6 Stochastic processes and sensitivity curves

In both frequentist and Bayesian approaches to statistical analysis, the likelihood plays a key role. This is the probability distribution from which the observed data has been drawn. In a gravitational wave context, we are typically concerned with analysing data from a noisy detector. The output from the detector, or detectors, is one or more real time series of measurements, $s_i(t)$. These measurements are a combination (usually assumed to be linear) of a signal part, $h_i(t)$, and a noise part, $n_i(t)$. The signal part is deterministic, depending only on the (unknown) parameters of the system, while the noise part is random. The likelihood is therefore a statement about the probability distribution from which the noise is drawn. The usual assumption is that the noise is generated by a **stationary**, **Gaussian random process**. In this section we will first define what this means, and discuss various approaches that are commonly used to summarise the noise properties and represent sensitivities to sources of different types.

6.1 Properties of random processes

A random process is a random sequence (often infinite in length) of values. Future values are not uniquely determined by current values, but by probability distributions that may be conditional on past values of the sequence. The observed random sequence is assumed to be drawn from *an ensemble of random processes* characterised by probability distributions

$$p_N(n_N, t_N; n_{N-1}, t_{N-1}; \ldots; n_2, t_2; n_1; t_1)\mathrm{d}n_N\mathrm{d}n_{N-1}\ldots\mathrm{d}n_2\mathrm{d}n_1.$$

The probability distribution could be anything, but it is usual to make some simplifying assumptions, which are well motivated by observed random processes, to make computations plausible. The most commonly made assumptions are that the random process is **stationary**, **Gaussian** and **ergodic**.

A **stationary** random process is one for which the joint probability distributions for finite sets of samples depend only on time differences, not absolute time. In other words

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

A random process is **Gaussian** if and only if all of its absolute probability distributions are Gaussian. In other words, for any set of N times, $\{t_1, \ldots, t_N\}$, we have

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

A ensemble of random process is **ergodic** if for any process, $n(t)$, drawn from the ensemble, the new ensemble defined by $\{n(t+KT) :\ k \in \mathbb{Z}\}$ has the same probability distributions.

To understand random processes, we are interested in both their mean values and the size of random fluctuations about the mean. We assume in the following (without loss of generality) that the mean of the random process is zero. Fluctuations about the mean can be characterised by the noise power (or variance), over a certain time interval $-T/2 < t < T/2$

$$\int_{-T/2}^{T/2} |n(t)|^2 \mathrm{d}t.$$

This quantity increase with time, linearly for stationary random processes. Therefore, it is more useful to work with the average value, referred to as the **mean power** or **mean square fluctuations**

$$P_n = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |n(t)|^2 \mathrm{d}t.$$

It is useful to analyse quantities in the Fourier domain, and so we define

$$n_T(t) = n(t)\mathbb{I}\left[|t| < T/2\right],$$

which is just the full series truncated to the interval of interest. This notation allows us to use Parseval's Theorem

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

and we see that the mean square fluctuations are given by

$$P_n = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} [n(t)]^2 = \lim_{T\to\infty} \frac{2}{T} \int_0^{\infty} |\tilde{n}_T(f)|^2 \mathrm{d}f.$$

This motivates the definition of the spectral density, $S_n(f)$, via

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

This is the **one-sided spectral density**, which assumes that the time series is real and hence we only need to consider positive frequencies. The **two-sided spectral density**, which is is also defined for negative frequencies, is one half of this.

The spectral density represents the power in the process at a particular frequency since we have

$$P_n = \int_0^{\infty} S_n(f) \mathrm{d}f.$$

Suppose we are interested in the properties of the process in time intervals of length Δt, with corresponding **bandwidth** $\Delta f = 1/\Delta t$. The mean square fluctuations at frequency f in intervals of length Δt, and averaged over all intervals of that length, are

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

Hence we see that the *root mean square fluctuations at frequency f and measured over a time Δt* are just $\Delta n(\Delta t, f)_{\mathrm{rms}} = \sqrt{S_n(f)\Delta f}$. The spectral density can be interpreted in this way as the size of mean square fluctuations at the specified frequency.

A property of a random process that is closely linked to the spectral density is the **auto-correlation function**. This is defined in the standard way

$$C(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} n(t) n(t+\tau) \mathrm{d}t.$$

For random processes that are ergodic (which implies they are also stationary), the averaging over time is equivalent to averaging 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). A consequence of this result is that the expectation value of noise products can be written

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

which is a statement that fluctuations of a stationary random process at different frequencies are uncorrelated with one another.

