1 The Likelihood Formalism

This section describes the mathematical likelihood formalism used in Skylab. First it introduces the log-likelihood approach, second the likelihood ratio test and the used test statistic and then describes the used optimizations.

1.1 The Log-Likelihood Approach

Skylab implements the two-component likelihood approach with a likelihood function $\mathcal{L}(n_\mathrm{s}, \vec{p}_\mathrm{s} \,|D)$ of the form

$$\mathcal{L}(n_\mathrm{s}, \vec{p}_\mathrm{s} \,|D) = \prod_{i=1}^{N} \left[\frac{n_\mathrm{s}}{N} \mathcal{S}_i(\vec{p}_\mathrm{s}) + (1 - \frac{n_\mathrm{s}}{N}) \mathcal{B}_i \right], \tag{1}$$

where n_s is the number of signal events, hence, $(N - n_\mathrm{s})$ the number of background events in the dataset D of N total events. The set of signal model parameters is denoted as $\vec{p}_\mathrm{s}$. For a point-like source model, the signal model parameters include the source position $\vec{x}_\mathrm{s}$ and the spectral index γ of the source flux. $\mathcal{S}_i(\vec{p}_\mathrm{s})$ and $\mathcal{B}_i$ is the value of the signal and background PDF for the ith data event, respectively.

The signal and background PDFs must incorperate the detector efficiency (yield), $\mathcal{Y}_i$, which, in general, depends on the celestrial direction, the energy, and the observation time of the data event.

For computational stability reasons the logarithm of the likelihood function of equation 1 is used in Skylab:

$$\log \mathcal{L}(n_\mathrm{s}, \vec{p}_\mathrm{s} \,|D) = \sum_{i=1}^{N} \log(...) \tag{2}$$

1.2 Likelihood Ratio Test and Test Statistic

For estimating the significance of an observation, the likelihood ratio Λ with respect to a null hypothesis of no observation, i.e. equation 1 at $n_\mathrm{s} = 0$ is tested:

$$\log \Lambda(n_\mathrm{s}, \vec{p}_\mathrm{s}) = \log \frac{L(n_\mathrm{s}, \vec{p}_\mathrm{s})}{L(n_\mathrm{s} = 0)} = \sum_{i=1}^{N} \log \left[1 + \frac{n_\mathrm{s}}{N} \left(\frac{\mathcal{S}_i(\vec{p}_\mathrm{s})}{\mathcal{B}_i} - 1 \right) \right] \tag{3}$$

By defining

$$\mathcal{X}_i(\vec{p}_\mathrm{s}) \equiv \frac{1}{N} \left(\frac{\mathcal{S}_i(\vec{p}_\mathrm{s})}{\mathcal{B}_i} - 1 \right), \tag{4}$$

this reads as:

$$\log \Lambda(n_\mathrm{s}, \vec{p}_\mathrm{s}) = \sum_{i=1}^{N} \log(1 + n_\mathrm{s} \mathcal{X}_i(\vec{p}_\mathrm{s})) \tag{5}$$

This leads to the test statistic TS

$$\mathrm{TS} = 2\mathrm{sgn}(n_\mathrm{s}) \log \Lambda(n_\mathrm{s}, \vec{p}_\mathrm{s}) \tag{6}$$

with separation of over- ($n_\mathrm{s} > 0$) and under-fluctuations ($n_\mathrm{s} < 0$).

1.3 Optimizations for Point-Like Sources

For point-source like signal hypotheses most of the events in the data sample will be far away from the hypothesised source, hence, the value of the signal PDF $\mathcal{S}_i$ will be zero or very close to zero. By selecting only the signal-contributing N' events from the sample, the likelihood ratio $\log \Lambda$ reads

$$\log \Lambda(n_\mathrm{s}, \vec{p}_\mathrm{s}) = \sum_{i=1}^{N'} \log(1 + n_\mathrm{s} \mathcal{X}_i(\vec{p}_\mathrm{s})) + (N - N') \log(1 - \frac{n_\mathrm{s}}{N}) \tag{7}$$

1.4 Signal & Background PDFs

The likelihood ratio function as given in equation (3) incorperates signal, $\mathcal{S}_i$, and background, $\mathcal{B}_i$, probability density functions (PDFs). Both PDFs can be factorized into a spatial (S_i), an energy ($\mathcal{E}_i$), and a time ($\mathcal{T}_i$) component.

