Chapter 1

Ranking Triggers by False Alarm Rate

Once triggers have been obtained from the matched filter, re-weighted using the χ^2 statistic, and vetted via coincidence test, they need to be evaluated for their statistical significance. In the low-mass CBC search we do this by computing a *false alarm rate* (FAR) for each coincident trigger we obtain. This chapter details how that is done. We begin by reviewing properties of a Poisson distribution. Next, in section 1.2, we model the expected number of coincident triggers from single-interferometer (IFO) distributions of a single template. In section **??** we show how to using time-slides allows us to estimate the background to measure a FAR. Section **??** describes how we do this across the template space. In section **??** we discuss limitations in using time-slides to model the background.

1.1 Poisson Process

If a system is a Poisson process, then on average, it will produce a mean number of events Λ in a length of time t. The probability of getting k events from this system in the same amount of time is:

$$P(k|\Lambda) = \frac{\Lambda^k e^{-\Lambda}}{k!} \tag{1.1}$$

Λ is known as the *rate parameter* of the system. It is a monotonically increasing function of time (for longer periods of time, we will expect a larger mean number of events). For a *stationary* source Λ will be linear in time. In other words, the mean number of events a stationary source would produce in several (hypothetical) iterations over the same period of time, τ, will be the same as the mean number produced in several iterations over any other period of equal duration. For a *non-stationary* source, the time dependence will be non-linear; i.e., two independent periods of time of equal duration would produce a different mean number of events. To get a duration-independent quantity, we define the source's *rate density*, $\mathcal{R}(t)$, which is the derivative of Λ with respect to time:

$$\mathcal{R}(t) \equiv \frac{\mathrm{d}\Lambda}{\mathrm{d}t} \tag{1.2}$$

For stationary sources, $\mathcal{R}(t)$ will be constant; for non-stationary sources, it will have a time dependence. The rate density is an *intrinsic* parameter of the source: it depends on the source's physical characteristics and is independent of the duration of time it is observed for. (Contrast this with Λ, which mixes the intrinsic parameters with the period of time that the source is observed, which is an *extrinsic* parameter). Whether or not a source produces a trigger in a given period of time is random; thus any measurement of the rate density is subject to uncertainty. We therefore define the "true" rate density as the value obtained if an infinite number of independent measurements with identical initial conditions were carried out over the same duration of time:

$$\widetilde{\mathcal{R}}(t) \equiv \lim_{N_m \to \infty} \frac{\sum_{k=1}^{N_m} \widehat{\mathcal{R}}(t)_k}{N_m} \tag{1.3}$$

In this analysis we will denote "true" rate densities by a tilde (e.g., $\widetilde{\mathcal{R}}$), measured values by a hat (e.g., $\widehat{\mathcal{R}}$), and average values by a bar (e.g., $\overline{\mathcal{R}}$).

If the source is stationary, then $\widetilde{\mathcal{R}}$ exists (at least theoretically; in practice, of course, we cannot perfrom an infinite number of measurements) and can be approximated by measuring $\mathcal{R}$ in a series of experiments. If we observe a source for a period of time t, then, from 1.2, the mean number of events produced during the experiment is:

$$\begin{aligned}
\Lambda &= \int_0^{\Lambda} \mathrm{d}\Lambda' \\
&= \int_0^{t} \widetilde{\mathcal{R}}(t')\,\mathrm{d}t' &(1.4)\\
&\underset{\widetilde{\mathcal{R}}(t')\to\widetilde{\mathcal{R}}}{=} \widetilde{\mathcal{R}}t &(1.5)
\end{aligned}$$

$\widetilde{\mathcal{R}}t$ is also the expectation value of the number of events:

$$\begin{aligned}
\mathbb{E}(N) &= \sum_{k=1}^{\infty} kP(k|\Lambda) \\
&= \sum_{k=1}^{\infty} k\frac{(\widetilde{\mathcal{R}}t)^k e^{-\widetilde{\mathcal{R}}t}}{k!} \\
&= \widetilde{\mathcal{R}}te^{-\widetilde{\mathcal{R}}t}\sum_{k=1}^{\infty}\frac{(\widetilde{\mathcal{R}}t)^{k-1}}{(k-1)!} \\
&= \widetilde{\mathcal{R}}te^{-\widetilde{\mathcal{R}}t}e^{\widetilde{\mathcal{R}}t} \\
&= \widetilde{\mathcal{R}}t &(1.6)
\end{aligned}$$