The spectral densities of a number of common noise processes are as follows

$$\begin{array}{ll} \textit{white noise spectrum} & S_n(f) = \text{const.} \\ \textit{flicker noise spectrum} & S_n(f) \propto 1/f \\ \textit{random walk spectrum} & S_n(f) \propto 1/f^2 \end{array} .$$

We conclude this section by noting that it is also possible to define a **cross-spectral density** between two separate random processes. This is defined via

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

and is the related via Fourier transform to the cross-correlation function of the two time series

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

6.2 Sensitivity curves

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 therefore natural to plot the spectral density to characterise the detector sensitivity. But - how should sources be presented on the same diagram? There is no unique way to do this. Different types of source are best represented in different ways.

6.2.1 Burst signals

Burst signals are by definition compact in time duration, and usually also in frequency duration. It is rare that burst signals can be represented by parametric models, and so they are quite like random processes. We can characterise the burst by its frequency, f, duration, Δt, bandwidth, Δf, and its mean square amplitude, a proxy for the signal power

$$\bar{P_h} = \frac{1}{\Delta t}\int_0^{\Delta t} |h(t)|^2\mathrm{d}t = h_c^2.$$

The square root of the mean square amplitude is called the **characteristic amplitude** of the burst. The power of the noise in the same bandwidth is $\Delta f S_n(f)$. The ratio of the power in the signal to the power in the noise is a measure of the detectability of the burst,

relative to random fluctuations in the instrument. This ratio is the **signal-to-noise ratio** squared of the burst

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

If the data is windowed and bandpassed in the vicinity of the burst, then we maximise the contribution of the burst to the data and the signal-to-noise ratio is the ratio of the root-mean-square (rms) signal contribution to the rms noise contribution. For a broad-band burst with $\Delta f \sim f$ we have

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

This motivates representing the sensitivity of a detector to bursts by plotting the quantity $fS_h(f)$ instead of the power spectral density. The detectability of a burst source with characteristic strain h_c can then be assessed by the height of h_c^2 above the curve.

6.2.2 Continuous sources

If instead of a burst we had a monochromatic gravitational wave source

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

then the signal power is constant over time

$$P_h = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |h(t)|^2 \mathrm{d}t = \frac{1}{2} h_0^2.$$

This power is concentrated at f_0. When observing a finite time series of length T, we can resolve frequencies to a precision $\Delta f \sim 1/T$. The noise power in that bandwidth is $S_n(f)/T$, which motivates representing the detector sensitivity curve by plotting

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

where ρ_{thresh} is the threshold signal-to–noise ratio needed for detection. This is called the **strain spectral density**. The advantage of rep[resenting sensitivity in this way is that the detectability of a source can be directly assessed by seeing if the source amplitude h_0 lies above or below the curve. The height above the curve is a direct estimate of the signal to noise ratio of the source. The disadvantage of this way of representing sensitivity is that it varies with the length of observation, so this must be specified. In the case of LIGO, this is not a problem, as the detectors take periodic breaks from observation. After each observing run, the length of observation is known and so the strain spectral density can be evaluated for each observing run after the fact, and used to represent the results.

An example of a strain spectral density curve is given in Fig. 23.

Finally, we note that rescaling the sensitivity according to the detection threshold is not the only type of rescaled spectral density that is encountered in the literature. The amplitude of a gravitational wave signal in a gravitational wave detector depends on the orientation of the source relative to the detector plane. The same source placed at different sky locations and orientations will have different signal-to-noise ratios. To avoid having to specify which particular choices are being made, it is useful to produce a **sky-averaged sensitivity curve**. To assess detectability of a source, its amplitude should then be assessed

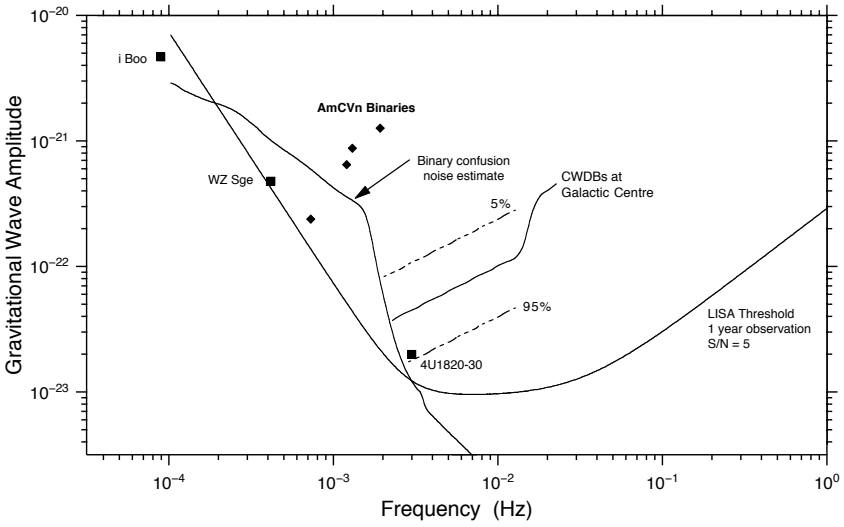

Figure 23: Strain spectral density curve for a 1 year observation with LISA and a detection threshold of $s/N = 5$. Reproduced from the LISA pre-phase A report.