The signal PDF can be written as

$$\mathcal{S}_i(\vec{p}_\mathrm{s}) = S_{\mathcal{S},i}(\vec{x}_i|\vec{p}_\mathrm{s,spatial}) \mathcal{E}_{\mathcal{S},i}(E_i|\vec{p}_\mathrm{s,energy}) \mathcal{T}_{\mathcal{S},i}(t_i|\vec{p}_\mathrm{s,time}), \tag{8}$$

where the signal model parameters $\vec{p}_\mathrm{s}$ can be divided into spatial, energy, and time parameters, i.e. $\vec{p}_\mathrm{s} = (\vec{p}_\mathrm{s,spatial}, \vec{p}_\mathrm{s,energy}, \vec{p}_\mathrm{s,time})$. The spatial component, S_i, can be identified as the point-spread-function (PSF) of the detector. For a point-like source model a convinient and widely used PSF is the gaussian function:

$$S_{\mathcal{S},i}(\vec{x}_i|\vec{p}_\mathrm{s,spatial}) \equiv S_i(r_i, \sigma_i|\vec{x}_\mathrm{s}) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{r_i^2}{2\sigma_i^2}\right), \tag{9}$$

where r_i is the space angle between the source position and the recorded reconstructed event direction. In equatorial coordinates, $\vec{x} = (\alpha, \delta)$, the cosine of r_i is given by

$$\cos(r_i) = \cos(\alpha_\mathrm{s} - \alpha_i) \cos(\delta_\mathrm{s}) \cos(\delta_i) + \sin(\delta_\mathrm{s}) \sin(\delta_i). \tag{10}$$

The data quantity σ_i describes the angular reconstruction uncertainty of the event, hence the PSF is narrower for well-reconstructed events, and wider for events which have a large reconstruction uncertainty.

When considering a power law as source flux model, the energy source parameters, $\vec{p}_\mathrm{s,energy}$, consists of the spectral index γ and possibly an energy cut-off parameter E_cut.

In analog to the signal PDF, the background PDF can be formulated as

$$\mathcal{B}_i = S_{\mathcal{B},i}(\vec{x}_i) \mathcal{E}_{\mathcal{B},i}(E_i) \mathcal{T}_{\mathcal{B},i}(t_i). \tag{11}$$

All the background PDF components can either be determined from the data itself or by using monte-carlo simulation.

1.5 Gradients of the Log-Likelihood Ratio

For maximizing the log-likelihood ratio function (equation 5), or minimizing the negative of it, the minimizer algorithm requires its partial derivatives of the fit parameters, $\vec{p}_{\rm s}$[1]. Hence, here we provide the expressions of the partial derivatives for equation 7. The partial derivative for $n_{\rm s}$ is given by

$$\frac{\partial \log\Lambda(n_{\rm s},\vec{p}_{\rm s})}{\partial n_{\rm s}} = \sum_{i=1}^{N'} \frac{\mathcal{X}_i(\vec{p}_{\rm s})}{1+n_{\rm s}\mathcal{X}_i(\vec{p}_{\rm s})} - \frac{N-N'}{N-n_{\rm s}}. \tag{12}$$

The partial derivative for an individual signal parameter, $p_{\rm s}$, is given by

$$\frac{\partial \log\Lambda(n_{\rm s},\vec{p}_{\rm s})}{\partial p_{\rm s}} = \sum_{i=1}^{N'} \frac{n_{\rm s}}{1+n_{\rm s}\mathcal{X}_i(\vec{p}_{\rm s})}\frac{\partial \mathcal{X}_i(\vec{p}_{\rm s})}{\partial p_{\rm s}}. \tag{13}$$

