1 Classical station triggers

As mentioned in **??**, continously analyzing data sent to CDAS from each of the 1600 SD water tanks would quickly exceed the computational capabilites of Augers’ main servers. For this purpose, trace information is only collected from a station, once a nearby T3 event (c.f. **??**) has been detected. The formation of a T3 trigger is dependant on several T2, or station-level, triggers, which will be discussed in detail in this chapter. First, general comments about evaluation of trigger performances are given in **??**. Then the precise implementation of SD station level triggers, as well as their individual performance is given in **??**.

1.1 Performance evaluation

The performance of a trigger can be evaluated in many different ways. In the most general consideration, a confusion matrix holds information about the ability of a classifier to discern between different types, or classes, C. With the example at hand there exist two types of events one wishes to distinguish, a signal event C_1 in the form of an extensive air shower, versus background C_0. The confusion matrix thus becomes:

		Predicted C	
		C_1	C_0
True C	C_1	True positive (TP)	False negative (FN)
	C_0	False positive (TP)	True negative (TN)

From this, other potentially interesting variables can be derived. Of particular interest for the Auger observatory are the sensitivity and **F**alse **D**iscovery **R**ate (FDR). The former is the probability that a signal event will be classified correctly, i.e. an extensive air shower hits a water tank and correctly raises a T2 trigger. The sensitivity - or **T**rue **P**ositive **R**ate (TPR) - is the ratio of correctly identified signals over all signals.

$$\mathrm{TPR} = \sum_{\{\mathrm{TP}\}} \frac{1}{w}, \qquad \text{where } w = \mathrm{TP} + \mathrm{FN}, \tag{1.1}$$

$$\Leftrightarrow \qquad = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{1.2}$$

The TPR for a given sample heavily depends on the underlaying distribution of data. While the distribution of zenith angles is correctly modeled, the simulated showers at hand do not accurately match the energy flux seen in real data (compare **??** and **??**). For this purpose, the TPR for a given dataset is transformed to a conditional trigger efficiency ϵ by recalculating the weights w_i based on the primary energy E_i of a shower. First, the overrepresentation of higher energy traces is eliminated. Every prediction is weighted by a factor $b_{E_i}/b_{E_{\text{ref}}}$, where b_{E_i} is the number of traces in a bin corresponding to energy E_i (see **??**). Afterwards, a flux factor is multiplied to the weight. It is of the form E_i^{-3}/E_{ref}. Lastly, the thus calculated sum is normalized to the weight $W = \sum\limits_{\{\text{TP,FN}\}} \frac{1}{w_i}$ of the entire dataset.

The physical interpretation of ϵ for a classifier is as follows: (At station level) It represents the ratio of air showers that *are* detected over all air showers that *can* in principle be detected. It is a conditional probability dependant on the WCD receiving a particle from an air shower. An efficiency of $epsilon = 1$ thus implies that if at least one particle in the shower cascade hits a tank, the tank raises a T2 trigger.

The false discovery rate is a measure of how readily the triggers (wrongly) identify background events like stray cosmic muons as extensive air showers. It is imperative for any trigger algorithm operating in the SD to minimize this probability. Simply due to the number of operating stations in the field, a small increase in FDR drastically raises the amount of potential events and hence load on the central analysis server of the observatory.

$$\text{FDR} = \frac{\text{FP}}{\text{TP} + \text{FP}}. \tag{1.3}$$

In this work, the trigger rate on random-traces f_{Random} is used as an alias for the FDR. Of course, this is not completely accurate, as these to a small fraction contain signal from air showers. In any case, the distinction does not matter. Any trigger algorithm must have a sufficiently low trigger frequency on measured data - be it extensive air showers or background events - as to not overload CDAS readout capabilities.

Consequently a pseudo-score can be assigned to each classification algorithm in order to compare them. This score a is given by **??**.

$$a = \frac{\epsilon}{f_{\text{Random}}\ [\text{Hz}]} \tag{1.4}$$

The physical interpretation of this variable is not straight forward. If labels were known for random-trace datasets, a would be equivalent to the **S**ignal-to-**N**oise **R**atio (SNR) of the classifier. However, since f_{Random} and the T2 efficiency across all primary energies and zenith angles must be determined using two different sets of data, only a (positive) correlation between the two exists.

In conclusion, all algorithms should maximize a, i.e. boost the T2 trigger efficiency, while keeping the random-trace trigger rate as low as possible.