for optimal orientation and sky location. The height of this optimal source above the sky averaged sensitivity is the average signal-to-noise ratio squared of a source of this type. For LIGO the sky averaged sensitivity is

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

while for LISA we have

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

The difference arises because of the 60° opening angle of the LISA arms ($\sin^2 60 = 3/4$).

6.2.3 Inspiralling sources

Inspiraling sources have to be treated differently to continuous sources. This is because they emit a finite amount of power in each frequency band and hence the Fourier transform at each frequency is also finite. Therefore

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

and so the strain spectral density of an inspiraling source is zero averaged over all time. Band-passing and windowing the data can recover some signal-to-noise ratio, as in the burst source case, but we can do better than that using **filtering**.

A filtered time series is defined from a kernel $K(t-t')$ via convolution

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

In the previous cases we considered signal-to-noise ratio as the ratio of the rms power in the presence of a signal to the rms power in the noise. We use an analogous definition for filtered

data, but now compare the amplitude of the filter output due to the signal only, to the rms amplitude of the filtered data in the presence of noise only

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

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

The choice of the kernel is arbitrary, but it makes sense to choose the kernel that is "best" in some sense. The best kernel is the one that maximises the signal-to-noise ratio. This is most easily found by working in the Fourier domain. We use the Fourier transform definition

$$\tilde{x}(f) = \int_{-\infty}^{\infty} x(t)\exp\left[-2\pi i f t\right]\,\mathrm{d}t.$$

From the convolution theorem, the Fourier transform of the filter output is

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

where $\tilde{K}(f)$ and $\tilde{h}(f)$ are the Fourier transform of the kernel and waveform respectively. We have also

$$w(t) = \int_{-\infty}^{\infty} \tilde{x}(f)\exp\left[2\pi i f t\right]\,\mathrm{d}f \quad \Rightarrow \quad w(0) = \int_{-\infty}^{\infty} \tilde{x}(f)\,\mathrm{d}f.$$

Similarly

$$\begin{aligned} N^2(0) &= \left\langle \left|\int_{-\infty}^{\infty} K(-t')n(t')\mathrm{d}t'\right|^2 \right\rangle = \left\langle \int_{-\infty}^{\infty} \tilde{K}(f)\tilde{n}(f)\mathrm{d}f \int_{-\infty}^{\infty} \tilde{K}^*(f')\tilde{n}^*(f')\mathrm{d}f' \right\rangle \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{K}(f)\tilde{K}^*(f')\left\langle \tilde{n}^*(f')\tilde{n}(f)\right\rangle \mathrm{d}f\mathrm{d}f' = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{K}(f)\tilde{K}^*(f')\delta(f-f')S_h(f)\mathrm{d}f\mathrm{d}f' \\ &= \int |\tilde{K}(f')|^2 S_h(f')\mathrm{d}f'. \end{aligned} \tag{92}$$

We deduce that the signal-to-noise ratio at zero lag is

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

which can also be written as

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

by introducing the noise-weighted inner product

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

This is of the form $\hat{\mathbf{e}}\cdot\mathbf{b}$, for a unit vector $\hat{\mathbf{e}}$ to be found. The inner product of two vectors of fixed length is maximised when they are parallel, i.e., $\hat{\mathbf{e}} \propto \mathbf{b}$. We therefore deduce that the choice which maximises the inner product is

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

This is the **Weiner optimal filter**. In the frequency domain the optimal filter is equal to the signal, weighted by the spectral density of the noise. A search using the optimal filter 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**. In practice we don't know exactly what the signal is, but the parameters of the signal must be estimated from the data. In LIGO/Virgo this is done by computing the output of the optimal filter for a large number of source parameter choices which define a **template bank**.

