LIGO, Virgo: typically < 1 s, but some signals up to O(100) s or possibly more.
- Potentially rich source of (astro-)physics.

Burst Sources (slide by C. Ott)



Burst Science (slide by C. Ott)



Outline of this Lecture

We will review the basic concepts on how we can detect and interpret GW signals *without* prior knowledge of the waveform.

- Standard matched filtering is a special case of the techniques used for bursts.

Topics

- detection
- waveform reconstruction
- glitch rejection
- sky position estimation

Further reading (*i.e.*, not covered in this talk!):

For a brief overview of potential sources, see:
http://bcc.impan.pl/13Gravitational/uploads/Sutton_sources.pdf

Excess Power Detection

Key Concept: Excess Power

Assumptions:

- Assume detector noise n is uncorrelated with GW signal.
- Assume detector responds linearly to GW signal.

Detector output:

$$d(t + \Delta t(\Omega)) = n(t + \Delta t(\Omega)) + F^{+} h_{+}(t) + F^{\times} h_{\times}(t)$$

Notation

- Assume the data have been time shifted to compensate for time-of-flight between detectors: $\Delta t(\Omega) = \frac{1}{c}(\mathbf{r}_0 - \mathbf{r}) \cdot \Omega$
- Treat each timeseries as a vector: $n(t) \to \mathbf{n}$, $h_{+\times}(t) \to \mathbf{h}_{+\times}$.

Then (suppressing $F^{+,\times}$):

$$\mathbf{d} \simeq \mathbf{n} + \mathbf{h}$$

Key Concept: Excess Power

The assumption that **n** and **h** are uncorrelated means

$$\langle \mathbf{n}^T \mathbf{h} \rangle = 0$$

where $\langle \ldots \rangle$ denotes the average over noise realisations.

Why is this useful? Because signal and noise add in quadrature in the energy:

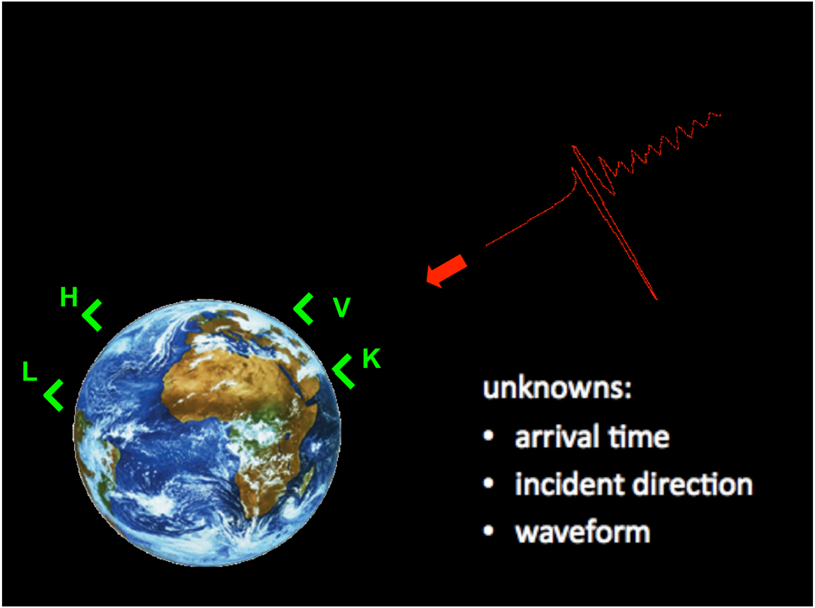
$$\begin{aligned} \text{energy} &= \mathbf{d}^T \mathbf{d} \\ \langle \text{energy} \rangle &= \langle \mathbf{n}^T \mathbf{n} \rangle + \underbrace{\langle \mathbf{h}^T \mathbf{h} \rangle}_{\textit{positive-definite}} + \underbrace{2\langle \mathbf{n}^T \mathbf{h} \rangle}_{\textit{zero}} > \langle \mathbf{n}^T \mathbf{n} \rangle \end{aligned}$$

Key Idea:

Any non-zero GW will cause *excess power* in the data regardless of its waveform. (But so will noise glitches!) [Anderson et al. (2001)]

Multi-Detector Coherent Analysis

Physical Point of View



Notation for discrete data

We'll work in the frequency domain unless specified otherwise.

measured data: d_{strain}

noise: n_{strain}

power spectrum: S, defined in k^{th} frequency bin by

$$\langle n^*_{\mathrm{strain}}[k] n_{\mathrm{strain}}[k'] \rangle = \frac{N}{2} \delta_{k,k'} S[k] \,.$$

whitened data: $d \equiv d_{\mathrm{strain}} / \sqrt{(N/2)S}$, and similarly for n

With these conventions the whitened noise has the simple normalisation

$$\langle n^*[k] n[k'] \rangle = \delta_{k,k'} \,.$$

Single-Sample, Known Sky Position

- Assume the sky position Ω of the GW is known, and time-shift the data streams appropriately.
- Work in the frequency domain. Consider one sample of whitened data from each of D detectors ($\sigma \equiv \sqrt{(N/2)S}$):

$$\underbrace{\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_D \end{bmatrix}}_{measured} = \underbrace{\begin{bmatrix} F_1^+(\Omega)/\sigma_1 & F_1^\times(\Omega)/\sigma_1 \\ F_2^+(\Omega)/\sigma_2 & F_2^\times(\Omega)/\sigma_2 \\ \vdots & \vdots \\ F_D^+(\Omega)/\sigma_D & F_D^\times(\Omega)/\sigma_D \end{bmatrix}}_{known\ (for\ each\ \Omega)} \underbrace{\begin{bmatrix} h_+ \\ h_\times \end{bmatrix}}_{unknown} + \underbrace{\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_D \end{bmatrix}}_{noise}$$

- Burst search: treat $h_+, h_\times$ as free parameters to be fit to the data.
  - linear
  - over-constrained (allows consistency tests)
  - sky-position dependent
  - repeat for each time-frequency data sample

Detection: the likelihood ratio

- Switch to matrix notation, where elements correspond to different detectors.

$$
\begin{aligned}
\mathbf{d} &= [\,\mathbf{F}_{+}\ \mathbf{F}_{\times}\,]\,\mathbf{h} + \mathbf{n} \\
&= \mathbf{F}\,\mathbf{h} + \mathbf{n}
\end{aligned}
$$

- Detection question: Is $\mathbf{d}$ inconsistent with $\mathbf{h} = 0$? Construct the likelihood ratio:

$$
L \equiv 2 \log \frac{p(\mathbf{d}|\mathbf{h})}{p(\mathbf{d}|0)}
$$

- Importance: Can show thresholding on L is the optimal strategy for detecting known $\mathbf{h}$.
  - Highest detection probability for fixed false alarm rate ("Neyman-Pearson" criterion).

Gaussian noise case

- Need to known statistical properties of **n**.
  - Assume noise is not correlated between detectors (realistic).
  - Assume noise is Gaussian (unrealistic).
- So probability of measuring data **d** in the absence of a GW ($\mathbf{h} = 0$) is

$$p(\mathbf{d}|0) \simeq \exp\left\{-\frac{1}{2}\mathbf{d}^T\mathbf{d}\right\}$$

- And probability of measuring data **d** given a GW **h** is

$$p(\mathbf{d}|\mathbf{h}) \simeq \exp\left\{-\frac{1}{2}(\mathbf{d} - \mathbf{F}\mathbf{h})^T(d - \mathbf{F}\mathbf{h})\right\}$$

Matched filter

- Likelihood ratio becomes

$$L = \underbrace{\mathbf{d}^T\mathbf{F}\mathbf{h} + (\mathbf{F}\mathbf{h})^T\mathbf{d}}_{\text{matched filter}} - \underbrace{(\mathbf{F}\mathbf{h})^T(\mathbf{F}\mathbf{h})}_{\substack{\text{independent of data} \\ \text{(ignored in matched filtering)}}}$$

- Typical matched filtering search procedure:
  - Construct template bank spanning all $\mathbf{h}$.
  - Matched filter with each template separately for each detector.
  - “Detection” if filter SNR > threshold in each detector.