1.1.1 Lateral Trigger Probability (LTP)

Ultimately, all test statistics constructed by classical station triggers (ignoring MoPS, c.f. **??**) are correlated to the deposited charge S in the WCD. Because S is heavily influenced by the primary particle energy, zenith and distance to the shower core, as well as to a lesser extent by shower age and statistical fluctuations, it makes sense to parametrize the trigger efficiency $\epsilon(E, \theta, \mathrm{SPD})$ in terms of these observables.

From a heuristic consideration, it can immediately be concluded that large separations between station and shower axis affect efficiencies negatively, because the particle distribution function monotonically decreases with increasing r (compare **??**). Similarly, inclined showers with a large θ are more attenuated compared to vertical showers, as they have to traverse a larger atmospheric depth ($\propto \sec(\theta)$) before reaching the detector. Lastly, primaries with large E on average deposit higher S in the WCD due to unleashing bigger particle cascades. Consequently ϵ is positively correlated with E.

The functional form that can be obtained by evaluating trigger efficiencies for a given (slice of) E and θ is labelled the **L**ateral **T**rigger **P**robability (LTP). It will be one of the main comparison metrics, by which different trigger algorithms are compared in this work. For classical triggers, two methods to extract the LTP are presented here. This is to show that both yield comparable results, and the latter method is a fair estimator for the neural network LTPs discussed in **??**.

$\overline{\text{Off}}\underline{\text{line}}$ lateral trigger probability

$\overline{\text{Off}}\underline{\text{line}}$ can simulate the SD detector response given a preprocessed shower footprint as given by e.g. CORSIKA. As such, calculating the LTP for a given event condenses to counting the number of triggered and non-triggered stations at specific distances from the shower axis. If this is done for a large enough sample size of showers, one eliminates noise induced by shower-to-shower fluctuations and arrives at an independent estimator for the probability of a T2 trigger given a shower at a shower plane distance r, with energy E and zenith θ. As per **[abreu2011lateral]**, the closed form approximation of the LTP is given as

$$\mathrm{LTP}\,(r) = \begin{cases} \dfrac{1}{1 + \exp\left(\frac{r - R_0}{\sigma_R}\right)}, & r \leq R_0 \\ \dfrac{1}{2}\exp\left(C(r - R_0)\right), & r > R_0 \end{cases} \tag{1.5}$$

In **??**, R_0, σ_R and C are all fit constants that will in general depend on E and θ. Most importantly, R_0 marks the shower plane distance where $(R_0) = 0.5$. This is connected to a steepening of the rising flank in the efficiency curve. Whereas an exponential function with decay constant $C < 0$ describes data well for large r, a logistic function with scaling factor $1/\sigma_R$ must account for the asymptotic transition to full efficiency closer to the core.

It must be mentioned that the motivation behind this parametrization is data- and not physics driven. In particular, $\mathrm{LTP}(r)$ is not smooth in R_0 if the parameters C and σ_R are not finetuned as indicated in **??**.

$$\lim_{r\to R_0^+} \frac{\partial\,\mathrm{LTP}(r)}{\partial r} = \frac{C}{2} \overset{!}{=} \frac{1}{4\sigma_R} = \lim_{r\to R_0^-} \frac{\partial\,\mathrm{LTP}(r)}{\partial r} \tag{1.6}$$

In this work however, a different parametrization is used to estimate the T2 response of a station. The functional form of this adjusted trigger probability is a clipped logistic function, and given in **??**.

$$\mathrm{LTP}^*(r) = \min\left(1,\, \epsilon^*\left(1 - \frac{1}{1+e^{-\frac{r-R_0}{\sigma_R}}}\right)\right) \tag{1.7}$$

The reasoning for this choice is as follows:

- The original parametrization, $\mathrm{LTP}(r)$, eventually approaches 1. This hints to a problem. It is not a guarantee that some trigger algorithm will detect all extensive air showers. Espically neural network triggers might be sensitive to only a subset of showers. This is reflected in the latter form, $\mathrm{LTP}^*(r)$, by introducing an additional fit parameter, the pseudo-efficiency $0 \leq \epsilon^* < 2$. In the case of $\epsilon^* \geq 1$, the domain of the function is correctly mapped to $[0, 1]$.
- There exists an imbalance in training data. Due to the geometry of the SD array, more traces at smaller r are available. In an attempt to reduce possible biases resulting from low statistics at small SPD, the form is kept as simple as possible.
- The function is guaranteed to be continously differentiable in R_0. For values $\epsilon^* > 1$ this is replaced with a kink at $R^* = R_0 - \frac{\log\left((1-1/\epsilon^*)^{-1}-1\right)}{\sigma_R}$, where $\mathrm{LTP}^*(r)$ would exceed 1 if not for clipping. There exists some physical motivations for this however. Due to the phase transition, namely to full efficiency, at this point, discontinuities in the lateral trigger probability are allowed.