The signal-to-noise ratio of the matched filtering search that uses the optimal filter is

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

which follows from the fact that

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

This result is proved as follows

$$\begin{aligned}\langle(\mathbf{h}_1|\mathbf{n})(\mathbf{h}_2|\mathbf{n})\rangle &= \Big\langle \int_{-\infty}^{\infty} \frac{\tilde{\mathbf{h}}_1(f)\tilde{\mathbf{n}}^*(f) + \tilde{\mathbf{h}}_1^*(f)\tilde{\mathbf{n}}(f)}{S_h(f)}\mathrm{d}f \int_{-\infty}^{\infty} \frac{\tilde{\mathbf{h}}_2(f')\tilde{\mathbf{n}}^*(f') + \tilde{\mathbf{h}}_2^*(f')\tilde{\mathbf{n}}(f')}{S_h(f')}\mathrm{d}f' \Big\rangle \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\tilde{\mathbf{h}}_1(f)\tilde{\mathbf{h}}_2^*(f')\langle\tilde{\mathbf{n}}^*(f)\tilde{\mathbf{n}}(f')\rangle + \tilde{\mathbf{h}}_1^*(f)\tilde{\mathbf{h}}_2(f')\langle\tilde{\mathbf{n}}(f)\tilde{\mathbf{n}}^*(f')\rangle}{S_h(f)S_h(f')}\mathrm{d}f\mathrm{d}f' \\ &\quad + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\tilde{\mathbf{h}}_1(f)\tilde{\mathbf{h}}_2(f')\langle\tilde{\mathbf{n}}^*(f)\tilde{\mathbf{n}}^*(f')\rangle + \tilde{\mathbf{h}}_1^*(f)\tilde{\mathbf{h}}_2^*(f')\langle\tilde{\mathbf{n}}(f)\tilde{\mathbf{n}}(f')\rangle}{S_h(f)S_h(f')}\mathrm{d}f\mathrm{d}f'.\end{aligned} \tag{93}$$

The terms on the final line vanish because $\langle\tilde{\mathbf{n}}(f)\tilde{\mathbf{n}}(f')\rangle = 0$, i.e., the size of fluctuations in the real and imaginary components of the noise are the same. The terms on the first line are simplified using $\langle\tilde{\mathbf{n}}^*(f)\tilde{\mathbf{n}}(f')\rangle = S_h(f)\delta(f-f')$

$$\begin{aligned}\langle(\mathbf{h}_1|\mathbf{n})(\mathbf{h}_2|\mathbf{n})\rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{[\tilde{\mathbf{h}}_1(f)\tilde{\mathbf{h}}_2^*(f') + \tilde{\mathbf{h}}_1^*(f)\tilde{\mathbf{h}}_2(f')]\delta(f-f')}{S_h(f')}\mathrm{d}f\mathrm{d}f' \\ &= \int_{-\infty}^{\infty} \frac{[\tilde{\mathbf{h}}_1(f)\tilde{\mathbf{h}}_2^*(f) + \tilde{\mathbf{h}}_1^*(f)\tilde{\mathbf{h}}_2(f)]}{S_h(f)}\mathrm{d}f,\end{aligned} \tag{94}$$

giving the result stated above.

The matched filtering signal-to-noise ratio simplifies to

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

which can also be written as

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

Plotting $S_h(f)$ and $f|\tilde{h}(f)|^2$ on a logarithmic frequency plot, the integral of the ratio of the two curves "by eye" gives an estimate of the signal-to-noise ratio squared.

For a source that has amplitude h_0 at frequency f, at which point the frequency derivative is $\dot{f}$, then the stationary phase approximation gives us the scaling

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

The analogy with the broad-band burst case described above motivates defining a characteristic strain, h_c, such that the signal-to-noise ratio squared is $h_c^2/(fS_h(f))$. The appropriate definition is

$$h_c = h_0\sqrt{\frac{2f^2}{\mathrm{d}f/\mathrm{d}t}} \sim f\tilde{h}(f).$$

Note also that since $\sqrt{\dot{E}} \sim \dddot{I}$ and $h \sim \ddot{I}/D$, we have $\sqrt{\dot{E}} \sim Dfh$ and hence

$$h_c \sim \frac{1}{D}\sqrt{\frac{\dot{E}}{\dot{f}}}$$

and this is an equality for monochromatic signals.

The characteristic strain is a measure of the signal-to-noise ratio accumulated while the frequency sweeps through a bandwidth equal to frequency. If we plot as a sensitivity curve the rms noise in a bandwidth equal to frequency, which is

$$h_n(f) \equiv \sqrt{f\langle S_h(f)\rangle_{\mathrm{SA}}}$$

then the signal-to-noise ratio accumulated as the inspiral proceeds from f to $2f$ is

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)^2_{f\to 2f} = \left[\frac{h_c(f)}{h_n(f)}\right]^2.$$

Therefore, plotting characteristic strain on the same plot gives a quick way to see how the signal-to-noise ratio of an inspiraling source builds up over the evolution. Note that plotting the characteristic strain only makes sense if the detector sensitivity is represented by $fS_h(f)$. If the detector sensitivity is represented by $S_h(f)$ then the quantity $h_c/\sqrt{f}$ should be used to represent the signal.