Life without templates: Waveform Reconstruction

- Burst search: Define best-fit waveform $\hat{\mathbf{h}}$ as that which maximizes L:

$$0 = \left.\frac{\partial L}{\partial h}\right|_{\mathbf{h}=\hat{\mathbf{h}}}$$

- Simple linear problem. Solution for any $\mathbf{d}$ is:

$$\hat{\mathbf{h}} = \mathbf{F}_{\mathrm{MP}}^{-1}\mathbf{d} = (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\mathbf{d}$$

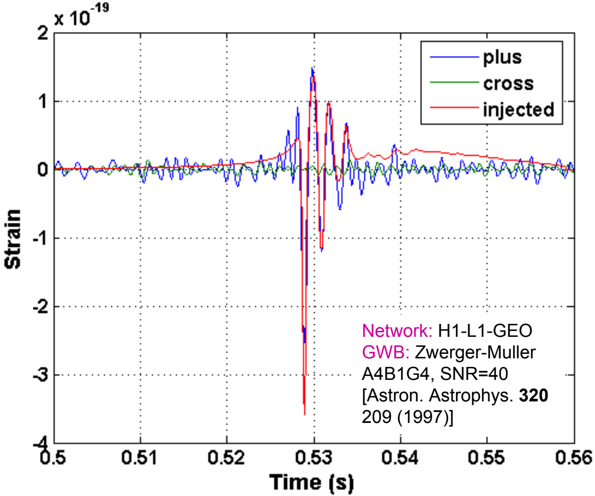
where $\mathbf{F}_{\mathrm{MP}}^{-1}$ is the Moore-Penrose inverse.

Caveats:

- Bayesians: We've assumed a flat (improper) prior on $\mathbf{h}$.
- $(\mathbf{F}^T\mathbf{F})$ tends to be singular, requiring regularisation.
- Number of free parameters $\simeq$ number of data points ...

More on these later.

Example: Supernova injection

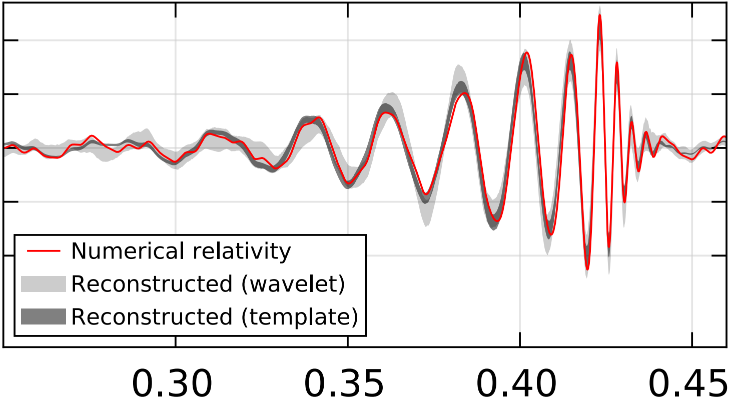


Recovered signal (blue) is a noisy, band-passed version of injected GWB signal (red)

Injected GWB signal has $h_x = 0$.

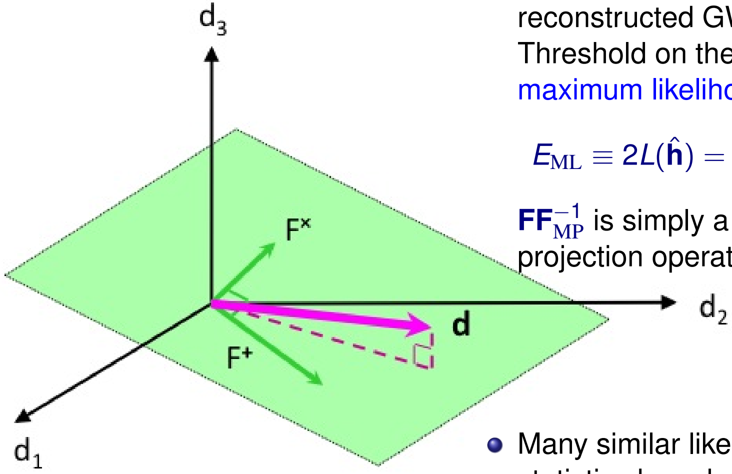
Recovered h_x (green) is just noise.

