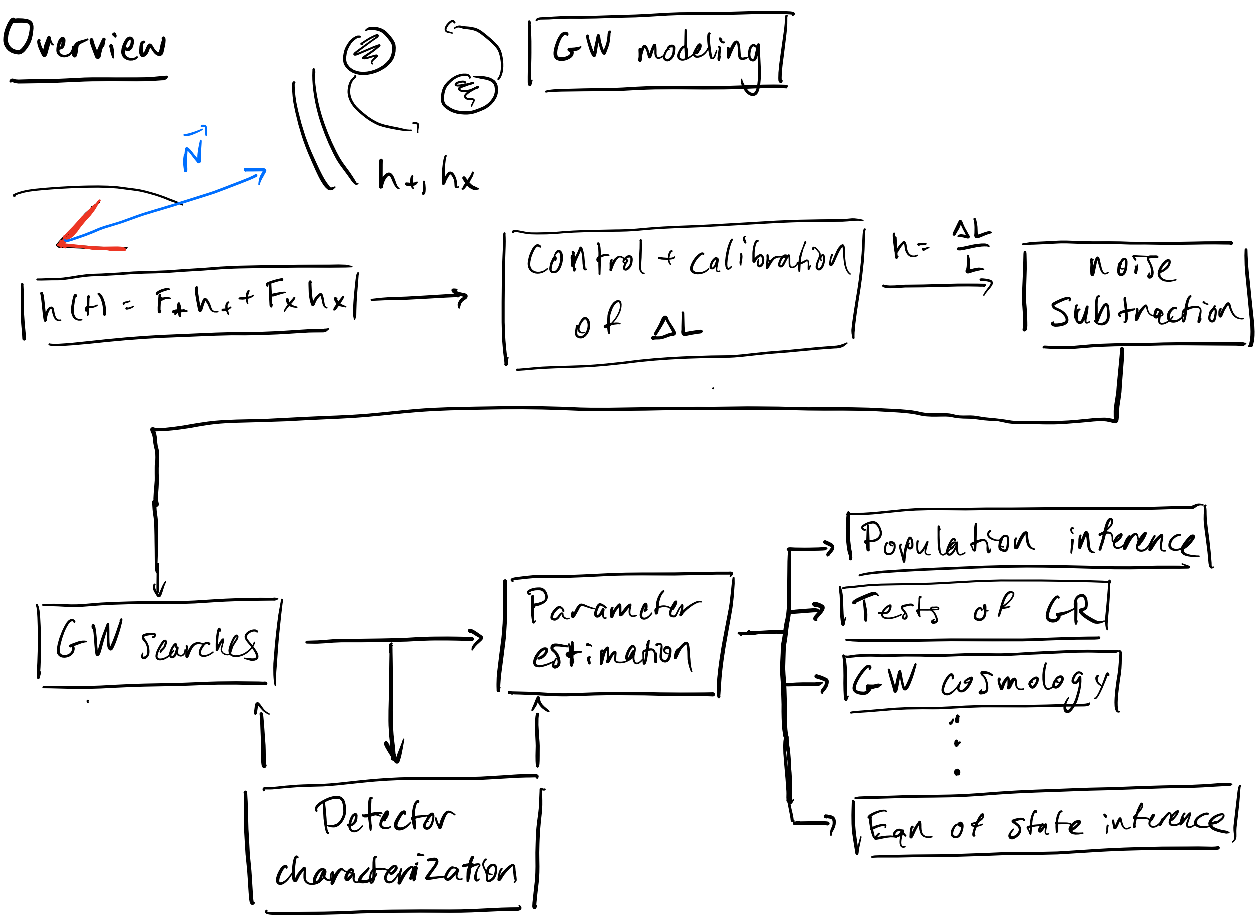
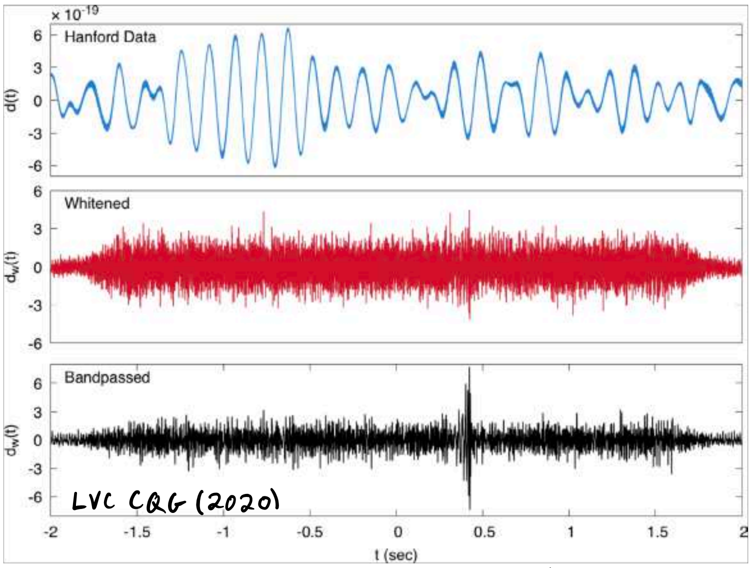
Introduction to GW data analysis