Thus, in a single experiment, we can get a measurement of $\widetilde{\mathcal{R}}$ by simply dividing the number of events produced by the period. Averaging these measured values over N_t experiments yields:

$$\begin{aligned}
\overline{\mathcal{R}} &= \frac{\sum_{k=1}^{N_t} t_k\widehat{\mathcal{R}}_k}{\sum_{k=1}^{N_t} t_k} \\
&= \frac{\sum_{k=1}^{N_t} t_k\left(\widehat{N}_k/t_k\right)}{\sum_{k=1}^{N_t} t_k} \\
&= \frac{1}{\mathrm{T}}\sum_{k=1}^{N_t} N_k &(1.7)
\end{aligned}$$

where t_k and $\widehat{N}_k$ are the duration and number of triggers produced in the k^{th} experiment, and $\mathrm{T} = \sum_{k=1}^{N_t} t_k$. If the duration of each trial has roughly the same duration, τ, then:

$$\begin{aligned}
\overline{\mathcal{R}} &= \frac{\sum_{k=1}^{\widehat{N}_t} N_k}{\tau N_t} &(1.8)\\
&= \frac{\overline{N}}{\tau} &(1.9)
\end{aligned}$$

Since $\overline{N}$ is the mean number of events produced by a Poisson process, the variance is simply $\overline{N}$. Thus, the error in $\overline{\mathcal{R}}$ is:

$$\delta\overline{\mathcal{R}} = \frac{1}{\tau}\sqrt{\frac{\overline{N}}{N_t}} \tag{1.10}$$

Once we have a measured value for $\mathcal{R}$ we can estimate the probability of getting k events from the source in any period t (assuming the source is stationary); it is:

$$P(k|\overline{\mathcal{R}}, t) = \frac{(\overline{\mathcal{R}}t)^k e^{-\overline{\mathcal{R}}t}}{k!} \tag{1.11}$$

with error:

$$\begin{aligned}
\delta P &= \left|\frac{\mathrm{d}P}{\mathrm{d}\overline{\mathcal{R}}}\right| \delta\overline{\mathcal{R}} \\
&= \frac{t}{\tau}\sqrt{\frac{\overline{N}}{N_t}} e^{-\overline{N}t/\tau} \left| \frac{\left(\overline{N}t/\tau\right)^{k-1}}{(k-1)!} - \frac{\left(\overline{N}t/\tau\right)^{k}}{k!} \right| \qquad (1.12)
\end{aligned}$$

In particular, we will be interested in the probability of getting one *or more* events. This is:

$$\begin{aligned}
P(k \geq 1|\overline{\mathcal{R}}, t) &= 1 - P(0|\overline{\mathcal{R}}, t) \\
&= 1 - e^{-\overline{\mathcal{R}}t} \qquad (1.13)
\end{aligned}$$

If the source is non-stationary, then an exact value of $\widetilde{\mathcal{R}}(t)$ does not exist. This is due to the fact that it is impossible to distinguish between fluctuations in the number of triggers produced by a random process from statistical variation and flutuations due to a changing rate density. Since we can only do a finite number of experiments, in order to get a better measurement of $\mathcal{R}$, we must observe it for a longer period of time. However, this results in a worse measurement of the time dependence of $\widetilde{\mathcal{R}}$, since any fluctuations that occur on a time scale smaller than the period of time we observe for will be indescernible.

In this analysis, we will assume that the sources we observe are roughly stationary over the periods we observe them for. Although the interferometers do change over time, we leave it to other indicators, such as environmental and instrumental monitors, to inform us when these changes happen so that we can adjust our observation time accordingly.

