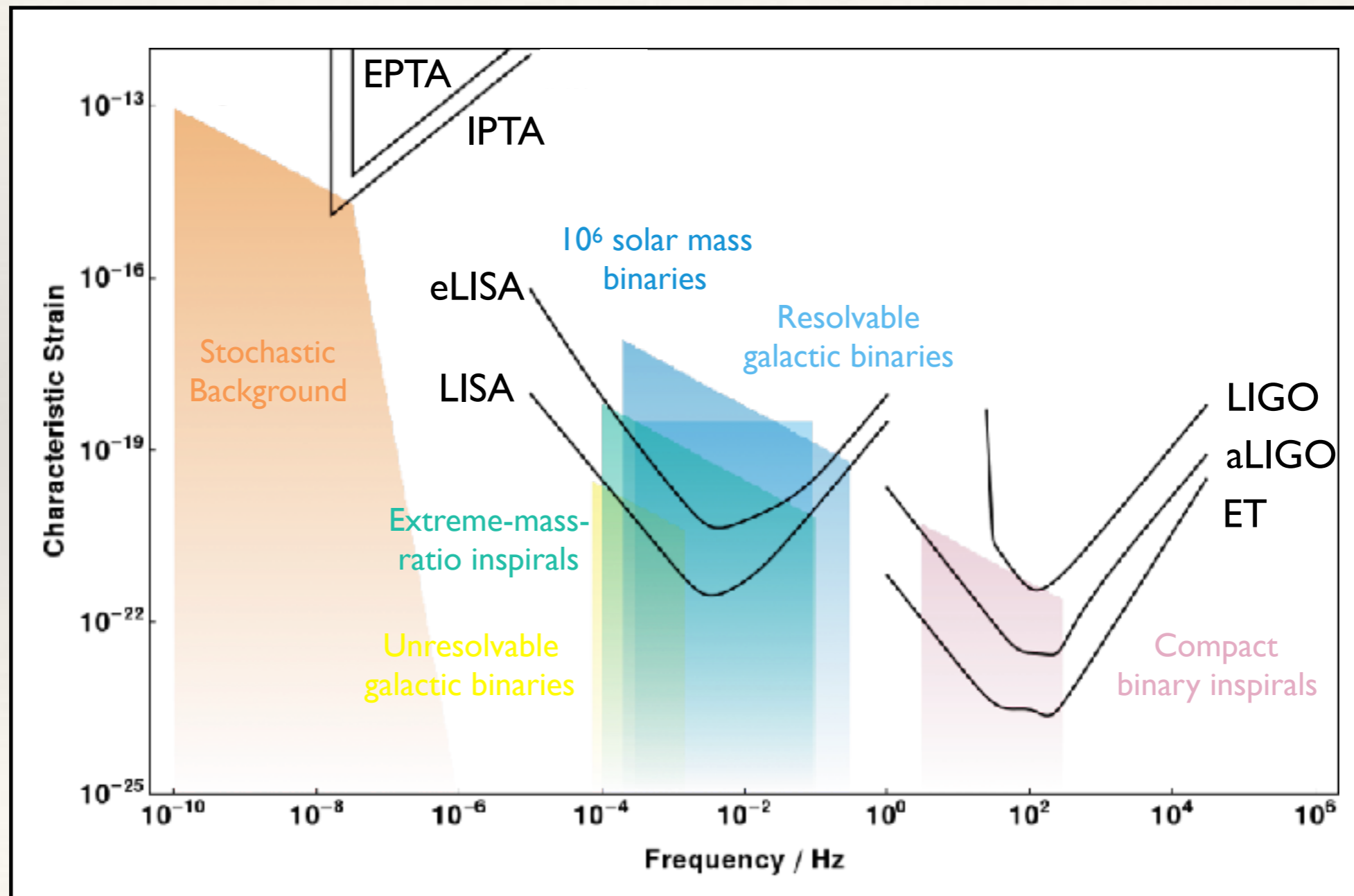


Making sense of data: introduction to statistics for gravitational wave astronomy

Lecture 7: stochastic processes and sensitivity curves

AEI IMPRS Lecture Course

Jonathan Gair jgair@aei.mpg.de



Principles of signal analysis

- ❖ Gravitational wave detectors are intrinsically noisy. The output $s(t)$ will consist of a (possible) signal $h(t)$ plus noise fluctuations $n(t)$.

$$s(t) = h(t) + n(t)$$

- ❖ The noise is a random process.
- ❖ Future values are not uniquely determined by initial data, but evolves according to some probabilistic model.
- ❖ We suppose the random process is drawn from an *ensemble of random processes characterised by probability distributions*

$$p_N(n_N, t_N; \dots; n_2, t_2; n_1, t_1) dn_N \dots dn_2 dn_1$$

Principles of signal analysis

- ❖ We typically make various useful assumptions about the properties of a random process
 - *Stationarity*: A stationary process is one for which the probability distributions depend only on time differences, not absolute time.

$$p_N(n_N, t_N + \tau; \dots; n_2, t_2 + \tau; n_1; t_1 + \tau) = p_N(n_N, t_N; \dots; n_2, t_2; n_1; t_1) \quad \forall \tau$$

- *Gaussianity*: A process is Gaussian if and only if all of its (absolute) probability distributions are Gaussian.

$$p_N(n_N, t_N; \dots; n_1; t_1) = A \exp \left[-\frac{1}{2} \sum_{j=1}^N \sum_{k=1}^N \alpha_{jk} (n_j - \bar{n}_j)(n_k - \bar{n}_k) \right]$$

- *Ergodicity*: An ensemble of stationary random processes is ergodic if for any process $n(t)$ drawn from the ensemble, the new ensemble $\{n(t+KT): K \text{ an integer}\}$ has the same probability distributions.