The partial derivative of $\mathcal{X}_i$ can be calculated using equation 4 and the expressions for the signal and background PDFs as given in equation 8 and 11, respectively. Depending on the type of fit parameter, i.e. spatial, energy, or time, the derivative of the PDF ratio, $\mathcal{R}_i(\vec{p}_{\rm s}) = \mathcal{S}_i(\vec{p}_{\rm s})/\mathcal{B}_i$, simplifies to the derivative of the respective type of PDF ratio:

$$\frac{\partial \mathcal{X}_i(\vec{p}_{\rm s})}{\partial p_{\rm s}} = \frac{1}{N}\frac{\partial \mathcal{R}_i(\vec{p}_{\rm s})}{\partial p_{\rm s}}, \tag{14}$$

with

$$\mathcal{R}_i(\vec{p}_{\rm s}) = \mathcal{R}_{S,i}(\vec{p}_{s,\rm spatial})\mathcal{R}_{\mathcal{E},i}(\vec{p}_{s,\rm energy})\mathcal{R}_{\mathcal{T},i}(\vec{p}_{s,\rm time}), \tag{15}$$

and

$$\frac{\partial \mathcal{R}_i(\vec{p}_{\rm s})}{\partial p_{\rm s,spatial}} = \frac{\partial \mathcal{R}_{S,i}(\vec{p}_{s,\rm spatial})}{\partial p_{\rm s,spatial}}\mathcal{R}_{\mathcal{E},i}(\vec{p}_{s,\rm energy})\mathcal{R}_{\mathcal{T},i}(\vec{p}_{s,\rm time}), \tag{16}$$

$$\frac{\partial \mathcal{R}_i(\vec{p}_{\rm s})}{\partial p_{\rm s,energy}} = \mathcal{R}_{S,i}(\vec{p}_{s,\rm spatial})\frac{\mathcal{R}_{\mathcal{E},i}(\vec{p}_{s,\rm energy})}{\partial p_{\rm s,energy}}\mathcal{R}_{\mathcal{T},i}(\vec{p}_{s,\rm time}), \tag{17}$$

$$\frac{\partial \mathcal{R}_i(\vec{p}_{\rm s})}{\partial p_{\rm s,time}} = \mathcal{R}_{S,i}(\vec{p}_{s,\rm spatial})\mathcal{R}_{\mathcal{E},i}(\vec{p}_{s,\rm energy})\frac{\mathcal{R}_{\mathcal{T},i}(\vec{p}_{s,\rm time})}{\partial p_{\rm s,time}}. \tag{18}$$

1.6 Multiple Datasets

With Skylab a set of J different data samples (datasets) D_j can be analyzed at once. Each data sample has its own detector signal efficiency $\mathcal{Y}_{{\rm s},j}$.

The composite likelihood function is the product of the individual dataset likelihood functions:

$$\log\Lambda = \sum_{j=1}^{J} \log\Lambda_j \tag{19}$$

[1] In this formalism we assume that $\vec{p}_{\rm s}$ consists solely of fit parameters for the signal, but in practice it can also consist of fixed signal parameters.

The total number of signal events n_s needs to get split-up into $n_{\mathrm{s},j}$ for the individual data samples. The distribution of n_s along the different data samples is based on the detector signal efficiency $\mathcal{Y}_{\mathrm{s},j}$ of each sample:

$$n_{s,j}(\vec{p}_\mathrm{s}) = n_\mathrm{s} \frac{\mathcal{Y}_{\mathrm{s},j}(\vec{p}_\mathrm{s})}{\sum_{j=1}^{J} \mathcal{Y}_{\mathrm{s},j}(\vec{p}_\mathrm{s})}, \tag{20}$$

where $\vec{p}_\mathrm{s}$ denotes the source parameters as above and contains the source position(s) $\vec{x}_\mathrm{s}$ and possible additional source parameters like for instance the spectral index γ.