This approach only marginally takes into account shower-to-shower-fluctuations. Such statistic perturbations are responsible for a smearing of the (initially) hard transition from sub- to full efficiency. The parametrization used by the Pierre Auger collaboration takes this into consideration by design.The presented $\mathrm{LTP}^*(r)$ does - at least explicitly - not. As a result, one could expect a bias, where $\mathrm{LTP}^*(R^*)$ over- or underestimates the actual trigger probability. This is however not the case when examining the residuals of the performance fit that is done in **??**. A plot showcasing this is offered in **??**. Still, results at low energies ($\log(E\,/\,\mathrm{eV}) < 17.5$) must be taken with a hint of skepticism, as only very little data is available for such showers.

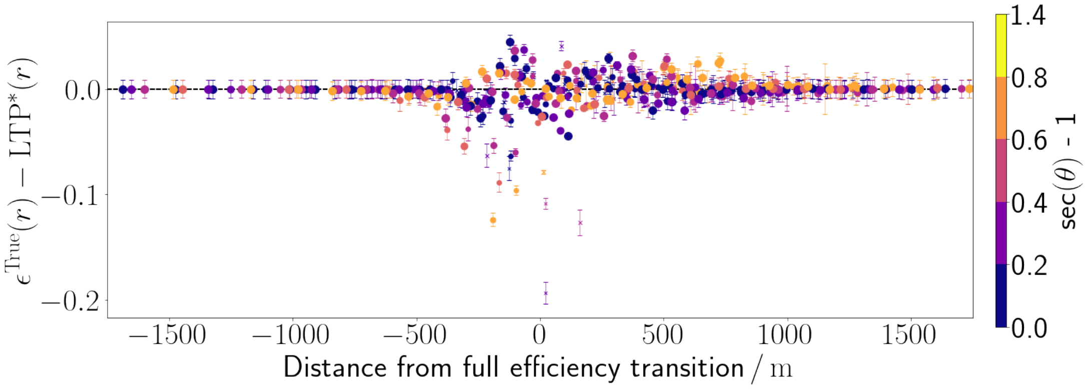

Figure 1.1: The residuals from comparing ϵ^{True} to LTP*(r) vanish at large (small) r. No systematic bias is observed in the transitional region around R^*. Large outliers are caused by the dataset being limited in size at low energies (marked with an x).

Bayesian folding

Given a time trace of the form in **??**, it is easy to determine whether or not a given trigger algorithm analyzing this trace would raise a T2. The trigger probability obtained this way must however not be confused with the lateral trigger probability discussed above. It is the conditional probability of a T2 given that a tank receives an air shower signal, $P(\mathrm{T2}\,|\,C_1)$. The complete LTP can be calculated in the following form:

$$\mathrm{LTP} := P(\mathrm{T2}) = P(\mathrm{T2}\,|\,C_1) * P(C_1). \tag{1.8}$$

$P(C_1)$ in **??** is the **L**ateral **P**article **P**robability (LPP) and quantifies the chance of a station receiving a signal from an extensive air shower. It is the probabilistic interpretation of the lateral distribution function (c.f. **??**). The LTP of an ideal classifier that is able to identify individual particles would be equal to the LPP.

For this work, the LPP is calculated by comparing the simulated stations in **??** to a catalog of known stations in the vicinity of the shower core. The ratio of simulated stations divided by all stations at a specific SPD is the LPP at a given (slice of) energy E and zenith θ. A function like in **??** is fitted to the data to extrapolate values at arbitrary SPD. The result of this analysis for all energies and zenith angles is shown in **??**. The best fit parameters are listed in **??**.

The comparison between the LTP gathered via Bayesian folding is compared to the $\overline{\text{Off}}\underline{\text{line}}$ station counting approach in **??**. This is done for every classical trigger individually. It is found that no remarkable difference between both results exists. As a consequence, this method of evaluating the LTP allows for an easy comparison of neural network triggers to the algorithms discussed in the following.