B. P. Abbott et al., PRL 116, 061102 (2016)



- Fit by `BayesWave` pipeline, which includes priors on amplitude, clustering, etc. [Cornish & Littenberg (2015)].
- Match with best-fit template: 94% (SNR=24).

Maximum likelihood detection statistic

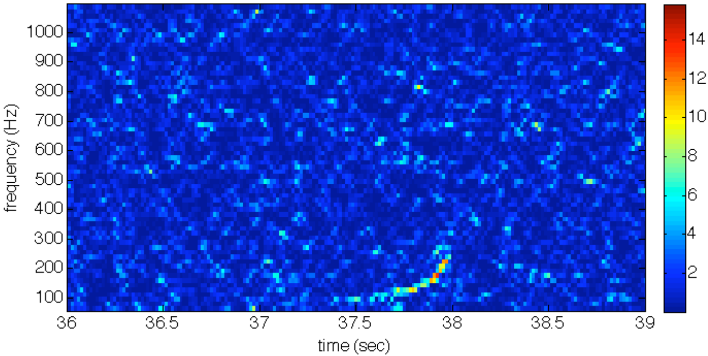


- How significant is the reconstructed GW? Threshold on the maximum likelihood:

$$E_{\mathrm{ML}} \equiv 2L(\hat{\mathbf{h}}) = \mathbf{d}^T \mathbf{F}\mathbf{F}_{\mathrm{MP}}^{-1}\mathbf{d}$$

$\mathbf{F}\mathbf{F}_{\mathrm{MP}}^{-1}$ is simply a 2D projection operator.

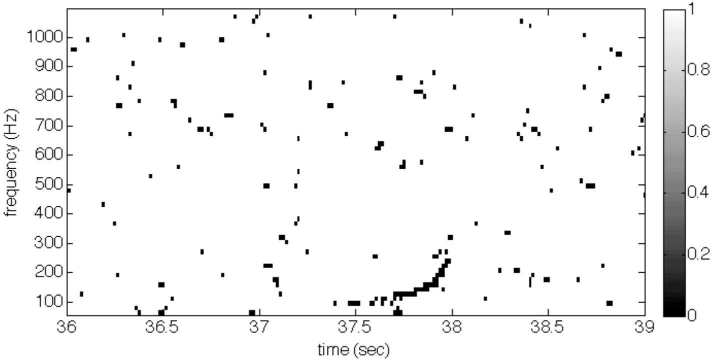
- Many similar likelihood statistics have been proposed and used.



A generic GW burst will be spread over multiple data samples.
Standard approach: threshold and cluster. E.g.:

- Zero-out most (e.g. 99%) of the lowest-value pixels.
- Group together remaining pixels which share an edge or corner (next-nearest neighbors).
- Sum energy (likelihood) over all pixels in a given cluster.

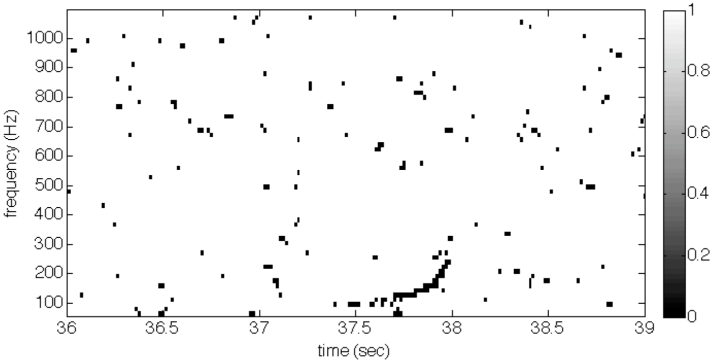
Time-Frequency Analysis: Clustering



A generic GW burst will be spread over multiple data samples. Standard approach: threshold and cluster. E.g.:

- Zero-out most (e.g. 99%) of the lowest-value pixels.
- Group together remaining pixels which share an edge or corner (next-nearest neighbors).
- Sum energy (likelihood) over all pixels in a given cluster.