1.2 Modelling the Expected Number of Coincident Triggers

Consider a coincident event produced by a network of detectors in an experiment with *combined* signal-to-noise ratio (SNR) $\rho^\dagger$. We wish to know the significance of the event; that is, we want to know the probability that the event was created by a specific source. To determine that probability we have two choices: we can compare the event to the distribution of the desired source's triggers or we can compare the event to a distribution of triggers from all other, *background*, sources, which gives the *false alarm probability*. Since we do not know *a priori* the distribution of gravitational-wave triggers in the IFOs, in compact binary coalesence (CBC) searches we aim to compute false alarm probabilities. We define the false alarm probibility, P_F, as being the probability of getting one or more triggers with a SNR $\geq \rho^\dagger$ from a background distribution of triggers in the time searched, $\mathrm{T_f}$ (where "f" is used for *foreground*). From equation 1.13 this is:

$$P_F(k \geq 1 | \mathcal{F}(\rho^\dagger), \mathrm{T_f}) = 1 - e^{-\mathcal{F}(\rho^\dagger)\mathrm{T_f}} \tag{1.14}$$

$\mathcal{F}(\rho^\dagger)$ is the false alarm rate of a trigger with SNR $\rho^\dagger$. It is the rate density of all background coincident triggers; i.e., it is the rate density of all coincident triggers occuring from every source except gravitational waves. Since we do not have an analytic model for the interferometers' noise sources we must measure $\mathcal{F}(\rho^\dagger)$ directly. To see how this is done, we first consider the true rate density of coincident triggers from *all* sources in the IFOs (gravitational waves included), $\widetilde{\mathcal{R}}_{\mathrm{all}}(\rho^\dagger)$. We then model the coincidence algorithm to derive this value. For simplicity, we only consider a single template; multiple templates are discussed in section **??**.

Assuming stationary sources, in a single experiment of duration $\mathrm{T_f}$, $\widetilde{\mathcal{R}}_{\mathrm{all}}(\rho^\dagger)$ is given by:

$$\widetilde{\mathcal{R}}_{\mathrm{all}}(\rho^\dagger) = \frac{\mathbb{E}\left(N_{\mathrm{uncorr}}(\rho^{\dagger 2} \leq \sum_i \rho_i^2), \mathrm{T_f}\right) + \mathbb{E}\left(N_{\mathrm{corr}}(\rho^{\dagger 2} \leq \sum_i \rho_i^2), \mathrm{T_f}\right)}{\mathrm{T_f}} \tag{1.15}$$

Here, $\mathbb{E}\left(N_{\mathrm{uncorr}}(\rho^{\dagger 2} \leq \sum_i \rho_i^2, \mathrm{T_f})\right)$ is the expected number of triggers from all *uncorrelated* sources that give a combined ρ^2 greater than or equal to $\rho^{\dagger 2}$ in time $\mathrm{T_f}$. Uncorrelated means that a source that causes a trigger in one detector has no effect on the others. Conversely, $\mathbb{E}\left(N_{\mathrm{corr}}(\rho^{\dagger 2} \leq \sum_i \rho_i^2, \mathrm{T_f})\right)$ is the expected number of *correlated* coincident triggers, which means they come from a source that causes triggers all of the detectors.

If we have N_d detectors and N_s independent trigger sources in each detector, then the expected number of *uncorrelated* sources is given by:

$$\begin{aligned}
&\mathbb{E}\left(N_{\mathrm{uncorr}}(\rho^{\dagger 2} \leq \sum_i \rho_i^2, \mathrm{T_f})\right) \\
&= \int\limits_{\substack{\rho^{\dagger 2} \leq \sum_i \rho_i^2, \\ \rho_i \geq a\ \forall i}}^{\infty} \int\limits_{\substack{|t_i - t_l| \leq \mathfrak{T}_{i,l} \\ \forall l \neq i}}^{\mathrm{T_f}} \prod_i^{N_d} \sum_j^{N_s} \mathbb{E}\left(n_{i,j}(\rho_i, t_i)\right)\, \mathrm{d}t_i\, \mathrm{d}\rho_i \qquad (1.16) \\
&= \int_{\mathcal{S}(\rho^\dagger)} \int_{\mathcal{T}(\mathrm{T_f})} \prod_i^{N_d} \sum_j^{N_s} \sum_{k=1}^{\infty} k\, P_{i,j}(k|\rho_i, t_i)\, \mathrm{d}t_i\, \mathrm{d}\rho_i \qquad (1.17)
\end{aligned}$$

$\mathbb{E}\left(n_{i,j}(\rho_i, t_i)\right)$ in equation 1.16 and $P_{i,j}(k|\rho_i, t_i)$ in equation 1.17 are, respectively, the expected number of triggers, and the probability of getting k triggers, from the j^{th} source in the i^{th} detector with SNR ρ_i at time t_i. (In going from equation 1.16 to equation 1.17 we have used the fact that the expected number of triggers produced by a source with probability distribution function (PDF) $P(k)$ is $\sum_{k=1}^{\infty} k\, P(k)$.)