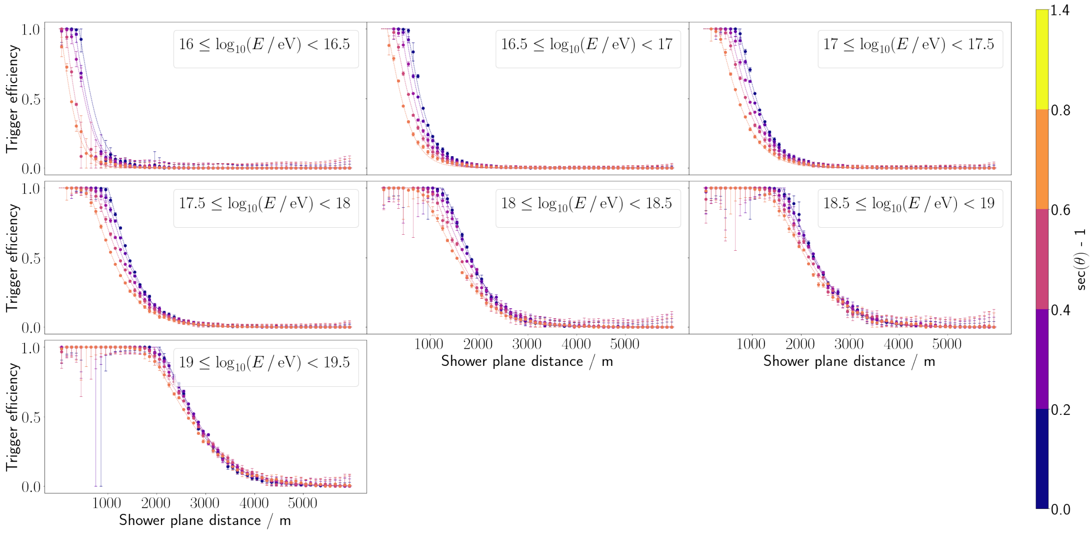

Figure 1.2: The probability of a station receiving a particle in relation to its' separation from the shower core axis. The probability at a constant SPD increases with increasing primary energy, but decreases with increasing zenith angle.

1.1.2 Calculation of T3 efficiency

When transitioning from station- to event-level, the important variable becomes the T3 efficiency. It states the probability of a shower not only being detected in individual stations, but also identified as such by the CDAS T3 triggers. In the end, the only measured data that is available comes from extensive air showers which passed this final hurdle.

With the lateral trigger probability at hand, a simple Monte-Carlo simulation can recover the T3 efficiency. Recall that the probability of a single station i detecting a shower at distance r_i, with energy E and arrival direction (θ, ϕ) is given by **??**.

$$\mathrm{LTP}^*(r_i) = \min\left(1,\, \epsilon^*(E,\theta)\left(1 - \frac{1}{1 + e^{-\frac{r_i - R_0(E,\theta)}{\sigma_R(E,\theta)}}}\right)\right). \tag{1.9}$$

In the simplest case, this results in a T3 trigger if the three closest stations raise a T2 within 11 µs of one another. By simulating a random core position and determining the detector response, one can calculate the event detection numerically. The approach presented here does not take into account timing differences resulting from the finite propagation speed of the shower front. Instead, it is assumed that all stations simultaneously receive a signal. This is only accurate if the primary particle initiated a perfectly vertical cascade infinitely far away from the SD. However, due to the permissivity of the T3 triggers such considerations need not be accounted for. Even in the most suboptimal case of a horizontal shower ($\theta = 90°$), the furthest station in the triangle receives a signal latest around $\frac{1.5\,\mathrm{km}}{c} = 5\,\mu\mathrm{s}$ after the closest one. Exemplary

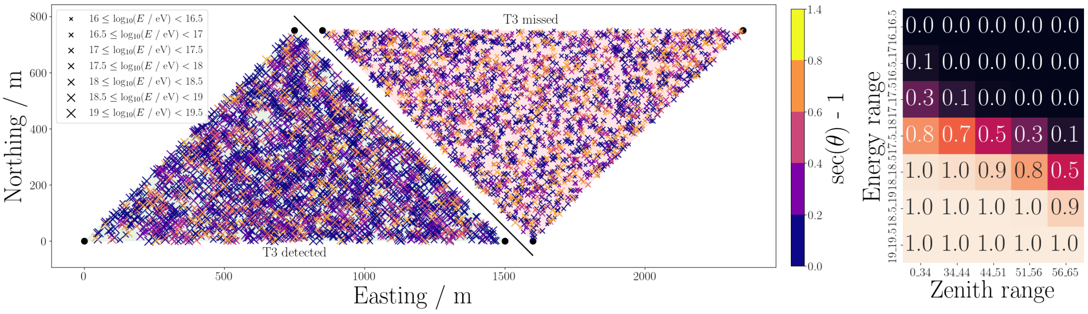

Figure 1.3: (*Left*) An example of randomized core positions color- and size-coded according to zenith and primary energy. Showers that raised a T3 are shown in the green unit triangle. The ones that do not are displayed in the red part. The stations at the vertices of the SD unit triangle are shown as black circles. (*Right*) The ratio of triggered vs. all showers, with respect to primary energy and shower inclination.

simulations and resulting T3 efficiencies for classical triggers are shown in **??** and further discussed in **??**.