Time-Frequency Analysis: Clustering



Bayesian Interpretation

Signal prior: colored Gaussian random process with unimodal time-frequency shape, uniform in size and shape [Was, PhD Thesis, 2011].

Sensitivity: Excess power vs. matched filter

- In Gaussian noise with a fixed time-frequency region of N pixels the maximum likelihood is non-central chi-squared distributed [Anderson et al. 2001, Sutton et al. 2010]:

$$2E_{\mathrm{ML}} \sim \chi^2_{4N}(\rho^2)$$

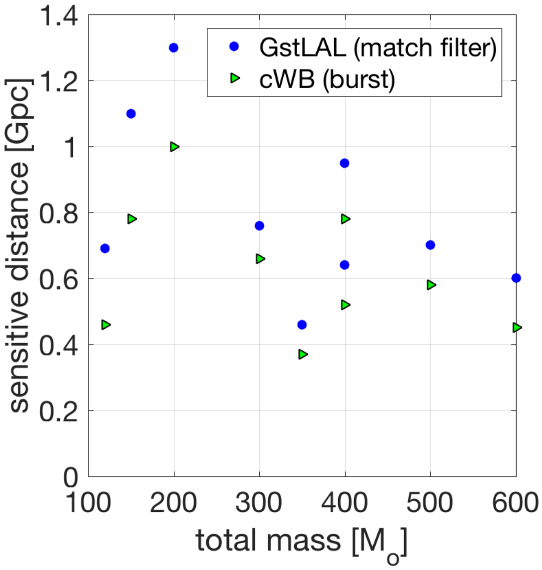
$$\begin{aligned}\mathrm{mean}\,[2E_{\mathrm{ML}}] &= \rho^2 + 4N \\ \mathrm{st.dev.}\,[2E_{\mathrm{ML}}] &= (8N)^{\frac{1}{2}}\end{aligned}$$

Here ρ^2 is the expected SNR of a matched filter:

$$\rho^2 = 2\sum_{pixels}(\mathbf{Fh})^T(\mathbf{Fh}) .$$

- Matched filter ($N \to 2$) more sensitive by amplitude factor $\propto N^{1/4}$ (volume factor $\propto N^{3/4}$).
  - $N \sim 100$ for low-mass binary mergers.
  - $N \sim 1$ for high-mass binary mergers.

Example: IMBBH Searches



Search for BBH systems in LIGO O1 data [Abbott et al., 1704.04628].

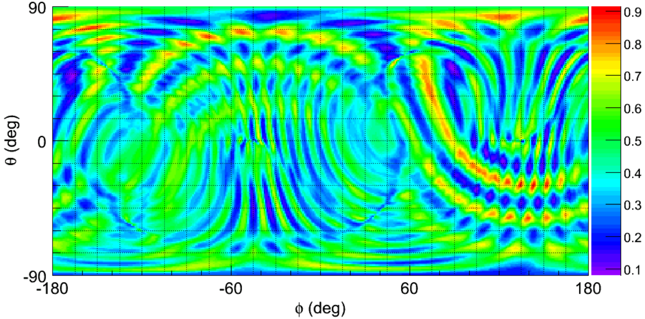
Burst sensitive range (sensitive volume) typically 80% (50%) that of matched filter.

Sky Localisation Accuracy

Unknown Sky Position

Typically we don't know the incident sky direction Ω of the GW *a priori*.

- Coherent analysis depends on Ω through antenna responses $F^{+}(\Omega)$, $F^{\times}(\Omega)$ and through arrival time delay between detectors.
- Standard approach: Repeat analysis over a grid of Ω covering the entire sky. Gives likelihood map $E(\Omega)$.