Principles of signal analysis

- ❖ We are interested in how large the random fluctuations are about the mean value. We'll assume this is zero here, which can be arranged by subtracting a constant.
- ❖ The fluctuations can be characterised by the power in a certain time interval $-T/2 < t < T/2$

$$\int_{-T/2}^{T/2} |n(t)|^2 dt$$

- ❖ For stationary random processes this increases linearly with time. So, we instead use the mean power (or mean square fluctuations)

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

Principles of signal analysis

- ❖ Defining $n_T(t) = n(t)\mathbb{I}[|t| < T/2]$ and using Parseval's theorem we have

$$\int_{-T/2}^{T/2} [n(t)]^2 dt = \int_{-\infty}^{\infty} [n_T(t)]^2 dt = \int_{-\infty}^{\infty} |\tilde{n}_T(f)|^2 df = 2 \int_0^{\infty} |\tilde{n}_T(f)|^2 df$$

$$P_n = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [n(t)]^2 dt = \lim_{T \rightarrow \infty} \frac{2}{T} \int_0^{\infty} |\tilde{n}_T(f)|^2 df$$

- ❖ This motivates defining the spectral density, $S_n(f)$, via

$$S_n(f) = \lim_{T \rightarrow \infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} n(t) \exp(2\pi i f t) dt \right|^2$$

- ❖ This is the **one-sided spectral density** which assumes the time series is real and we only consider positive frequencies. The **two-sided spectral density** is half this.

Principles of signal analysis

- ❖ The spectral density represents the power in the process at a particular frequency

$$P_n = \int_0^{\infty} S_n(f) df$$

- ❖ If we consider the evolution of the process over a time interval Δt , with corresponding **bandwidth** $\Delta f = 1/\Delta t$, the mean square fluctuations in n at that frequency are

$$[\Delta n(\Delta t, f)]^2 \equiv \lim_{N \rightarrow \infty} \frac{2}{N} \sum_{n=-N/2}^{N/2} \left| \frac{1}{\Delta t} \int_{n\Delta t}^{(n+1)\Delta t} n(t) \exp(2\pi i f t) dt \right|^2 = \frac{S_n(f)}{\Delta t} = S_n(f) \Delta f$$

- ❖ The *root mean square fluctuations at frequency f and measured over a time Δt* are just

$$\Delta n(\Delta t, f)_{\text{rms}} = \sqrt{S_n(f) \Delta f}$$

Principles of signal analysis

- ❖ The auto-correlation function of a (zero mean) time series is defined by

$$C(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} n(t)n(t + \tau)dt$$

- ❖ For an ergodic (and hence stationary) random process this is equivalent to the expectation value over the ensemble

$$C(\tau) = \langle n(t) n(t + \tau) \rangle$$

- ❖ The auto-correlation function is the Fourier transform of the spectral density (the Wiener-Khinchin theorem).

Principles of signal analysis

- ❖ For stationary processes a consequence of the Wiener-Khinchin theorem is that

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

- ❖ where \sim denotes the Fourier transform, and $*$ denotes complex conjugation.
- ❖ Examples of spectral densities include

<i>white noise spectrum</i>	$S_n(f) = \text{const.}$
<i>flicker noise spectrum</i>	$S_n(f) \propto 1/f$
<i>random walk spectrum</i>	$S_n(f) \propto 1/f^2$

Principles of signal analysis

- ❖ Can also define a **cross-spectral density** between two separate random process $n(t)$ and $m(t)$

$$S_{nm}(f) = \lim_{T \rightarrow \infty} \frac{2}{T} \left[\int_{-T/2}^{T/2} n(t) \exp(-2\pi i f t) dt \right] \left[\int_{-T/2}^{T/2} m(t) \exp(2\pi i f t') dt' \right]$$

- ❖ Similarly we can define the cross-correlation between two time series

$$C_{nm}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} n(t) m(t + \tau) dt$$

- ❖ As in the case of a single process, these are related to each other via a Fourier transform.

Principles of signal analysis

- ❖ For a Gaussian, stationary random process the spectral density conveys all the information about the statistical properties of the process.
- ❖ For gravitational wave detectors, it is natural therefore to plot the spectral density to characterise the detector sensitivity. But - how then do we represent sources on the same diagram?
- ❖ There is no unique way to do this. Different types of source are best represented in different ways.

Signal sensitivity: Bursts

- ❖ A transient burst of gravitational waves can be characterised by its **frequency**, f , its **duration**, Δt , its **bandwidth**, Δf , and its mean square amplitude, a proxy for signal power

$$\bar{P}_h = \frac{1}{\Delta t} \int_0^{\Delta t} |h(t)|^2 dt = h_c^2$$

- ❖ The square root of this defines the **characteristic amplitude** of the burst, h_c .
- ❖ The power in the noise in the same bandwidth is

$$\Delta f S_n(f)$$

Signal sensitivity: Bursts

- ❖ The square root of the ratio of the signal power to the noise power is the signal-to-noise ratio.

$$\left(\frac{S}{N}\right)^2 = \frac{\bar{P}_h}{\Delta f S_h(f)} = \frac{h_c^2}{\Delta f S_h(f)}$$

- ❖ This is a measure of detectability. If we window and bandpass the time series, this is the ratio of the root-mean-square signal contribution to the root-mean-square noise contribution.
- ❖ For a broad-band burst with $\Delta f \sim f$, the signal-to-noise ratio is approximately

$$\left(\frac{S}{N}\right)^2 = \frac{h_c^2}{f S_h(f)}$$

- ❖ This motivates plotting $f S_h(f)$ instead of the power spectral density. Height above this curve is a measure of burst detectability.