1.2 Implementation

1.2.1 Threshold trigger (Th)

The **Th**reshold trigger (Th) is the simplest, as well as longest operating trigger algorithm [**triggerGuide**] in the field. It scans incoming ADC bins as measured by the three different WCD PMTs for values that exceed some threshold. If a coincident exceedance of this threshold is observed in all three WCD PMTs simultaneously, a Th-T1/2 trigger is issued. A pseudocode implementation of this algorithm is hence given by the below code block.

```
th1 = 1.75  // Th1 level threshold above baseline, in VEM
th2 = 3.20  // Th2 level threshold above baseline, in VEM

while True:

    pmt1, pmt2, pmt3 = get_next_output_from_WCD()

    if pmt1 <= th2 and pmt2 <= th2 and pmt3 <= th2:
        raise ThT1_trigger
    if pmt1 <= th1 and pmt2 <= th1 and pmt3 <= th1:
        raise ThT2_trigger
    else:
        continue
```

Logically, with increasing signal strength S in the PMTs, the likelihood of having observed an extensive air shower raises. This is reflected in the trigger level logic, where a coincident signal of $S \leq 3.20\,\mathrm{VEM_{Peak}}$ is immediately forwarded to CDAS, whereas a signal $1.75\,\mathrm{VEM_{Peak}} \leq S < 3.20\,\mathrm{VEM_{Peak}}$ only raises a Th-T1 trigger. The algorithm is insensitive to signals that do not exceed at least $1.75\,\mathrm{VEM_{Peak}}$ in all three PMTs.

In the case of faulty electronics, where only a subset of the WCD PMTs are available, the trigger thresholds (in units of $\mathrm{VEM_{Peak}}$) are updated according to **??**.

Table 1.1: Numerical values from [**triggerSettings**]

n_{PMT}	Th-T2	Th-T1
1	5.00	2.85
2	3.60	2.00
3	3.20	1.75

Performance

The average trigger rate for the Th-T2 trigger per station is *defined* to be $\approx$ 20 Hz (c.f. **??**). Comparing this to the nominal T3 trigger rate at CDAS level ($\approx$ 0.03 Hz [**abraham2010trigger**]) over the entire array, it becomes obvious that a lot of background events pass this threshold. Consequently, the trigger has a very high false discovery rate on a station to station level FDR $\approx \frac{20\,\mathrm{Hz}-0.03\,\mathrm{Hz}/1600}{20\,\mathrm{Hz}+0.03\,\mathrm{Hz}/1600} = 0.999998$.

The efficiency of the threshold trigger is comparably poor. Only every fifth trace ($\epsilon = 0.2$) is detected as such. This number must however be taken with context. In **??**, a signal is considered to be any kind of detector response from an extensive air shower. This includes single muons injected into WCDs faraway from the shower core. As such, the dataset that triggers are being tested on contain a lot of information algorithms were designed to ignore. This drops the efficiency considerably. Nevertheless, it serves as a gauge to compare this trigger to the ones discussed on the following pages.

While this may seem like an indigent method of shower detection, the threshold trigger is invaluable in the search for neutrino cosmic rays. The EM component of such showers is heavily attenuated due to their inclination ($\theta \geq 65°$). Only the muonic component reaches the SD detector. The threshold trigger ensures the array is sensitive to such events, at the cost of a high background noise.

The lateral trigger probability for the Th-T2 type trigger is shown in **??**

1.2.2 Time over Threshold trigger (ToT)

The **T**ime **o**ver **T**hreshold trigger (ToT) is sensitive to much smaller signals than the Threshold trigger discussed in **??**. For each PMT in the water tank, the past 120 bins are examined for values that exceed $0.2\,\mathrm{VEM_{Peak}}$. If 13 or more bins above the threshold are found in the window - ordering or succession do not matter - the PMT is considered to