The regions of integration over ρ and t ($\mathcal{S}$ and $\mathcal{T}$, respectively) for a two-detector network are shown in Figure **??**. The SNR integral is carried out such that the quadruture sum of the single-IFO SNRs are $\geq \rho^\dagger$, with a lower-cutoff at a. This lower cut-off is the SNR cut that we impose in the matched-filter search; typically $a = 5.5$. The time integral is carried out such that for each point in time in a given detector, t_i, we only integrate between $t_i \pm \mathfrak{T}_{i,l}$ in every other detector. $\mathfrak{T}_{i,l}$ is the duration of the coincidence window between the i^{th} and the l^{th} detector for the template we are considering. As dicussed in chapter **??** it is an SNR-dependent quantity. However, since it is mostly dominated by the light-travel time between the i^{th} and l^{th} IFO, here we approximate it to be a constant for each (i, l) pair of detectors.

$P_{i,j}(k|\rho_i, t_i)$ is a two-dimensional PDF: it has some distribution in time and some distribution in ρ. If we assume the sources' time and ρ dependence are independent of each other, then:

$$P_{i,j}(k|\rho_i, t_i) = P(k|\widetilde{r}_{i,j}, t_i) P_{i,j}(\rho_i) \qquad (1.18)$$

In words: the probability of getting k triggers from the j^{th} source in the i^{th} detector with SNR ρ_i in time t_i is the probability of getting k triggers from a source with rate-desnity $\widetilde{r}_{i,j}$ in time t_i times the probability of getting a trigger with SNR ρ_i from the source. If we again assume that all of the sources are stationary Poisson processes in the time domain, then:

$$\sum_{k=1}^{\infty} kP(k|\widetilde{r}_{i,j}, t_i)\,\mathrm{d}t_i = \widetilde{r}_{i,j}\,\mathrm{d}t_i \tag{1.19}$$

With this assumption, the time integral can be carried out independent of the integral over ρ_i; it gives the volume bounded by a hyper-surface, which we designate $\mathcal{V}$:

$$\mathcal{V}(\mathrm{T_f}) \equiv \int\limits_{\substack{|t_i - t_l| \leq \mathfrak{T}_{i,l} \\ \forall l \neq i}}^{\mathrm{T_f}} \prod_{i}^{N_d} \mathrm{d}t_i \tag{1.20}$$

$\mathcal{V}$ has units of $[\mathrm{time}]^{N_d}$. Plugging this into equation 1.16 we have:

$$\mathbb{E}\left(N_{\mathrm{uncorr}}(\rho^{\dagger 2} \leq \sum_i \rho_i^2, \mathrm{T_f})\right) = \mathcal{V}(\mathrm{T_f}) \int_{\sum_i \rho_i^2 \geq \rho^{\dagger 2}} \prod_{i}^{N_d} \sum_{j}^{N_s} \widetilde{r}_{i,j} P_{i,j}(\rho_i)\,\mathrm{d}\rho_i \tag{1.21}$$

To estimate the expected number of coincident triggers with combined SNR $\geq \rho^\dagger$ from *correlated* sources, we assume that the sources create a trigger with the same ρ at the same time (modulo the light-travel time) in all detectors. In this case, the expected number is the quadruture sum over the individual probabilities:

$$\begin{aligned}
&\mathbb{E}\left(N_{\mathrm{corr}}(\rho^{\dagger 2} \leq \sum_i \rho_i^2, \mathrm{T_f})\right) \\
&= \sum_{j=1}^{N_s} \sqrt{\sum_{i=1}^{N_d} \left(\int_{\rho^\dagger/\sqrt{N_d}}^{\infty} \int_0^{\mathrm{T_f}} \mathbb{E}\left(n_{i,j}(\rho_i, t_i)\right)\,\mathrm{d}t_i\,\mathrm{d}\rho_i \right)^2} \qquad (1.22) \\
&= \mathrm{T_f} \sum_{j=1}^{N_s} \sqrt{\sum_{i=1}^{N_d} \left(\widetilde{r}_{i,j} \int_{\rho^\dagger/\sqrt{N_d}}^{\infty} P_{i,j}(\rho_i)\,\mathrm{d}\rho_i \right)^2} \qquad (1.23)
\end{aligned}$$