By defining the sample weight factor

$$f_j(\vec{p}_\mathrm{s}) \equiv \frac{\mathcal{Y}_{\mathrm{s},j}(\vec{p}_\mathrm{s})}{\sum_{j=1}^{J} \mathcal{Y}_{\mathrm{s},j}(\vec{p}_\mathrm{s})} \tag{21}$$

with the property

$$\sum_{j=1}^{J} f_j = 1 \tag{22}$$

equation 20 reads

$$n_{s,j}(\vec{p}_\mathrm{s}) = n_\mathrm{s} f_j(\vec{p}_\mathrm{s}) \tag{23}$$

The detector signal efficiency $\mathcal{Y}_{\mathrm{s},j}(\vec{p}_\mathrm{s})$ depends on the signal model parameters $\vec{p}_\mathrm{s}$, hence, on the source location(s) $\vec{x}_\mathrm{s}$ and the spectral index γ. It can be calculated via the detector effective area and the source flux.

For a single point source the sample weight factor can be calculated via the effective area $A_{\mathrm{eff},j}(E)|_{\vec{x}_\mathrm{s}}$ at the source location of each data sample, and the differential flux $\frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}$ of the source.

$$f_j(\vec{p}_\mathrm{s}) = \frac{\int_0^\infty \mathrm{d}E A_{\mathrm{eff},j}(E)|_{\vec{x}_\mathrm{s}} \frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E, \vec{p}_\mathrm{s})}{\sum_{i=1}^{J} \int_0^\infty \mathrm{d}E A_{\mathrm{eff},i}(E)|_{\vec{x}_\mathrm{s}} \frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E, \vec{p}_\mathrm{s})} \tag{24}$$

Using the sample weight factor $f_j(\vec{p}_\mathrm{s})$ the likelihood ratio of equation (19) with equation (5) can now be written as

$$\log \Lambda(n_\mathrm{s}, \vec{p}_\mathrm{s}) = \sum_{j=1}^{J} \sum_{i=1}^{N} \log(1 + n_\mathrm{s} f_j(\vec{p}_\mathrm{s}) \mathcal{X}_i(\vec{p}_\mathrm{s})) \tag{25}$$

For multiple point sources, i.e. a stacking of K point sources with positions $\vec{x}_{\mathrm{s},k}$, the sample weight factor of each single source needs to be taking into account. Thus, f_j can be written as the sum of the products of the sample weight factor $f_j(\vec{p}_{\mathrm{s},k})$ for source k and the relative strength $f_k(\vec{p}_{\mathrm{s},k})$ of the kth source in all data samples compared to all the other sources in all data samples.

$$f_j(\vec{p}_\mathrm{s}) = \sum_{k=1}^{K} f_j(\vec{p}_{\mathrm{s},k}) f_k(\vec{p}_{\mathrm{s},k}) \tag{26}$$

The relative strength of source k can be written as

$$f_k(\vec{p}_{\mathrm{s},k}) = \frac{\sum_{i=1}^{J} \int_0^\infty \mathrm{d}E A_{\mathrm{eff},i}(E)|_{\vec{x}_{\mathrm{s},k}} \frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E, \vec{p}_{\mathrm{s},k})}{\sum_{\kappa=1}^{K} \sum_{i=1}^{J} \int_0^\infty \mathrm{d}E A_{\mathrm{eff},i}(E)|_{\vec{x}_{\mathrm{s},\kappa}} \frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E, \vec{p}_{\mathrm{s},k})} \tag{27}$$

The combination of equation 24 with $\vec{p}_\mathrm{s} \equiv \vec{p}_{\mathrm{s},k}$, hence, $\vec{x}_\mathrm{s} \equiv \vec{x}_{\mathrm{s},k}$ and equation 27, leads to the final expression for f_j for multiple sources:

$$f_j(\vec{p}_\mathrm{s}) = \frac{\sum_{k=1}^{K} \int_0^\infty \mathrm{d}E A_{\mathrm{eff},j}(E)|_{\vec{x}_{\mathrm{s},k}} \frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E, \vec{p}_{\mathrm{s},k})}{\sum_{i=1}^{J} \sum_{k=1}^{K} \int_0^\infty \mathrm{d}E A_{\mathrm{eff},i}(E)|_{\vec{x}_{\mathrm{s},k}} \frac{\mathrm{d}\Phi_\mathrm{s}}{\mathrm{d}E}(E, \vec{p}_{\mathrm{s},k})} \tag{28}$$

One should note that by definition the numerator of equation (28) is one of the summands of the sum in the denumerator, i.e. for $i = j$.