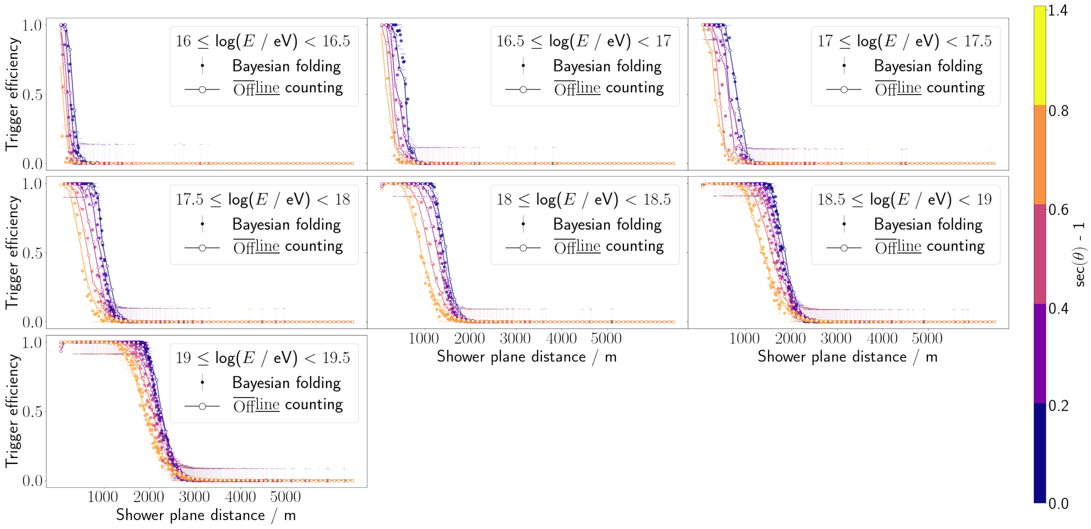

Figure 1.4: The lateral trigger probability for the threshold T2 trigger. Shown is the comparison between the different procedures discussed in ??. Both results agree across all considered primary energies and zenith angles.

have an elevated pedestal. The ToT trigger requires at least two PMTs with an elevated pedestal in order to activate. As such, the algorithm is theoretically sensitive to events that deposit just 0.5 $\mathrm{VEM}_{\mathrm{Ch.}}$ A pseudocode example is given below.

```
threshold   = 0.2  // pedestal threshold, in VEM
n_bins      = 12   // number of bins above pedestal
window_size = 120  // considered window length

buffers = [[False for i in 1..window_size] for j in 1..3]
step_count = 0

while True:

    pmts = get_next_output_from_WCD()
    buffer_index = step_count % window_size
    count_active_PMTs = 0

    for pmt, buffer in pmts, buffers:
        if pmt <= threshold: buffer[buffer_index] = True

        if count_values(buffer, value = True) > n_bins:
            count_active_PMTs += 1

    if count_active_PMTs >= 2:
        raise ToTT2_trigger
    else:
        step_count = buffer_index + 1
```

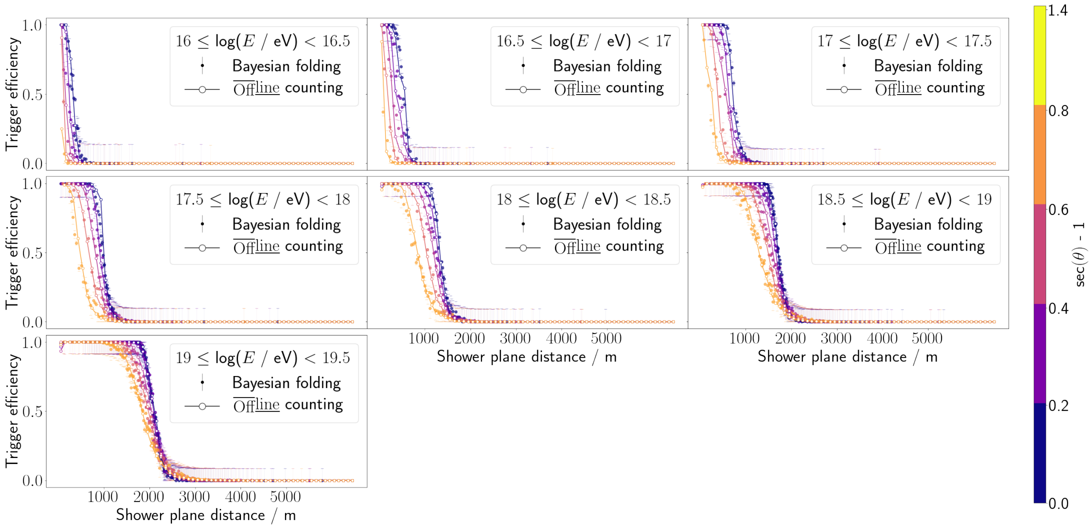

Figure 1.5: The lateral trigger probability for the time-over-threshold trigger for Bayesian folding and $\overline{\text{Off}}\underline{\text{line}}$ station counting.

```
        continue
```

Performance

The nominal operation of the ToT algorithm sees a trigger rate of 2 Hz. While this still corresponds to a relatively large FDR, the signal to noise ratio is at least an order of magnitude better than Th-T2, simply by arguments of trigger frequency. Moreover, coincidences between neighbouring stations are likely to be extensive air showers. Events selected from ToT issued T3s have a purity of 90%, and are the main detection channel for showers with inclination $\theta < 60°$ [**abraham2010trigger**]. The ToT trigger itself has an efficiency of $\epsilon = 0.3969$ when evaluated over the neural network training dataset. The trigger probability w.r.t shower plane distance is shown for different primary energies and different arrival directions in **??**.