In going from 1.22 to 1.23 we have used equation 1.18 and have assumed the correlated sources are stationary. Since all triggers will occur in each detector within the light-travel time between them the constraints are removed from the time integral, making it simply $\mathrm{T_f}$. Likewise, as we have assumed that each source will create triggers with the same ρ in all detectors, there are no constraints on the SNR integral and the single-IFO SNR is reduced to $\rho^\dagger/\sqrt{N_d}$.

We assume that the only source that can create correlated triggers across detectors is from gravitational waves. Thus, the sum over the number of sources in equation 1.23 becomes a single term, with $j = \mathrm{GW}$. This is a bit of an oversimplfication: due to the antenna patterns of the detectors and their varying sensitivies, a gravitational wave will only generate a trigger in all detectors with the same SNR for certain sky locations and orientations. However, since we are interested in computing background rates in this analysis and not gravitational wave (GW) rates, we will assume that the detectors are roughly co-located with similar sensitivity. Under this assumption, the $P_{i,\mathrm{GW}}$ is the same for all detectors; thus the integral can be pulled out of the sum, and we have:

$$\mathbb{E}\left(N_{\mathrm{GW}}(\rho^{\dagger 2} \leq \sum_i \rho_i^2, \mathrm{T_f})\right) = \mathrm{T_f}\widetilde{r}_{\mathrm{GW}}(\vec{\theta})\sqrt{N_d}\int_{\rho^\dagger/\sqrt{N_d}}^{\infty} P_{\mathrm{GW}}(\rho)\,\mathrm{d}\rho \tag{1.24}$$

Note that $\widetilde{r}_{i,j}$ has been replaced by $\widetilde{r}_{GW}(\vec{\theta})$. This is the rate of CBCs in the universe; as it depends on the parameters of the source, $\vec{\theta}$, we have made the parameter dependence explicit, even though we are still only considering a single template.

The integral in equation 1.24 gives the sensitivity of the detectors to GW sources. To compute it, we need the dependence of the sensitivy as a function of ρ. This can be obtained as follows: we wish to know the number of sources the detector is sensitive to from here to some distance $D^\dagger$. If we assume that binary sources are distributed uniformly throughout space (which is a valid assumption for distances greater than $\sim 10\,\mathrm{Mpc}$ and less than $\sim 1\,\mathrm{Gpc}$ [**?**]), then this is:

$$(2.26)^{-3}\int_0^{D^\dagger} D^2\,\mathrm{d}D\int \mathrm{d}\Omega \tag{1.25}$$

where $\mathrm{d}\Omega$ is the solid angle. The factor $(2.26)^{-3}$ comes from approximating the volume enclosded by the detector's antenna pattern (which is peanut shaped) by a sphere with a radius equal to the detector's "horizon distance". The "horizon distance" is the distance to a binary with optimal orientation and location (i.e., it is the length of the longest part of the peanut) [?]. As discussed in chapter **??**, the sensitive distance is related to SNR by:

$$D = \frac{\sigma_{\rm GW}}{\rho} \tag{1.26}$$

(The need for the "GW" subscript will become clear later.) Plugging this into the above, and setting $\rho^\dagger/\sqrt{N_d} = \sigma_{\rm GW}/D^\dagger$ we have:

$$(2.26)^{-3}\int_0^D D'\,\mathrm{d}D' \int d\Omega = \frac{4\pi}{(2.26)^3}\int_{\rho^\dagger/\sqrt{N_d}}^{\infty} \frac{\sigma_{\rm GW}^3\,\mathrm{d}\rho}{\rho^4} \tag{1.27}$$