Aaron Zimmerman

Overview



Part 1: Noise and the "easy" search challenge



• Conventions: $\hat{h}(f) = \int_{-\infty}^{\infty} h(t)\, e^{-2\pi i f t}\, dt$

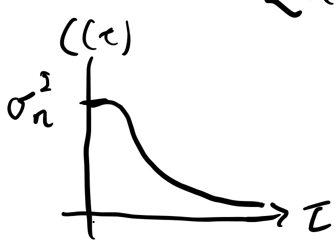
$h(t) = \int_{-\infty}^{\infty} \hat{h}(f)\, e^{+2\pi i f t}\, df$

h real $\rightarrow \hat{h}^*(f) = \hat{h}(-f)$

• $d(t) = h(t) + n(t)$

* Stationary + ergodic + zero mean $\langle n \rangle = 0$

* $C(\tau) = \langle n(t)\, n(t+\tau) \rangle \qquad C(0) = \sigma_n^2$



* $\langle \hat{n}(f)\, \hat{n}^*(f') \rangle = \frac{1}{2} S_n(f)\, \delta(f'-f)$

$\uparrow$ one-sided PSD

* $[n] = 1 \quad [\hat{n}] = T = \frac{1}{Hz} = [S_n]$

* RMS fluctuations in n @ f over T

  resolution $\Delta f = \frac{1}{T}$

  $n_{rms} = \sqrt{S_n(f)\,\Delta f}$

* Stretch $T \to$ DFT $\to$ repeat + average Welch

- Linear Filter

$$z(t) = \int K(\tau)\, d(t+\tau)\, d\tau \qquad z(0) = z$$

* Maximize

$$SNR = \frac{\langle z \rangle_{signal}}{\left[\langle z^2 \rangle_{noise}\right]^{1/2}}$$

* $\hat{K}(f) = \frac{2\,\hat{g}(f)}{S_n(f)}$

$$\langle z \rangle_{signal} = \int_{-\infty}^{\infty} df\; K^*(f)\,(\hat{h} + \overset{0}{\langle \hat{n} \rangle})$$

$$= 2\int_{-\infty}^{\infty} \frac{g^*\hat{h}}{S_n}\, df$$

$$= 2\int_{0}^{\infty} \frac{(g^*h + g h^*)}{S_n}\, df$$

$$= 4\,\mathrm{Re}\int_0^\infty \frac{g^* h}{S_n}\,df = \langle g | h \rangle$$

- $$\langle Z^2 \rangle_{noise} = \iint df\,df'\, K^*(f)\, K^*(-f')\, \underbrace{\langle n(f)\, n(-f) \rangle}_{\frac{1}{2} S_n \delta(f-f')}$$

$$= \frac{2\cancel{4}}{\cancel{2}}\int_{-\infty}^{\infty} \frac{g^* g}{S_n^{\cancel{2}}}\,\cancel{S_n}\,df$$