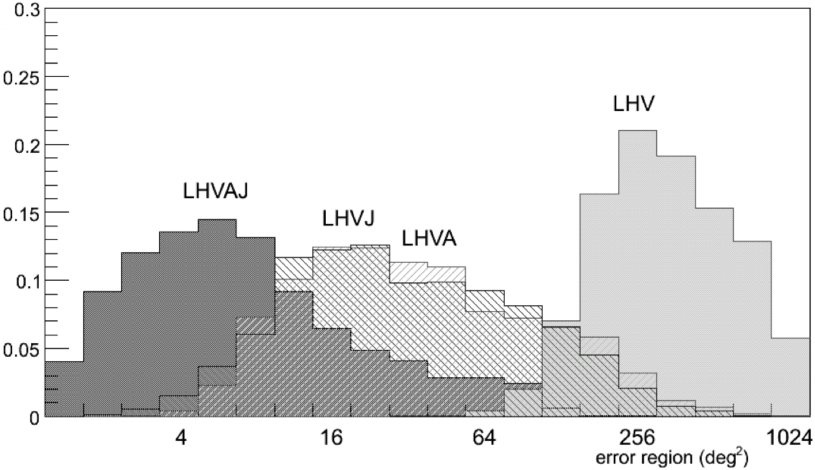


Ex: coherent WaveBurst sky map for SG235Q9 GWB at $(\theta, \phi) = (-30^{\circ}, 144^{\circ})$.

Sky Localisation Accuracy

Klimenko et al., 1101.5408 predictions for 90% containment regions for advanced detector networks.

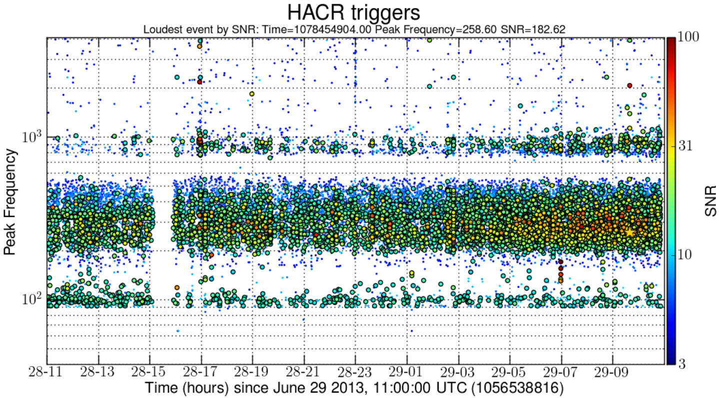
- test signals: mix of white-noise bursts and sine-Gaussians



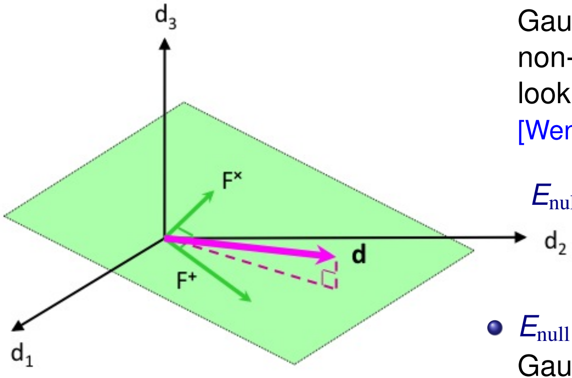
Background Glitch Rejection

Non-Stationary Background

Real detector noise isn't Gaussian or stationary.



The "Null Stream"



- Residual data with signal $\hat{h}$ removed should be Gaussian. Can reject non-Gaussian glitches by looking at the *null energy* [Wen and Schutz (2005)]:

$$E_{\text{null}} \equiv \mathbf{d}^T \underbrace{\left(\mathbf{I} - \mathbf{F}\mathbf{F}_{\text{MP}}^{-1}\right)}_{\textit{projection}} \mathbf{d}$$

- $E_{\text{null}} \sim \chi^2_{2N(D-2)}$ in Gaussian noise.
- Analogous to χ^2 test for inspirals [Allen (2005)].