In the definition of characteristic strain, $h_c = h_0\sqrt{2f^2/\dot{f}}$, the term inside the square root is the number of cycles the inspiral spends in the vicinity of the frequency f. Papers that discuss matched filtering often include the statement that the signal to noise ratio is enhanced by the number of cycles spent in the vicinity of a certain frequency. This is what they are referring to.

In Fig. 24 we give an example of a plot of the characteristic strain, reproduced from Finn and Thorne (2000). The figure shows the characteristic strain of various extreme-mass-ratio inspiral sources detectable by LISA.

6.2.4 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. However, there are two caveats. Firstly, the "power" we have been talking about so far has not been

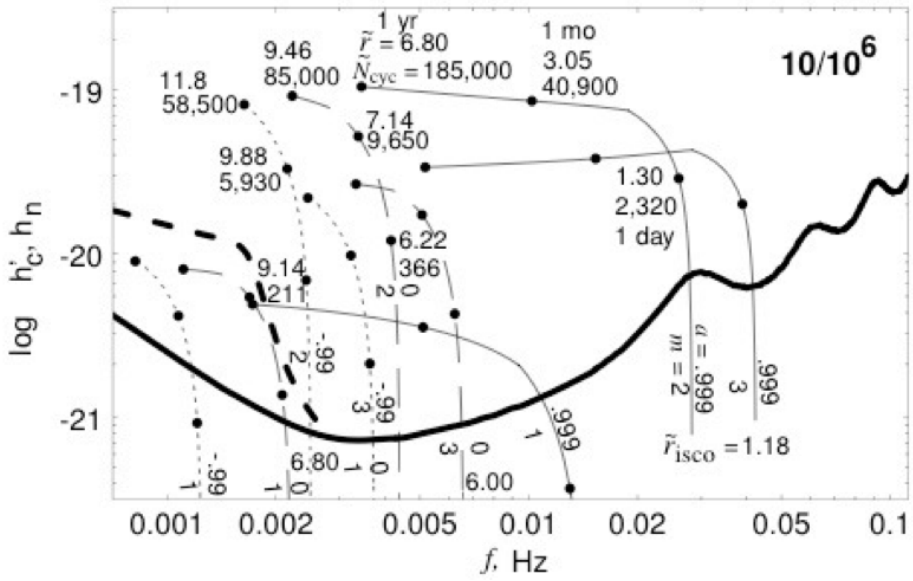

Figure 24: Characteristic strain for a number of typical extreme-mass-ratio inspiral sources observed by the classic (5 km arm length) LISA interferometer. All inspirals are circular, with $10^6 M_\odot$ central black holes and observed at a distance of 1Gpc. Curves are labelled by the spin of the central black hole, a, and the mass of the inspiraling object, m. Points on the curve correspond to 1 year, 1 month and 1 day prior to merger. The numbers above the points are the radius of the orbit (in units of M) at that time, and the number of gravitational wave cycles remaining until plunge. Reproduced from Finn and Thorne (2000).

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). When comparing to the noise power spectral density which is an energy density, it would be preferable to use something that represents a physical energy density if possible. Secondly, plotting two power spectral densities does not convey any information about their distinguishability. It would be preferable to represent a background in a way that conveys the detectability of the background at a glance.

The energy density carried by a gravitational wave is given by

$$\frac{\mathrm{d}E}{\mathrm{d}t\mathrm{d}A} \propto \dot{h}_+^2 + \dot{h}_\times^2.$$

Therefore, to obtain a physical energy density we should consider the time derivative of the strain. Differentiation with respect to time brings down a factor of frequency and so the energy spectral density is $f^2 S_h(f)$. Fluctuations of the energy spectral density in a bandwidth equal to frequency are then $f^3 S_h(f)$.

The energy density of an astrophysical or cosmological stochastic background, per logarithmic frequency interval, is often expressed as a fraction of the closure density of the Universe via

$$\Omega_{\rm GW} = \frac{8\pi G}{3H_0^2}\frac{\mathrm{d}E_{\rm GW}}{\mathrm{d}\ln f} \propto f^2 h_c^2(f).$$

The last equality defines the characteristic strain for a background, since, as argued above, a plane wave of frequency f and amplitude h_c carries an energy density $f h_c$. In the examples below we will show how to calculate the energy density for an astrophysical population of sources.

To represent backgrounds in a way that conveys their detectability directly, one can use *power-law sensitivity curves* (Thrane and Romano 2013). These are not uniquely defined, as they require some assumptions to be made about data analysis procedures and the threshold required for a detection using the defined procedure. However, given these assumptions, the procedure is as follows.