$$= 4\int_0^\infty \frac{|g|^2}{S_n}\,df = \langle g | g \rangle$$

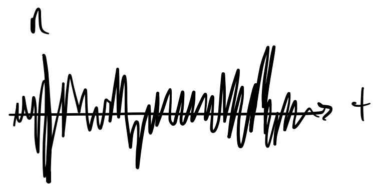
$$SNR = \frac{\langle g | h \rangle}{\sqrt{\langle g | g \rangle}} = \langle \hat{g} | h \rangle$$

$$\hat{g} = c\tilde{h}$$ Matched filter

$$\hat{K} = \frac{2\tilde{g}(f)}{S_n}$$

Example: White noise

$n(t_i) \sim N(0, \sigma)$ $\qquad \langle n(t)\, n(t') \rangle = \sigma^2 \delta(t - t')$

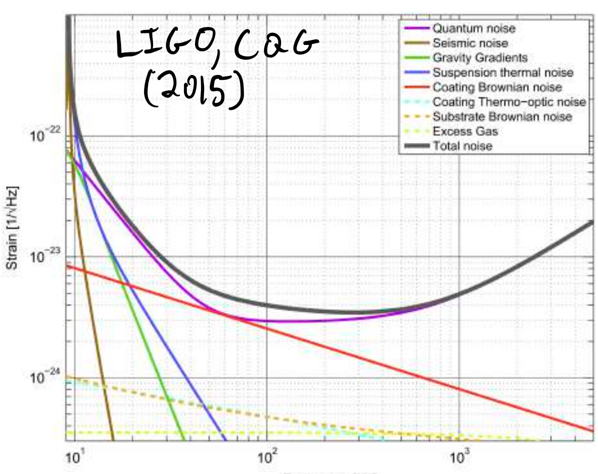


Example: Colored Gaussian noise

$\hat{n}(f_i) \sim N(0, \sigma_i)$ $\qquad \sigma_i^2 = \frac{1}{2} \frac{S_n(f)}{\Delta f}$

$$p(\hat{n}(f)) \propto \prod_i \exp\left[-\frac{1}{2}\left(\frac{2|\hat{n}|^2 \Delta f}{S_n}\right)\right] = \exp\left[-\frac{1}{2}\langle \hat{n} | \hat{n} \rangle\right]$$

$ASD \approx \sqrt{PSD}$

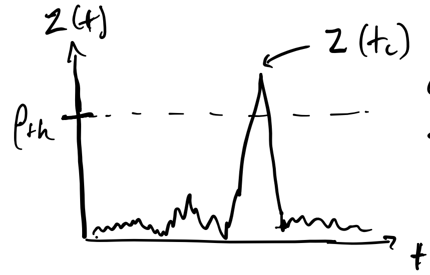


Part 2: Real challenge

- Don't know h, have to scan over parameter space

- Max over time: use

$$z(t) = 2\int_{-\infty}^{\infty} \frac{\hat{h}^* \hat{d}}{S_n} e^{2\pi i f t}\, df$$

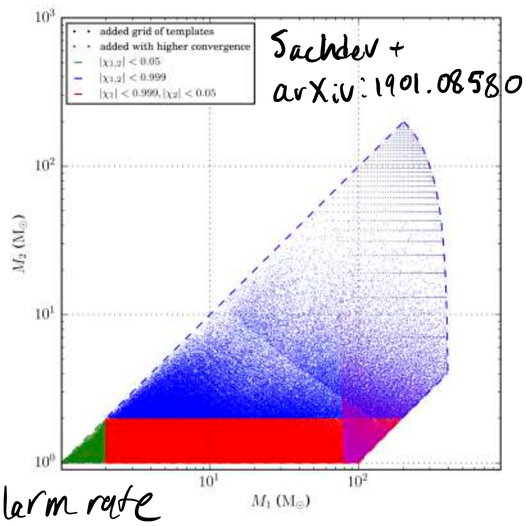


- Max over phase: use a cosine-like and sine-like template

$$z = \langle \hat{h}_s + i\hat{h}_c \mid d \rangle, \qquad |z(t_c)| \text{ maximizes over unknown phase}$$