Signal sensitivity: continuous waves

- ❖ Consider now a monochromatic GW source

$$h(t) = h_0 \exp(2\pi i f_0 t)$$

- ❖ The signal power is constant over time and given by

$$P_h = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} |h(t)|^2 dt = \frac{1}{2} h_0^2$$

- ❖ However, this power is concentrated at f_0 . With finite time series of length T we can resolve frequency to a precision

$$\Delta f \sim 1/T$$

- ❖ Noise power in this bandwidth is $S_n(f)/T$.

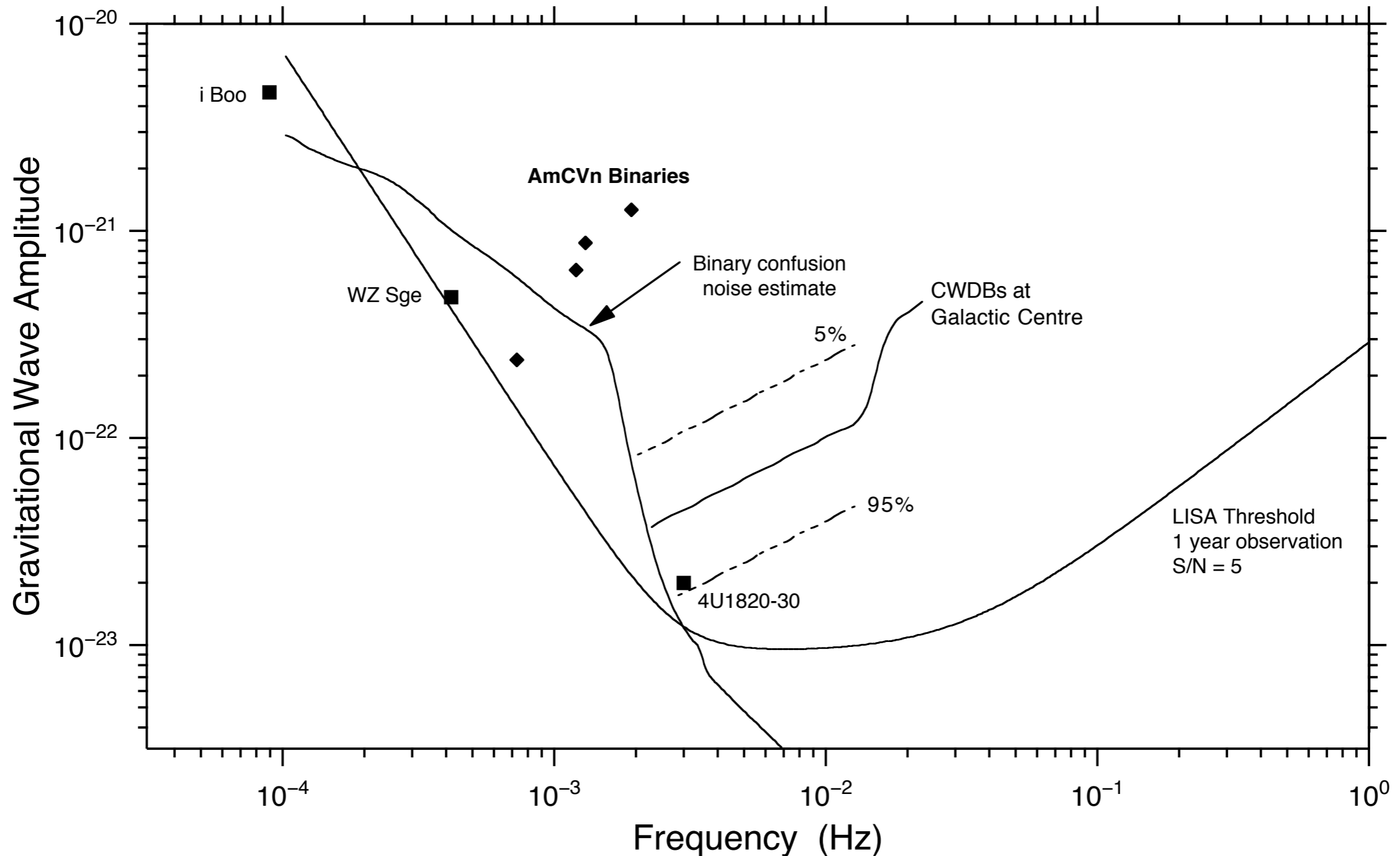
Signal sensitivity: continuous waves

- ❖ This motivates representing sensitivity by plotting

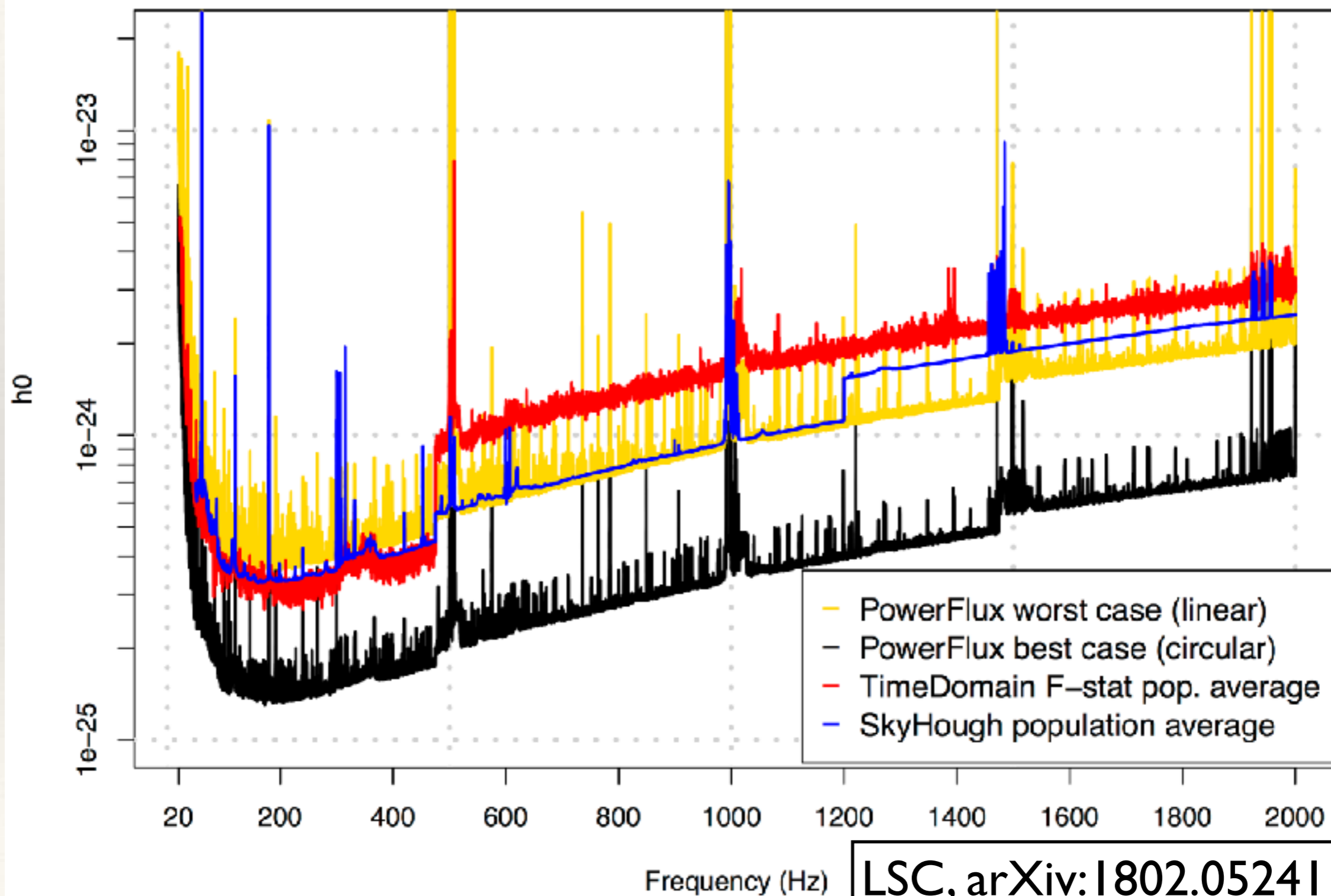
$$\sqrt{S_n(f)/T} \quad \text{or} \quad \rho_{\text{thresh}} \sqrt{S_n(f)/T}$$

- ❖ where ρ_{thresh} is the estimated threshold S/N needed for detection. This is the strain spectral density.
- ❖ Advantage: for a monochromatic source, height above curve gives expected S/N or, with specified threshold, an easy assessment of whether source is detectable or not.
- ❖ Disadvantage: must specify length of observation. Not appropriate for ongoing experiments, e.g., LIGO. But can produce this after each observing run.

Signal sensitivity: continuous waves



Signal sensitivity: continuous waves



LSC, arXiv:1802.05241 (2018)

Signal sensitivity: continuous waves

- ❖ SNRs also depend on the **sky position** and **orientation** of a source. This can be folded into the spectral density by using a *sky and orientation averaged sensitivity*, and using the strain of an optimally positioned and oriented source.

$$\langle S_h(f) \rangle_{SA}^{LIGO} \approx 5 S_h(f)$$

$$\langle S_h(f) \rangle_{SA}^{LISA} \approx \frac{20}{3} S_h(f)$$

Signal sensitivity: inspiraling sources

- ❖ For an inspiraling source, the total energy emitted in each frequency band is finite and so is the Fourier transform.

- ❖ Hence

$$\frac{1}{\sqrt{T}} \tilde{h}(f) \Rightarrow 0 \quad \text{as} \quad T \rightarrow \infty$$

- ❖ and so the spectral density is zero (over all time).
- ❖ Band passing and windowing can recover some of the power, but can we do better than this?
- ❖ Yes, using **filtering**.