Thus:

$$P_{\rm GW}(\rho) = \frac{4\pi\sigma_{\rm GW}^3}{(2.26)^3}\rho^{-4} \tag{1.28}$$

which gives:

$$\mathbb{E}\left(N_{\rm GW}(\rho^{\dagger 2} \leq \sum_i \rho_i^2, \mathrm{T_f})\right) = \frac{4}{3}\pi\frac{\sqrt{N_d}\,\sigma_{\rm GW}^3}{(2.26)^3(\rho^\dagger/\sqrt{N_d})^3}\widetilde{r}_{\rm GW}(\theta)\mathrm{T_f} \tag{1.29}$$

(Note that we only have one $\sigma_{\rm GW}^3$ in the equation. This is because we assumed that all the detectors had the same sensitivity. If we had not made this assumption, the $\sqrt{N_d}\,\sigma_{\rm GW}^3$ term would instead be $\sqrt{\sum_i^{N_d}\sigma_{i,\rm GW}^6}$.) Adding this term to equation 1.21 and dividing by $\mathrm{T_f}$ gives the overall rate-density $\widetilde{\mathcal{R}}_{\rm all}(\rho^\dagger)$.

As an example, consider two arbitrary detectors, H and L, that are roughly co-located and have similar sensitivies. Let us assume that each IFO has only one non-gravitational wave source, which we call "noise", and that this noise is Gaussian distributed in ρ with some constant rate-density in time:

$$\mathbb{E}\left(n_{i,\rm noise}(\rho_i, t_i)\right)\,\mathrm{d}t_i\,\mathrm{d}\rho_i = \frac{\widetilde{r}_{i,\rm noise}}{\sqrt{2\pi}\sigma_{i,\rm noise}}e^{-\rho_i^2/2\sigma_{i,\rm noise}^2}\,\mathrm{d}t_i\,\mathrm{d}\rho_i \tag{1.30}$$

Given the large tail in the SNR distribution, as seen in chapter **??**, it may seem absurd to make this assumption. However, if we use *New SNR* for our ranking statistic (which we still label ρ) instead of *SNR*, then, as seen in Figure **??** in Chapter **??**, this is not such a bad assumption, especially if the template we are considering is from a binary neutron star. The caveat is that the $\sigma_{i,\rm noise}$ in the above equation is not the same as the $\sigma_{\rm GW}$ in equations 1.26–1.29. Although $\sigma_{\rm GW}$ is also the variance of the matched filter in Gaussian noise, that noise is for an *idealized* detector that has a stationary Gaussian distribution in SNR. The noise we are considering here comes from the actual detectors; as such, $\sigma_{i,\mathrm{noise}}$ can only be determined by fitting a Gaussian to the New SNR distribution of the detector data. How, then, do we determine σ_{GW}? Contrary to noise, we will assume that gravitational waves match the template well, resulting in $\chi^2 \approx 1$. In this case, New SNR reduces to SNR, and so we can use equation 1.26 to relate D to ρ. One other caveat is that we currently do not use a New SNR cut as we do in SNR. Due to the cut in SNR, the distribution in Figure **??** ceases to be Gaussian and falls off at low New SNR. For simplicity, we assume the same cut in New SNR here, and set $a = 5.5$.

With these assumptions in mind, the rate density of coincident triggers from all sources would be:

$$\begin{aligned}
\widetilde{\mathcal{R}}_{\mathrm{all}}(\rho^\dagger) \quad = & \\
& \frac{\mathfrak{T}_{\mathrm{HL}}(2\mathrm{T_f} - \mathfrak{T}_{\mathrm{HL}})}{\mathrm{T_f}} \frac{\widetilde{r}_{\mathrm{H,noise}}\widetilde{r}_{\mathrm{L,noise}}}{2\pi\sigma_{\mathrm{H,noise}}\sigma_{\mathrm{L,noise}}} \int_{\mathcal{S}} e^{-(\rho_{\mathrm{H}}^2/2\sigma_{\mathrm{H,noise}}^2 + \rho_{\mathrm{L}}^2/2\sigma_{\mathrm{L,noise}}^2)}\,\mathrm{d}\rho_{\mathrm{H}}\,\mathrm{d}\rho_{\mathrm{L}} \\
+ \quad & \frac{4}{3}\pi\frac{\sqrt{2}\sigma_{\mathrm{GW}}^2}{(2.26)^3(\rho^\dagger/\sqrt{2})^3}\widetilde{r}_{\mathrm{GW}}(\vec{\theta})
\end{aligned} \tag{1.31}$$