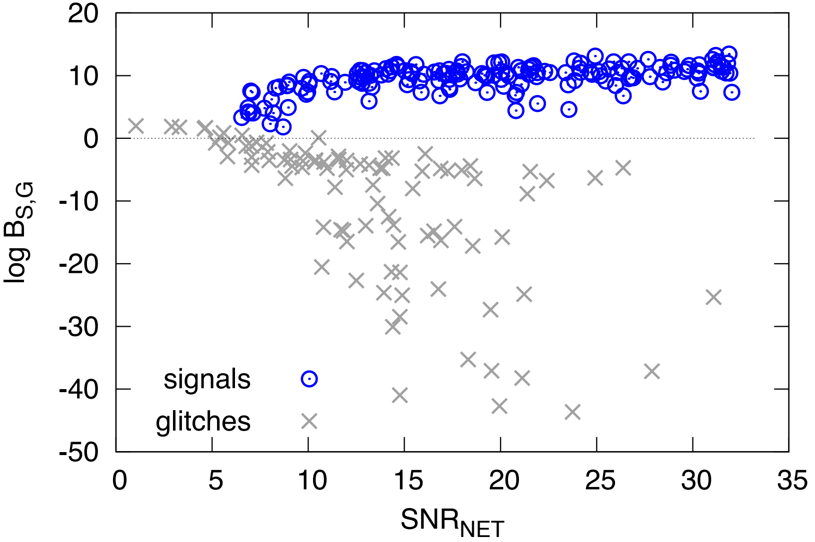
Advanced Glitch Rejection

Instead of assuming Gaussian noise, re-derive optimal statistic using a quantitative model for glitches. E.g.:

- Locally Optimum Detectors [Creighton 1999, Allen et al. 2002].
- Maximum likelihood analysis using Middleton glitch model [Principe & Pinto 2009, 2017].
- Bayesian model selection with a similar glitch model [Littenberg et al. (2016)]

Example: Bayes Factors for Signal vs. Gltich

[Littenberg et al. (2016)]



Summary

- We have powerful techniques for detecting and studying GW bursts that don't rely on detailed knowledge of the waveform.
  - detection
  - background rejection
  - waveform reconstruction
  - source localization on the sky
- Full exploitation relies crucially on having a network of detectors at several sites (at least 3).
- More work is still needed! In particular:
  - Searches are still limited by non-Gaussian backgrounds.
  - Searches (esp. Bayesian analyses) are computationally expensive.
  - How will we interpret eventual detections without reference models?

Reading List

1. Gursel and Tinto (1989) Phys. Rev. D 40 3884 [waveform reconstruction, sky localisation]
2. Flanagan and Hughes (1998) Phys. Rev. D 57 4566 [maximum likelihood formulation]
3. Creighton (1999) Phys. Rev. D60 021101 [glitch robustness]
4. Anderson et al. (2001) Phys. Rev. D 63 042003 [excess power]
5. Sylvestre (2002) Phys. Rev. D 66 102004 [clustering]
6. Allen et al. (2002) Phys. Rev. D 65, 122002 [glitch robustness]
7. Klimenko et al. (2005) Phys. Rev. D 72 122002 [regularisation]
8. Wen and Schutz (2005) Class. Quant. Grav. 22 S1321 [glitch rejection]
9. Rakhmanov (2006) Class. Quant. Grav. 23 S673 [regularisation]
10. Chatterji et al. (2006) Phys. Rev. D 74 082005 [glitch rejection]

Reading List (continued)

11. Klimenko et al. (2008) Class. Quant. Grav. 25:114029 [coherent WaveBurst]
12. Searle et al. (2008) Class. Quant. Grav. 25 114038 [Bayesian formulation]
13. Searle et al. (2009) Class. Quant. Grav. 26 155017 [Bayesian formulation]
14. Principe and Pinto (2009) Class. Quant. Grav. 26 045003 [glitch robustness]
15. Sutton et al. (2010) New J Phys. 12 053034 [X-Pipeline]
16. Klimenko et al. (2011) Phys.Rev.D83:102001 [sky localisation]
17. Cornish and Littenberg (2015) Class. Quant. Grav. 32, 135012 [BayesWave]
18. Littenberg et al. (2016) Phys. Rev. D94 044050 [glitch robustness]
19. Principe and Pinto (2017) Phys. Rev. D95 082006 [glitch robustness]