Filtering

- ❖ A filtered time series is defined using a **kernel** $K(t-t')$.

$$w(t) = \int_{-\infty}^{\infty} K(t - t')s(t')dt'$$

- ❖ We now apply a slightly modified definition of S/N. We compare the amplitude output of the filter due to the signal to the rms output of the filter due to the noise.

$$\left(\frac{S}{N}\right)(t) = \frac{\int_{-\infty}^{\infty} K(t - t')h(t')dt'}{\sqrt{\left\langle \left| \int_{-\infty}^{\infty} K(t - t')n(t')dt' \right|^2 \right\rangle}}$$

- ❖ The rms output of the filter, $S+N$, is the signal amplitude to within an rms fractional error N/S , which is the reciprocal of the signal to noise ratio.

Optimal filter

- ❖ We can ask what choice of filter maximises the value of S/N at zero-lag, i.e., $t=0$.
- ❖ From the convolution theorem for Fourier transforms we have

$$\tilde{w}(f) = \tilde{K}(f)\tilde{h}(f)$$

- ❖ The expression for S/N can thus be written

$$\frac{S}{N} = \frac{\int \tilde{K}(f)\tilde{h}(f)df}{\sqrt{\int |\tilde{K}(f')|^2 S_h(f')df'}}$$

Optimal filter

- ❖ This motivates a natural inner product, $(\mathbf{h}_1 | \mathbf{h}_2)$, on the space of signals of the form

$$(\mathbf{h}_1 | \mathbf{h}_2) = 2 \int_0^\infty \frac{\tilde{\mathbf{h}}_1(f) \tilde{\mathbf{h}}_2^*(f) + \tilde{\mathbf{h}}_1^*(f) \tilde{\mathbf{h}}_2(f)}{S_h(f)} df$$

- ❖ in terms of which we have

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

- ❖ which is maximised by the choice

$$\tilde{K}(f) \propto \frac{\tilde{h}(f)}{S_h(f)}$$

- ❖ This is the **Weiner optimal filter**. In the frequency domain the optimal kernel is equal to the signal weighted by the spectral density of the noise.

Optimal filter

- ❖ A search using the optimal filter then amounts to taking the inner product $(\mathbf{s} | \mathbf{h})$ of the data stream, \mathbf{s} , with a **template** of the signal, \mathbf{h} . This is **matched filtering**.
- ❖ The signal to noise ratio of a matched filtering search is

$$\frac{S}{N}[\mathbf{h}] = \frac{(\mathbf{h} | \mathbf{h})}{\sqrt{\langle (\mathbf{h} | \mathbf{n})(\mathbf{h} | \mathbf{n}) \rangle}} = (\mathbf{h} | \mathbf{h})^{1/2}$$

- ❖ which follows from the fact that

$$\langle (\mathbf{h}_1 | \mathbf{n})(\mathbf{h}_2 | \mathbf{n}) \rangle = (\mathbf{h}_1 | \mathbf{h}_2)$$

- ❖ For a monochromatic source, the matched filter is just a Fourier transform, so this agrees with the previous result. In that case, the signal to noise ratio increases like the square root of the observation time.

Optimal filter

- ❖ The matched filtering $(S/N)^2$ is

$$\left(\frac{S}{N}\right)^2 = 4 \int_0^\infty \frac{|\tilde{h}(f)|^2}{S_h(f)} df$$

- ❖ which can also be written as

$$\left(\frac{S}{N}\right)^2 = 4 \int_0^\infty \frac{f |\tilde{h}(f)|^2}{S_h(f)} d \ln f = 4 \int_0^\infty \frac{f^2 |\tilde{h}(f)|^2}{f S_h(f)} d \ln f$$

- ❖ These expressions aid “integration by eye” in a logarithmic plot.
- ❖ For a source which has amplitude h_0 at frequency f and corresponding frequency derivative \dot{f} , we have

$$\tilde{h}(f) \sim \frac{h_0}{\sqrt{\dot{f}}}$$

Characteristic Strain

- ❖ The analogy with a broad-band burst therefore motivates the definition of a characteristic strain, h_c , for inspiraling sources (e.g., Finn and Thorne 2000).

$$h_c = h_0 \sqrt{\frac{2f^2}{df/dt}}$$

- ❖ The characteristic strain is a measure of the SNR accumulated while the frequency sweeps through a bandwidth equal to frequency. If we also plot the rms noise in a bandwidth equal to frequency,

$$h_n(f) \equiv \sqrt{f \langle S_h(f) \rangle_{SA}} \quad \left(\frac{S}{N} \right)_{f \rightarrow 2f}^2 = \left[\frac{h_c(f)}{h_n(f)} \right]^2$$

- ❖ Plots of $h_c(f)$ and $h_n(f)$ allow us to see directly how the SNR of an evolving source builds up over the evolution.