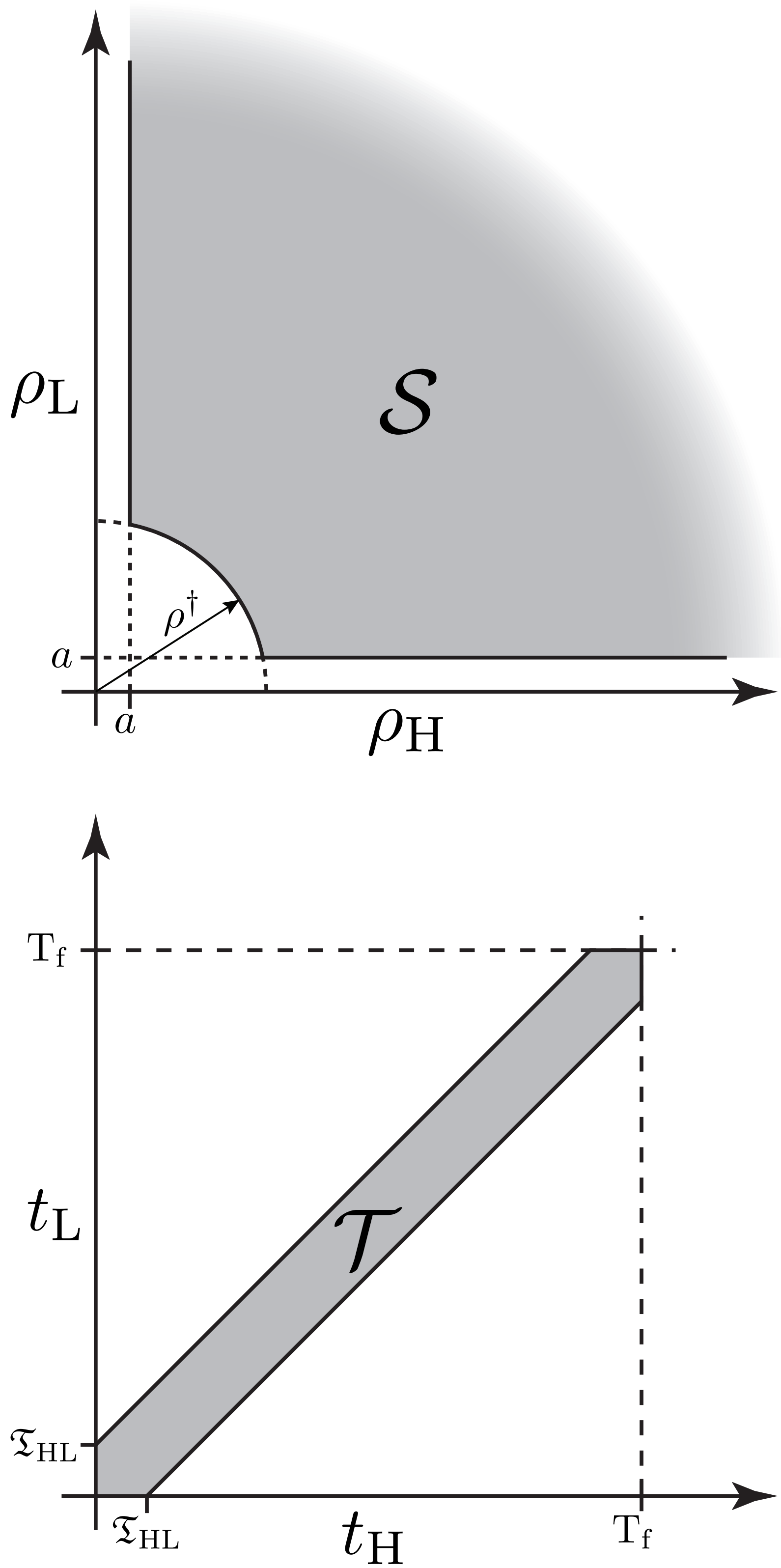

Figure 1 : The regions of integration in ρ, and time for a network of two detectors, H and L.