Fundamentally we have

A) Experiment that gives us $d_{raw}(t_n)$

B) GR + math tricks that give us a waveform model over a parameter space $\vec{\theta} = (m_1, m_2, \vec{\chi}_1, \vec{\chi}_2, \iota, \varphi, t_c, D_L, \alpha, \delta, \psi, \lambda_1, \lambda_2, e)$

From these we get our three fundamental pieces of data analysis:

1) d_n = noise-subtracted data series from Wiener filter & glitch masking on $d_{raw}(t_n)$
→ depends on noise model $n(t)$

2) $P(f_k) = \mathcal{F}\left[\langle d(t)\, d(t+\tau)\rangle\right] \equiv$ "Stationary (assumed) Power Spectral Density"
= F.T. of autocovariance of d_n samples
→ estimated using, e.g., Welch's Method

3) $h(\vec{\theta})$ = modeled signal produced at the detector due to GW from merger of compact object binary w/ parameters $\vec{\theta}$

AND with our "overlap" (inner product) defined (in continuum limit) by $\langle a|b\rangle = 4\,\mathrm{Re}\left[\int_0^\infty df\, \frac{\tilde{a}(f)\, \tilde{b}^*(f)}{P(f)}\right]$
we want to compute the SNR squared:

$$\rho^2(\vec{\theta}) = \frac{\langle d|h(\vec{\theta})\rangle^2}{\langle h(\vec{\theta})|h(\vec{\theta})\rangle}$$

"WHITENED" DATA & TEMPLATE

to make overlap math easier, we define a WHITENED series by [now considering discrete time grid w/ $\Delta t = t_{n+1} - t_n$]

$$\tilde{X}_w(f) \equiv \sqrt{\frac{2\Delta t}{P(f)}}\,\tilde{X}(f) \quad \& \quad X_w(t) = \mathcal{F}^{-1}\left[\tilde{X}_w(f)\right]$$

so that

$$\langle d \,|\, h(\vec{\theta})\rangle = \frac{1}{N}\sum_{f_k} \tilde{d}_w(f_k)\,\tilde{h}^*_w(f_k;\vec{\theta})$$

$$= \sum_{\substack{\text{time samples}\\ t_n}} d_w(t_n)\,h_w(t_n;\vec{\theta})$$

Two important tricks/derivations:

$\boxed{1}$ separate the "coalescence time" parameter t_0 by recognizing that in frequency space we have

$$\tilde{h}(f;\vec{\theta}=(\vec{\theta}_0,t_0)) = \mathcal{F}\left[h(t-t_0;\vec{\theta}_0)\right]$$

$$= \tilde{h}(f;\vec{\theta}=(\vec{\theta}_0,0))\,e^{-2\pi i f t_0} \Longrightarrow$$

$$\langle d / h(t-t_0;\vec{\theta}_0)\rangle(t_0) = \frac{1}{N}\sum_{f_k}\tilde{d}_w(f_k)\,\tilde{h}^*_w(f_k;\vec{\theta}_0,0)\,e^{2\pi i f t_0}$$

$$= \mathcal{F}^{-1}\left[\tilde{d}_w\,\tilde{h}^*_w(\vec{\theta}_0,0)\right](t_0)$$

$\boxed{2}$ log likelihood is $\ln\mathcal{L}(\vec{\theta}) = -\langle d|d\rangle + (d/h) - \frac{1}{2}(h/h)$

$\rightarrow$ write $h = \dfrac{h_1}{D_L}$ where h_1 is $h(\ldots, D_L = 1\,\text{Mpc})$ & D_L is luminosity distance

$$\Rightarrow \frac{\partial \ln\mathcal{L}}{\partial D_L} = \frac{\partial}{\partial D_L}\left(\frac{\langle d|h_1\rangle}{D_L} - \frac{1}{2}\frac{\langle h_1|h_1\rangle}{D_L^2}\right) = 0$$

$$\Rightarrow D_L^{max} = \frac{\langle h_1/h_1\rangle}{\langle d|h_1\rangle} \Rightarrow \underset{\substack{\uparrow\\ \text{drop const. } \langle d/d\rangle}}{\Delta\ln\mathcal{L}}(D_L^{max}) = \frac{1}{2}\frac{\langle d|h_1\rangle^2}{\langle h_1|h_1\rangle} = \frac{\rho^2}{2}$$

In fact, we use ρ^2 as a detection statistic precisely because it is a monotonic function of the likelihood $\mathcal{L}(\vec{\theta};d) \equiv P(d|\vec{\theta}) = P(\vec{\theta}|d)\,\dfrac{\Pi(\vec{\theta})}{\Pi(d)}$

which is what the Neyman-Pearson lemma tells us will be the optimal statistic [max "power" $\Leftrightarrow$ min False Negatives for fixed False Positive Rate]

Note that it is common to denote by $\mathcal{L}$ the rescaling "$\mathcal{L}$" $= Z\,P(d|\vec{\theta})$ for any positive normalization constant Z ($\nabla_{\vec{\theta}} Z = \vec{0}$ & $Z > 0$)

since we typically only compute likelihood ratios $\leftrightarrow$ log likelihood differences