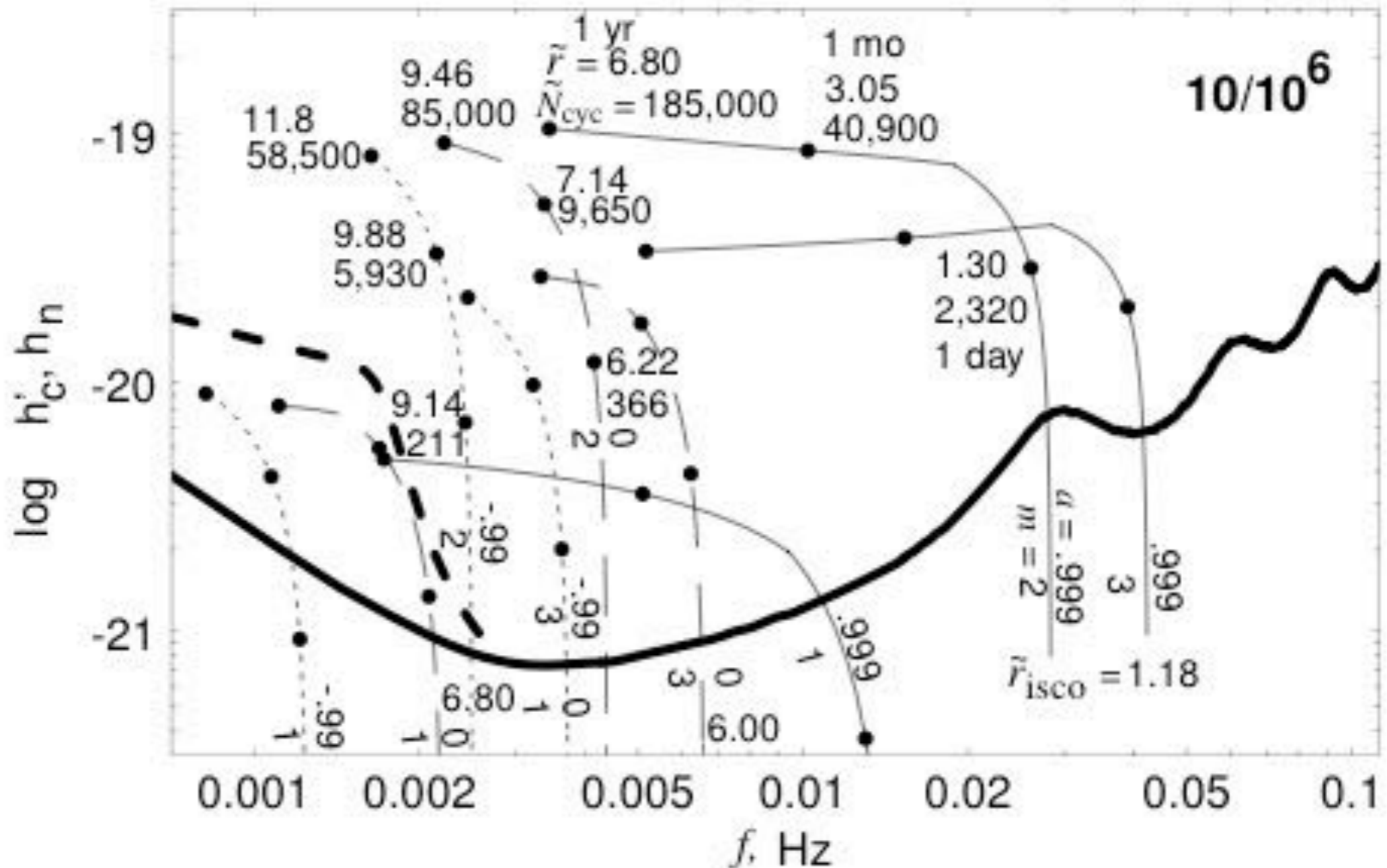
Characteristic Strain

- ❖ In the definition of characteristic strain

$$h_c = h_0 \sqrt{\frac{2f^2}{df/dt}}$$

- ❖ the term inside the square root is equal to the number of cycles the inspiral spends in the vicinity of the frequency f .
- ❖ You will read papers in which people talk about S/N being enhanced by the number of cycles spent in the vicinity of a certain frequency. This is what they mean.
- ❖ Note: plotting characteristic strain only makes sense if you are also plotting $f S_h(f)$. If you are plotting $S_h(f)$ directly your strain should be a factor of \sqrt{f} lower.

Characteristic Strain



Build up of SNR for EMRIs observed by LISA (Finn & Thorne 2000)

Representing stochastic backgrounds

- ❖ Stochastic backgrounds are characterised by a spectral density, so it is natural to compute the power spectral density and plot it on the same axes as the detector PSD.
- ❖ There are two caveats.
 - Firstly, the “power” we have been talking about so far has not been a power in a physical sense since we have not specified any units for the time series (and indeed for GW strain this is dimensionless). Better to use something that represents a physical energy density if possible.
 - Plotting two PSDs does not convey any information about their distinguishability. Can we represent backgrounds in a way that allows the reader to assess detectability at a glance?

Representing stochastic backgrounds

- ❖ The energy density carried by a gravitational wave is

$$\frac{dE}{dt dA} \propto \dot{h}_+^2 + \dot{h}_\times^2$$

- ❖ Therefore, we should consider the time derivative of the strain series to get a physical energy.
- ❖ The corresponding spectral density is $f^2 S_h(f)$ and fluctuations in a bandwidth equal to frequency are $f^3 S_h(f)$.
- ❖ Energy densities in astrophysical and cosmological backgrounds are often expressed as a fraction of the closure density of the Universe

$$\Omega_{\text{GW}} = \frac{8\pi G}{3H_0^2} \frac{dE_{\text{GW}}}{d \ln f} \propto f^2 h_c^2(f)$$

Representing stochastic backgrounds

- ❖ Suppose background is generated by an astrophysical population of sources with coming volume density $N(z)$. Then, total energy density in background today is