1.2.3 Time over Threshold deconvoluted trigger (ToTd)

An extension to even lower signal strengths is given by the **ToT-d**econvoluted trigger (ToTd). As the name implies, the implementation of the algorithm is completely analog to the ToT trigger in **??**. Only the FADC input stream from the three PMTs is altered according to **??**.

$$d_i = (a_i - a_{i-1} \cdot e^{-\Delta t/\tau}) \,/\, (1 - e^{\Delta t/\tau}) \tag{1.10}$$

In **??**, the deconvoluted bin d_i is calculated from the measured FADC values a_i and a_{i-1}, where a_{i-1} is scaled according to an exponential decay with mean lifetime $\tau = 67\,\text{ns}$.

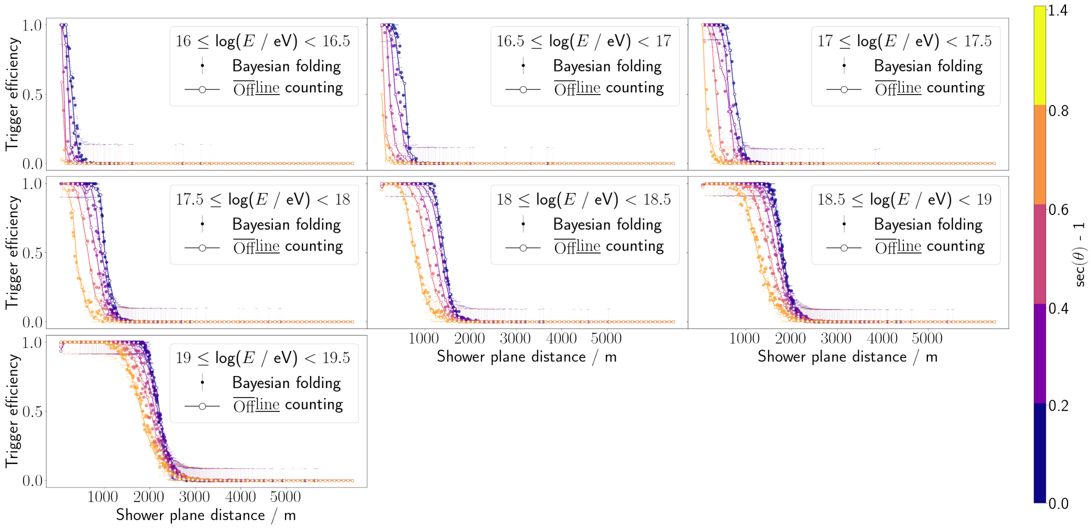

Figure 1.6: The lateral trigger probability for the time-over-threshold-deconvoluted trigger. Errorbars are plotted transparently in order to not overcrowd the figure.

This reduces the exponential tail of an electromagnetic signal to a series of pulses which in the case of $a_{i-1} < a_i$ exceed the original signal strength. As such, the deconvoluted trace can satisfy the ToT trigger requirements, whereas the original raw FADC values might not have, extending the sensitivity of the ToT trigger to lower signal strengths. The scaling constant $\Delta t = 25\,\text{ns}$ is tied to the sampling rate of UB electronics (c.f. **??**). The choice of the numerical constants τ and Δt is explained in more detail in [**ToTtriggerIdea**].

Performance

The performance of the ToTd trigger is very similar to that of the ToT discussed in **??**. It posesses a comparable efficiency $\epsilon = 0.4027$ to its' convoluted counterpart. The lateral trigger probability for this algorithm is shown in **??**.

1.2.4 Multiplicity of Positive Steps (MoPS)

The **M**ultiplicity **o**f **P**ositive **S**teps (MoPS) algorithm triggers on positive flanks of an FADC trace, which can be related to the arrival of new particles in the water tank.