2 Detector Signal Efficiency

The detector signal efficiency $\mathcal{Y}_{\mathrm{s},j}(\vec{p}_{\mathrm{s},k})$ of a data sample j for a source k is defined as the integral over the energy of the product of the detector effective area and the differential flux $\frac{\mathrm{d}\Phi}{\mathrm{d}E}$ of the source:

$$\mathcal{Y}_{\mathrm{s},j}(\vec{p}_{\mathrm{s},k}) \equiv \int_0^\infty \mathrm{d}E A_{\mathrm{eff},j}(E)|_{\vec{x}_{\mathrm{s},k}} \frac{\mathrm{d}\Phi}{\mathrm{d}E}(E, \vec{p}_{\mathrm{s},k}) T_{\mathrm{live},j} \tag{29}$$

It is the mean number of signal events per steradian expected from a source at position $\vec{x}_\mathrm{s}$ with source parameters $\vec{p}_\mathrm{s}$. In the most-general case, the source position $\vec{x}_\mathrm{s}$ consists of three quantities: right-ascention, declination, and observation time, i.e. $\vec{x}_\mathrm{s} = (\alpha_\mathrm{s}, \delta_\mathrm{s}, t_\mathrm{obs})$.

2.1 Effective Area

In Skylab the effective area $A_{\mathrm{eff},j}$ of a data sample j is not calculated separately in order to avoid binning effects. However, the effective area can be calculated using the monte-carlo weights `mcweight`[2] of the simulation events. The monte-carlo weights have the unit GeV cm^2 sr. Using the monte-carlo weight, $w_{m,j}$, of the mth event of data sample j the effective area is given by the sum over the event weights divided by the solid angle and the energy range ΔE of the summed selected events:

$$A_{\mathrm{eff},j}(E) = \frac{\sum_{m=1}^{M} w_{m,j}}{\Omega \Delta E} \tag{30}$$

[2]In IceCube known as "OneWeight", but which already includes the number of used MC files.

Table 1: IceCube specific detector signal efficiency implementation methods.

Name of Class	Description
I3FixedFluxDetSigEff	IceCube detector signal efficiency implementation method for a fixed flux model, which might contain flux parameters, but which are not fit in the likelihood maximization process. This implementation assumes that the detector effective area depends solely on the declination of the source. This method creates a spline function of given order for the logarithmic values of the $\sin(\delta)$-dependent detector signal efficiency. The constructor of this implementation method requires a $\sin(\delta)$ binning definition for the monte-carlo events and the order of the spline function.
I3PowerLawFluxDetSigEff	IceCube detector signal efficiency implementation method for a power law flux model, implemented by the `PowerLawFlux` class. This method creates a 2D spline function of given orders for the logarithmic values of the $\sin(\delta)$-dependent detector signal efficiency for a range of γ values. This implementation method supports multi-processing.

2.2 The DetectorSignalEfficiency Class

`DetectorSignalEfficiency` provides a detector signal efficiency class to compute the integral given in equation (29). The detector signal efficiency depends on the flux model and its source parameters, which might change during the likelihood maximization process. It is also dependent on the detector effective area, hence is detector dependent. Thus, `DetectorSignalEfficiency` must be provided with a detector signal efficiency implementation method derived from the `DetSigEffImplMethod` class.

Detector signal efficiency values can be retrieved via the call operator `__call__(src_pos src_params)`, which takes the celestrial position of the source and the additional source parameters as arguments.

2.2.1 The DetSigEffImplMethod Class

`DetSigEffImplMethod` is an abstract base class and defines the interface between the detector signal efficiency implementation method and the `DetectorSignalEfficiency` class.

Table 1 lists all available IceCube specific detector signal efficiency implementation methods and their description.

3 Implemented Log-Likelihood Models

This section describes the implemented log-likelihood models. [1]

References

[1] Jim Braun, Mike Baker, Jon Dumm, Chad Finley, Albrecht Karle, Teresa Montaruli. Time-Dependent Point Source Search Methods in High Energy Neutrino Astronomy. *Astropart.Phys.*, 33:175–181, 2010.