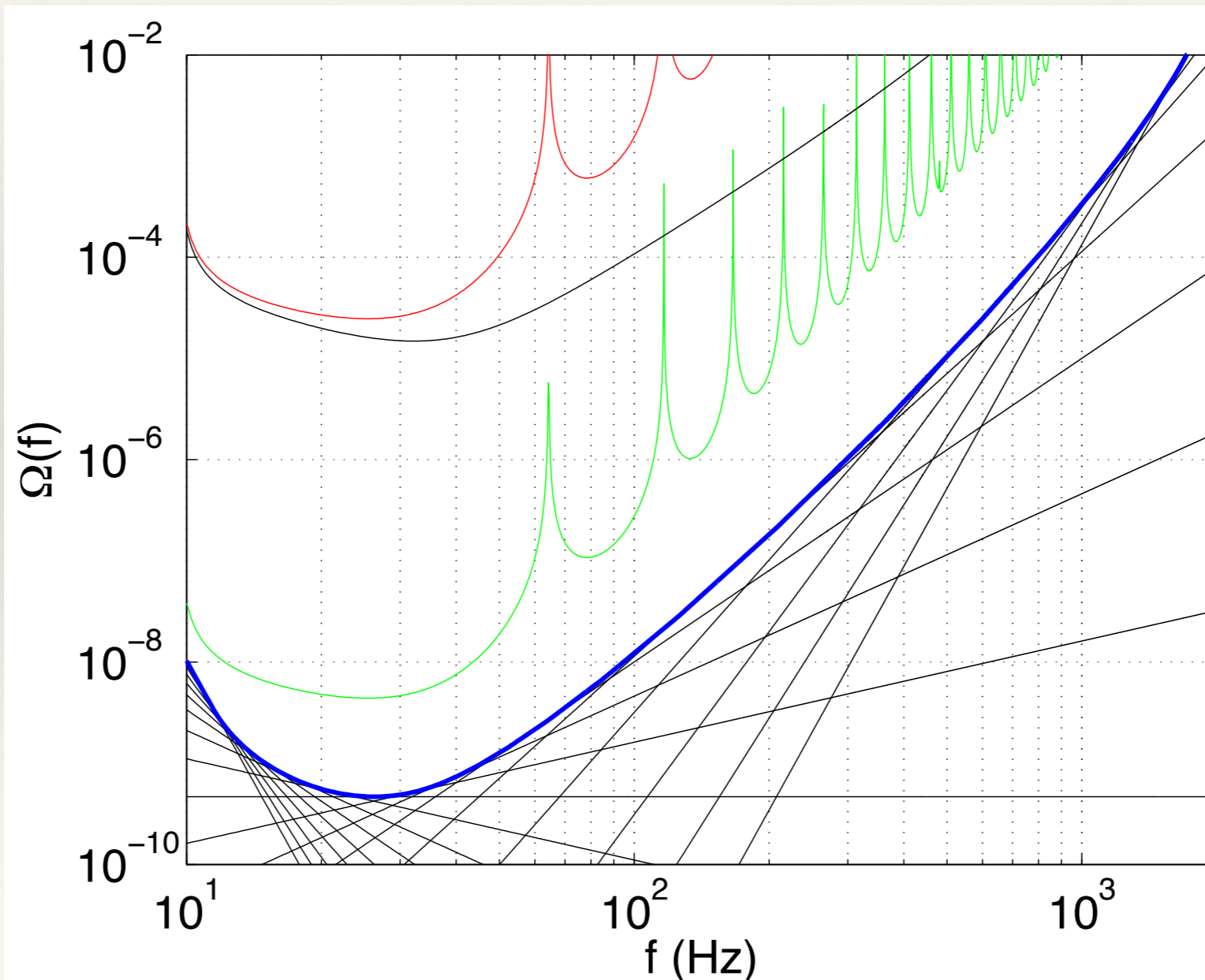
$$\mathcal{E}_{\text{GW}} = \int_0^\infty \rho_c c^2 \Omega_{\text{GW}} d \ln f = \int_0^\infty \int_0^\infty N(z) \frac{1}{(1+z)} \frac{dE}{df} f \frac{df}{f} dz$$

- ❖ We deduce (Phinney 2001, astro-ph/0108028)

$$\rho_c c^2 \Omega_{\text{GW}} = \frac{\pi c^2}{4 G} f^2 h_c^2(f) = \int_0^\infty \frac{N(z)}{1+z} \left(f_r \frac{dE}{df_r} \right)_{|f_r=f(1+z)} dz$$

Representing stochastic backgrounds

- ❖ Quick assessment of background detectability can be derived from power-law sensitivity curves (Thrane & Romano 2013). Requires assumptions about data analysis procedures.



Sensitivity curves: summary

❖ To summarise, there are four different types of sensitivity curve you might see in figures.