- For a given assumed power-law slope of a background, $\Omega_{\rm GW} \propto f^\beta$, compute the minimum amplitude, $A_{\rm min}(\beta)$, such that the background would be detectable by the defined procedure.
- Define the *power-law sensitivity curve*, $S_{\rm pl}(f)$, via

$$S_{\rm pl}(f) = \max\{A_{\rm min}(\beta) f^\beta : \beta \in [-\infty, \infty]\}.$$

The power-law sensitivity curve is the envelope of the minimal-detectable power-law backgrounds. It is a useful object to assess background detectability, since drawing a background of interest on the same figure gives an immediate indication of detectability. If the curve lies above the power-law sensitivity curve then it will be detectable (via the designated procedure) and otherwise it will not. An illustration of such a curve is given in Fig. 25.

6.3 Examples

We now estimate the leading order scaling of the quantities introduced above for some common astrophysical sources. Throughout we will make the usual choice of units to set $G = c = 1$.

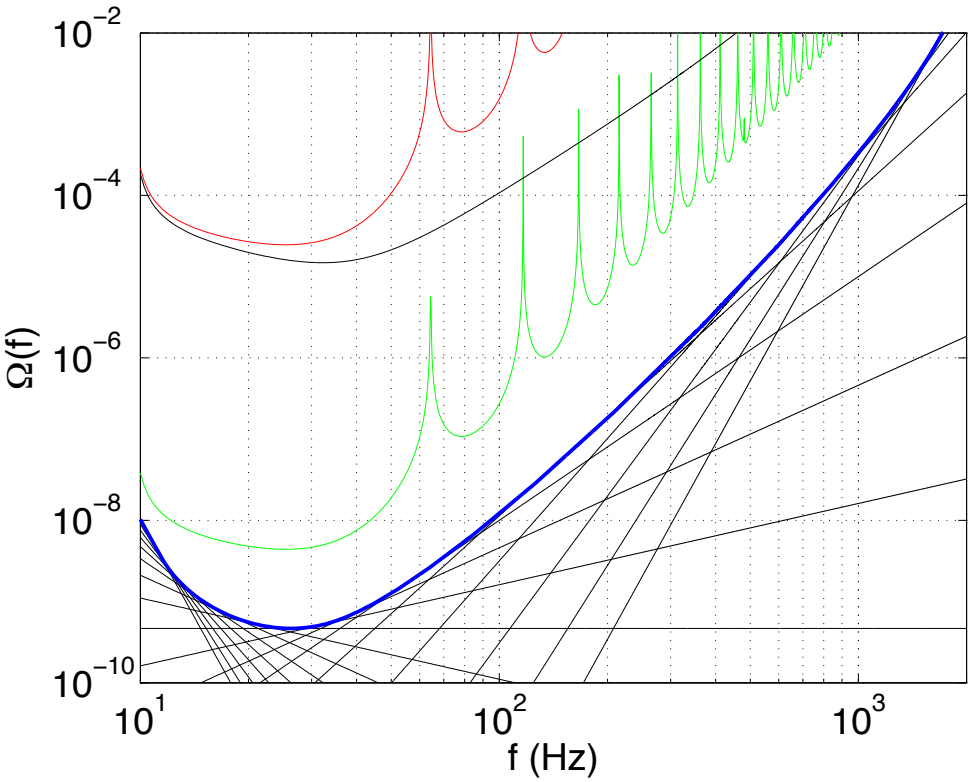

Figure 25: Power law sensitivity curve (in blue) for backgrounds detectable by ground-based interferometers, assuming the search is based on cross-correlation of the H1 and L1 detectors and the threshold for detection is a signal-to-noise ratio of 1. The solid black lines show the set of minimally-detectable power-laws that are used to generate the power-law sensitivity curve. The other curves show the detector strain spectral density instantaneously (red) and for a one year observation (green). Reproduced from Thrane and Romano (2013).

6.3.1 Single inspiraling compact binary

We consider first the case of compact binary coalescence. We assume that we have a circular binary with component masses M_1 and M_2 and separation r. We work in the Newtonian regime where the binary components are on Keplerian orbits. We denote the total mass, M, and reduced mass, μ, by

$$M = M_1 + M_2, \qquad \mu = \frac{M_1 M_2}{M_1 + M2}.$$

In the Newtonian two-body problem, the two objects each orbit around the centre of mass of the system. The two objects are distances r_1 and r_2 from the centre of mass respectively, where

$$r_1 M_1 = r_2 M_2 = \mu r.$$

The motion is also equivalent to that of a single body of mass μ orbiting in a fixed Newtonian potential of an object with mass M at a distance r. The orbital frequency is given by Kepler's laws

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

To estimate the scaling of the gravitational wave emission we need to estimate the quadrupole moment of the binary. This can be estimated from