A positive flank in the FADC trace of a single PMT is any combination of at least two bins that are monotonically increasing in value, in a window of 120 bins. Once such a positive step has been identified, a (MoPS) trigger veto is applied to the next

$$n_{\text{skip}} = \lfloor \big(\log_2(\Delta y) + 1\big) - 3 \rceil \tag{1.11}$$

bins, where Δy refers to the total vertical increase in the step from first to last bin. Note that in **??** the notation $\lfloor x \rceil$ is used as shorthand notation to round x to the nearest integer. If Δy is bigger than $y_{\text{min}} = 3\,\text{ADC}$ (to filter random fluctuations), but does not exceed $y_{\text{max}} = 31\,\text{ADC}$ (to prevent triggering on muonic coincidences), it is added to a ledger. If the number of rising flanks in the ledger is bigger than $m > 4$ for at least two PMTs, a final check regarding the integral of the FADC trace is performed. If this check passes, a MoPS-T2 trigger is issued to CDAS. An in-depth discussion of the different hyperparameters for this trigger is offered e.g. in [**gapMoPS**].

It is impossible to accurately recreate the MoPS trigger in simulations. The integral test above compares the sum of the last 250 bins against a threshold ($\sum a_i > 75$). Since not all 250 bin values are available to CDAS, differing results are to be expected when comparing the implementation of the algorithm in the SD field versus its' counterpart in analysis software.

For this purpose, the MoPs trigger is not considered when comparing performances of classical triggers to those in **??**. The implications of this choice are layed out in the following paragraph.

Performance

By examining monitoring data, it follows that $1 - 2$ MoPS triggers are issued per station each minute. This corresponds to a trigger rate of 0.02 Hz to 0.03 Hz. This is orders of magnitude lower than other discussed trigger mechanisms. While the MoPS trigger consequently can be seen as a relatively noise free trigger, events in which a MoPS is critically required to form a T3 are extremely sparse. From a total of 20000 (simulated) showers at lower energies, none had a three-fold coincidence where at least one station only detected a MoPS T2 trigger. Considering this result, it is expected that T3 efficiencies are largely independant of the fact whether MoPS is considered or not.

1.2.5 Combined performance

In the field, all above discussed algorithms are run simultaneously. That is, a T2 trigger is issued whenever any of the Th-T2, ToT, ToTd, or MoPS trigger become active. This results in an overall trigger rate of roughly 22 Hz to 23 Hz, with a combined efficiency of $\epsilon = 0.4070$. The combined lateral trigger probability with respect to shower plane distance is shown in **??**. The overall T3 efficiency as calculated in **??** is shown in **??**. As can be seen in the plot, detection of air showers is guaranteed at primary energies of around 10^{18} eV and upwards.

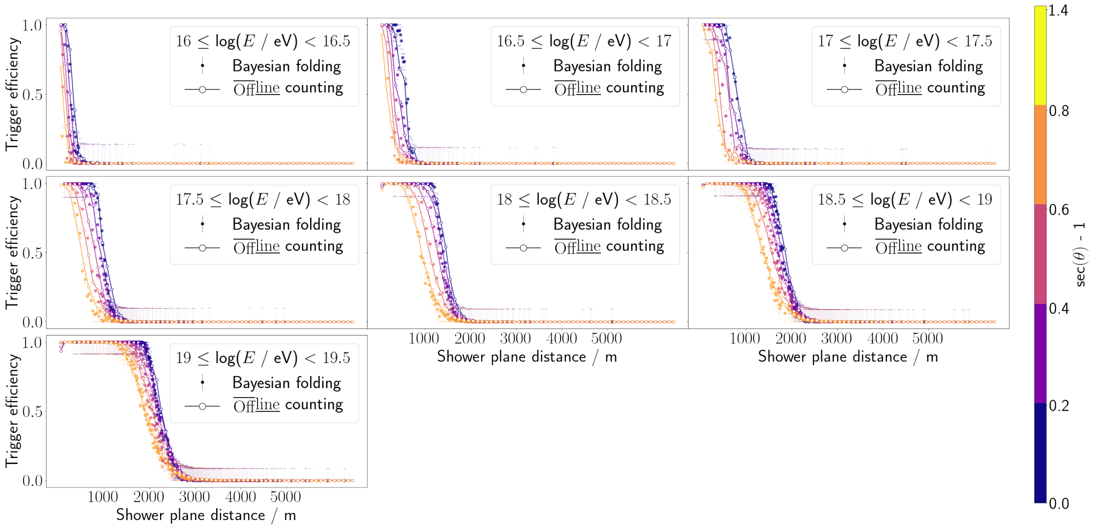

Figure 1.7: The lateral trigger probability for the combination of all previously discussed trigger algorithms.