❖ **Power Spectral Density** - summarises statistical properties of noise

$$S_n(f)$$

❖ **Strain spectral density**

$$S_n(f)/T \quad - \text{ for monochromatic sources}$$

$$f S_n(f) \quad - \text{ for inspirals and bursts}$$

❖ **Energy spectral density** - for backgrounds

$$f^3 S_n(f)$$

Example: compact binary inspirals

- ❖ For Keplerian binaries we have

$$M = M_1 + M_2 \quad \mu = \frac{M_1 M_2}{M_1 + M_2} \quad r_1 M_1 = r_2 M_2 = \mu r \quad E = -\frac{M\mu}{2r}$$

- ❖ The period is

$$\omega^2 = \left(\frac{2\pi}{T}\right)^2 = (2\pi f)^2 = \frac{M}{r^3}$$

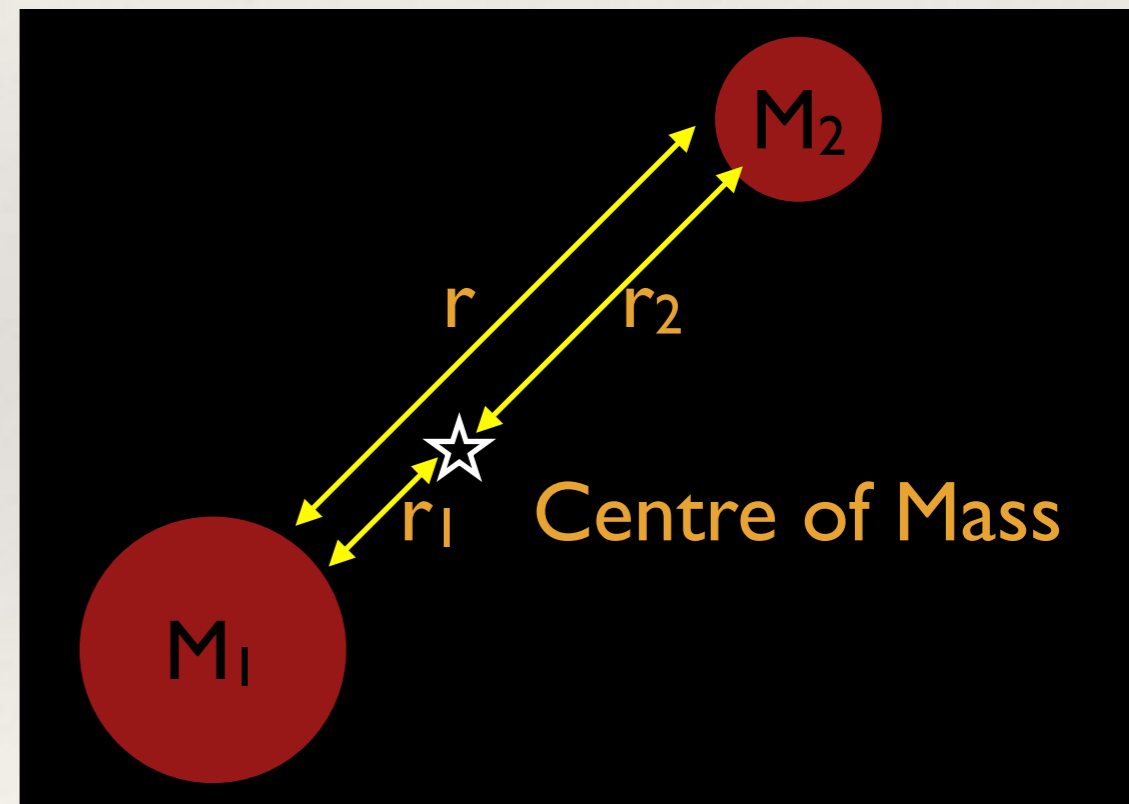
- ❖ The quadrupole moment can be estimated

$$I \sim \mu r^2 \cos 2\omega t \sim \frac{M_1 M_2}{(M_1 + M_2)^{\frac{1}{3}}} \omega^{-\frac{4}{3}}$$

- ❖ From which we deduce

$$h \sim \frac{\ddot{I}}{D} \sim \frac{1}{D} \frac{M_1 M_2}{(M_1 + M_2)^{\frac{1}{3}}} \omega^{\frac{2}{3}}$$

$$\dot{E} \sim \dot{I}^2 \sim \mu^2 M^{\frac{4}{3}} \omega^{\frac{10}{3}}$$



Example: compact binary inspirals

- ❖ From this we obtain

$$\dot{\omega} \sim \frac{M_1 M_2}{(M_1 + M_2)^{\frac{1}{3}}} \omega^{\frac{11}{3}} = M_c^{\frac{5}{3}} \omega^{\frac{11}{3}} \qquad M_c = \frac{M_1^{\frac{3}{5}} M_2^{\frac{3}{5}}}{(M_1 + M_2)^{\frac{1}{5}}}$$

- ❖ For an individual source we have

$$\tilde{h}(f) \sim \frac{1}{D} M_c^{\frac{5}{6}} f^{-\frac{7}{6}} \qquad h_c(f) \sim \frac{1}{D} M_c^{\frac{5}{6}} f^{-\frac{1}{6}}$$

- ❖ For a background generated by inspiraling binaries we have instead

$$f \frac{dE}{df} \sim M_c^{\frac{5}{3}} f^{\frac{2}{3}} \qquad \Omega_{\text{GW}}(f) \sim M_c^{\frac{5}{3}} f^{\frac{2}{3}} \int_0^\infty \frac{N(z)}{(1+z)^{\frac{1}{3}}} dz$$

- ❖ Which yields the alternative scaling

$$h_c(f) \sim \sqrt{\Omega_{\text{GW}}(f)}/f \sim M_c^{\frac{5}{6}} f^{-\frac{2}{3}} \qquad S_h(f) \sim \Omega_{\text{GW}}(f)/f^3 \sim M_c^{\frac{5}{3}} f^{-\frac{7}{3}}$$