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

At leading order, the gravitational wave strain scales like the second time derivative of the quadrupole moment divided by the distance to the source

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

The rate of energy loss scales like the third time derivative of I squared and so this has the scaling

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

Finally, we need to know how the energy relates to the orbital separation or equivalently the orbital frequency. In the Newtonian limit this follows from

$$E = -\frac{M\mu}{2r} = -\frac{\mu (M\omega)^{\frac{2}{3}}}{2}$$

from which we deduce

$$\dot{E} \sim -\mu M^{\frac{2}{3}} \omega^{-\frac{1}{3}} \dot{\omega}. \tag{95}$$

Combining this with expression (6.3.1) we obtain

$$\dot{\omega} \sim \mu M^{\frac{2}{3}} \omega^{\frac{11}{3}} = \frac{M_1 M_2}{(M_1 + M_2)^{\frac{1}{3}}} \omega^{\frac{11}{3}} = M_c^{\frac{5}{3}} \omega^{\frac{11}{3}}$$

where we have introduced the chirp mass

$$M_c = \frac{M_1^{\frac{3}{5}} M_2^{\frac{3}{5}}}{(M_1 + M_2)^{\frac{1}{5}}}.$$

We can now determine the scaling of the various quantities introduced in the previous section. From Eq. (6.2.3) and recalling $\omega = 2\pi f$, we obtain the Fourier domain amplitude

$$\tilde{h}(f) \sim \frac{h_0}{\sqrt{\dot{f}}} \sim \frac{1}{D}\frac{M_c^{\frac{5}{3}} f^{\frac{2}{3}}}{M_c^{\frac{5}{6}} f^{\frac{11}{6}}} = \frac{1}{D} M_c^{\frac{5}{6}} f^{-\frac{7}{6}}.$$

We can also deduce the characteristic strain

$$h_c(f) \sim \frac{1}{D} M_c^{\frac{5}{6}} f^{-\frac{1}{6}}.$$

6.3.2 Eccentric binaries

Eccentric binaries have gravitational wave emission at multiple harmonics of the orbital frequency (Peters and Matthews 1963). The flux of radiation at frequency nf, where n is the orbital frequency, is

$$\dot{E}_n = \frac{32}{5}\mu^2 M^{\frac{4}{3}} (2\pi f)^{\frac{10}{3}} g(n,e)$$

where $g(n,e)$ is given by

$$g(n,e) = \frac{n^4}{32}\left\{\left[J_{n-2}(ne) - 2eJ_{n-1}(ne) + \frac{2}{n}J_n(ne) + 2eJ_{n+1}(ne) - J_{n+2}(ne)\right]^2 \right.$$
$$\left. +(1-e^2)\left[J_{n-2}(ne) - 2J_n(ne) + J_{n+2}(ne)\right]^2 + \frac{4}{3n^2}\left[J_n(ne)\right]^2\right\} \quad (96)$$

where $J_n(x)$ is the Bessel function of the first kind. The characteristic strain for an individual harmonic is therefore

$$h_{c,n}(f) = \frac{1}{\pi D}\sqrt{\frac{2\dot{E}_n(f/n)}{n\dot{f}(f/n)}} \sim M_c^{\frac{5}{6}} f^{-\frac{7}{6}} n^{\frac{2}{3}} \sqrt{g(n,e)}$$

where the argument (f/n) indicates that in order to get the contribution at frequency f from the n'th harmonic, it must be evaluated when the orbital frequency had the lower value of f/n.

It is normal to represent the contributions form individual waveform harmonics on a "waterfall plot". An example is shown in Figure 26 which is reproduced from Barack and Cutler (2004).

6.3.3 Stochastic backgrounds

The energy density in a gravitational wave background was defined in equation (6.2.4). If this background is generated by a population of individual sources, the total background can be estimated by integrating the contribution from each component in the background. The quantity of relevance is the total energy density in gravitational waves today, $\mathcal{E}_{\rm GW}$. If the sources are identical, have number density $n(z)$ and each generate a differential energy density $\mathrm{d}E/\mathrm{d}f$, then we have

$$\mathcal{E}_{\rm GW} = \int_0^\infty \rho_c c^2 \Omega_{\rm GW} \mathrm{d}\ln f = \int_0^\infty \int_0^\infty N(z)\frac{1}{(1+z)}\frac{\mathrm{d}E}{\mathrm{d}f} f \frac{\mathrm{d}f}{f}\mathrm{d}z,$$