- Max over intrinsic parameters: use a bank of templates

  * current banks tile in $m_1, m_2, \chi_{1z}, \chi_{2z}$

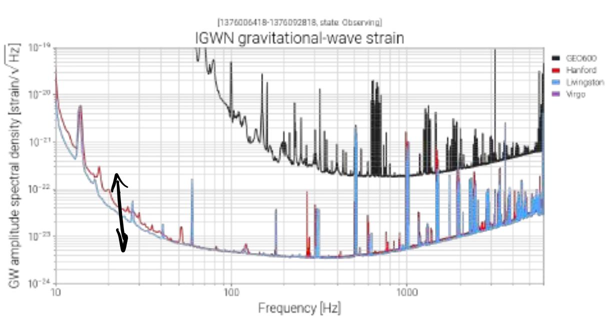


Sachdev + arXiv:1901.08580

- Given Gaussian noise, can compute or fit $p(|z| \mid \text{noise})$ and get False alarm probability / false alarm rate of a given candidate

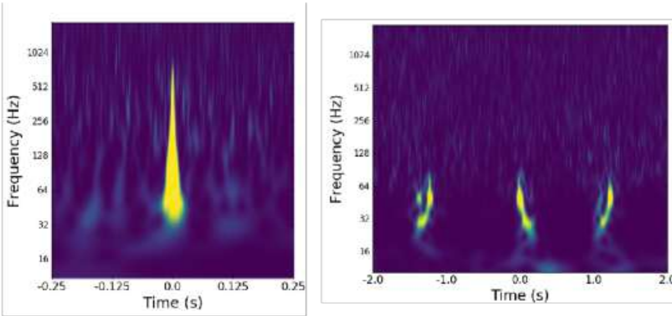
$$MM = 1 - \underbrace{\max_{\Delta t, \Delta\phi_c} \langle \hat{h}_1 \mid \hat{h}_2 \rangle}_{\text{match}}$$

- Reality check:



gw-openscience.org

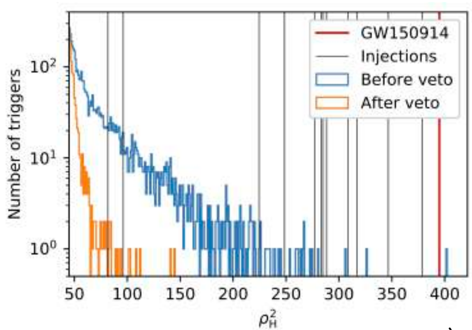
- Noise not stationary
  $\Rightarrow$ update $S_n(f)$ regularly



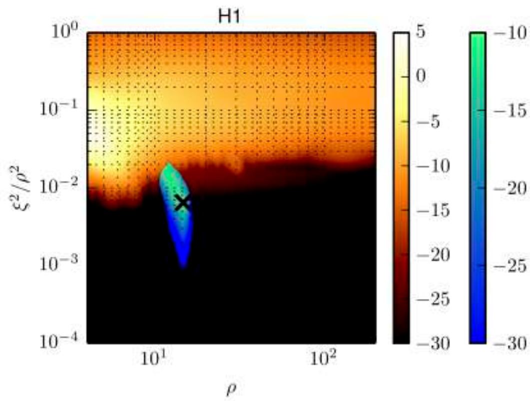
Zevin + CQG (2017)

- Not even slowly varying Gaussian: glitches

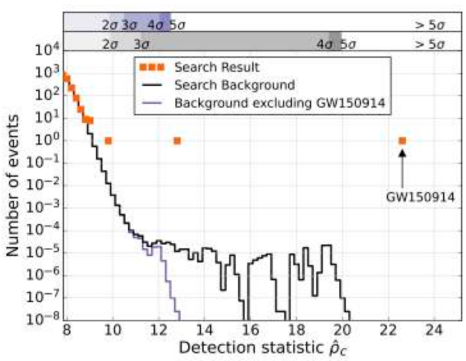
=> modify detection statistic



Venumadhav + PRD (2019)