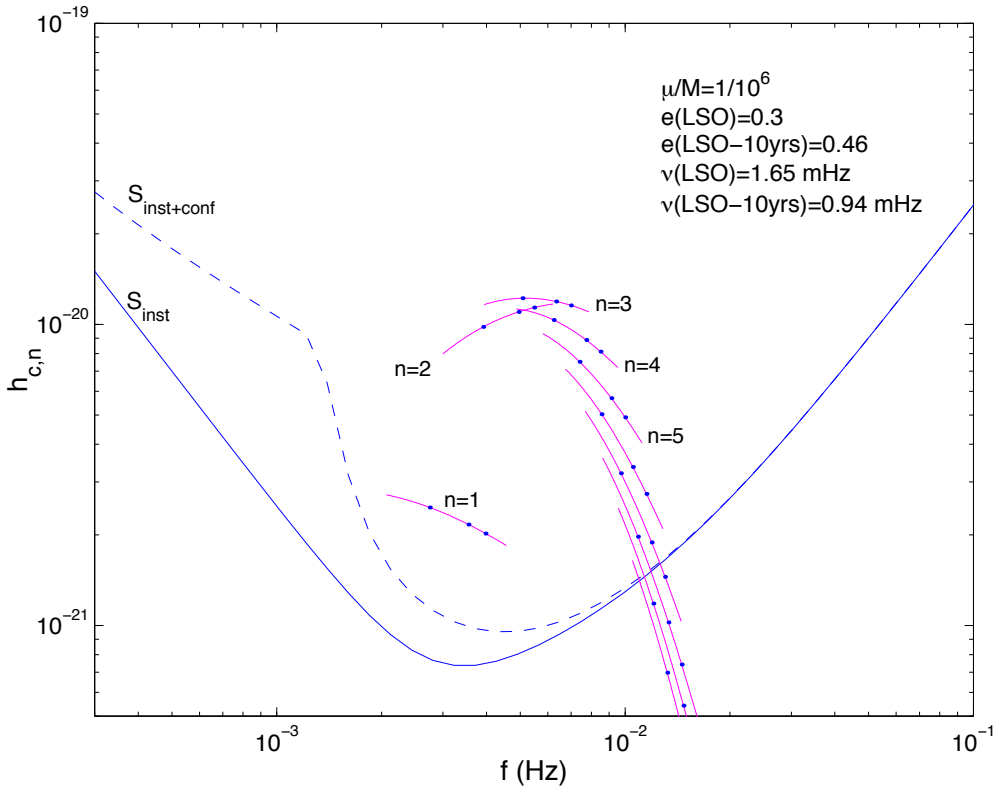

Figure 26: Characteristic strain of each harmonic in a the extreme-mass-ratio inspiral of a $1M_\odot$ black hole into a $10^6 M_\odot$ black hole with eccentricity of 0.3 at plunge. Figure reproduced from Barack and Cutler (2004).

where the factor of $(1+z)$ accounts for the fact that the energy density today is redshifted relative to the energy density at emission. We deduce

$$\rho_c c^2 \Omega_{\rm GW} = \frac{\pi}{4}\frac{c^2}{G} f^2 h_c^2(f) = \int_0^\infty \frac{N(z)}{1+z}\left(f_r \frac{{\rm d}E}{{\rm d}f_r}\right)_{|f_r = f(1+z)} {\rm d}z \tag{97}$$

where the latter quantity is evaluated at the rest frame frequency, $f_r = (1+z)f$.

For a stochastic background generated by inspiraling binary sources, from Eq. (95), we have at leading order

$$f\frac{{\rm d}E}{{\rm d}f} \sim M_c^{\frac{5}{3}} f^{\frac{2}{3}}.$$

Plugging this into Eq. (97) we obtain

$$\Omega_{\rm GW}(f) \sim M_c^{\frac{5}{3}} f^{\frac{2}{3}} \int_0^\infty \frac{N(z)}{(1+z)^{\frac{1}{3}}} {\rm d}z. \tag{98}$$

We see that the energy spectral density of the background is

$$S_h(f) \sim \Omega_{\rm GW}(f)/f^3 \sim M_c^{\frac{5}{3}} f^{-\frac{7}{3}}$$

and the characteristic strain is

$$h_c(f) \sim \sqrt{\Omega_{\rm GW}(f)}/f \sim M_c^{\frac{5}{6}} f^{-\frac{2}{3}}.$$

In this case the characteristic strain scales like $f^{-2/3}$, while in the case of a single compact binary coalescence we had a scaling that was $f^{-1/6}$. This difference arises because the definition of characteristic strain relates to the signal to noise ratio that can be obtained in a search for the source of interest. For individual sources, we can perform matched filtering and enhance the signal to noise ratio coherently by the square root of the number of cycles (approximately $\sqrt{f}$, which explains the difference between $f^{-2/3}$ and $f^{-1/6}$). This is not possible for incoherent backgrounds where we can only predict the power at each frequency, not the phase.