Messick + PRD (2016)



Simple example:

$$\hat{\rho}_c = \rho \min\left[1, \left(\frac{1}{2} + \frac{\chi_r^6}{2}\right)^{1/6}\right]$$

penalty for bad agreement

Part 3: Parameter estimation

$d - h = n \qquad \theta = \{m_1, m_2, \ldots\}$

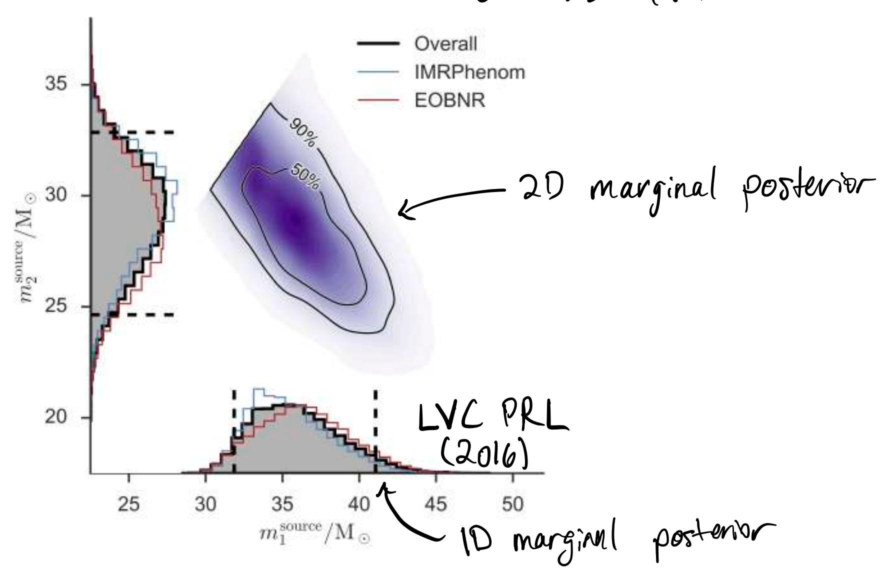
$\mathcal{L}(d|\theta) \qquad d - h(\theta) = n$

$\shortparallel$

$p(\hat{n}|h(\theta)) \propto \exp\left[-\frac{1}{2}\langle n|n\rangle\right] = \exp\left[-\frac{1}{2}\langle d-h(\theta)|d-h(\theta)\rangle\right]$

$p(\theta|d)$ ← measurement of θ

GW150914



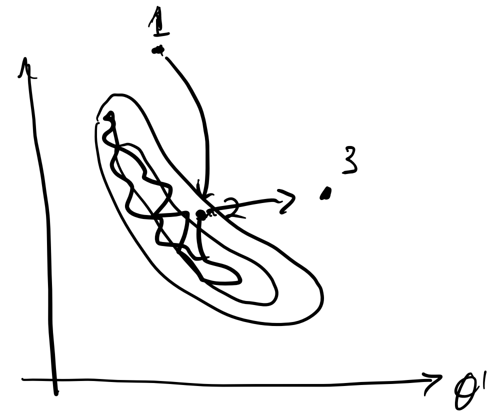
$$p(\theta|d) = \frac{\mathcal{L}(d|\theta)\, p(\theta)}{Z}$$

* samples in practice

$$\theta_i \sim p(\theta|d)$$

$$\langle m_1\rangle = \int m_1\, p(\theta|d)\, d\theta \approx \frac{1}{N}\sum_{i=1}^{N} m_{1,i}$$

Example: Metropolis-Hastings



(1) θ_1 propose $\theta_2 \sim Q(\theta_2 | \theta_1)$

$$Q(\theta_2|\theta_1) = Q(\theta_1|\theta_2)$$

(2) compute $r = \dfrac{p(\theta_2|d)}{p(\theta_1|d)}$

(3) accept the jump w/ prob $\min[1, r]$

* Reality: compute correlation τ of the chain
  * decimate