



The optimal on-source region size for detections with counting-type telescopes



S. Klepser

DESY, D-15738 Zeuthen, Germany

ARTICLE INFO

Article history:

Received 2 December 2016

Revised 14 December 2016

Accepted 13 January 2017

Available online 17 January 2017

Keywords:

Gamma-ray astronomy

Statistics

Li&Ma statistics

Source detection

Selection cuts

ABSTRACT

Source detection in counting type experiments such as Cherenkov telescopes often involves the application of the classical Eq. (17) from the paper of Li & Ma (1983) to discrete on- and off-source regions. The on-source region is typically a circular area with radius θ in which the signal is expected to appear with the shape of the instrument point spread function (PSF). This paper addresses the question of what is the θ that maximises the probability of detection for a given PSF width and background event density. In the high count number limit and assuming a Gaussian PSF profile, the optimum is found to be at $\zeta_{\infty}^2 \approx 2.51$ times the squared PSF width $\sigma_{\text{PSF}39}^2$. While this number is shown to be a good choice in many cases, a dynamic formula for cases of lower count numbers, which favour larger on-source regions, is given. The recipe to get to this parametrisation can also be applied to cases with a non-Gaussian PSF. This result can standardise and simplify analysis procedures, reduce trials and eliminate the need for experience-based ad hoc cut definitions or expensive case-by-case Monte Carlo simulations.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Classical on-off detection techniques are still widely applied for ground-based gamma-ray observatories like H.E.S.S., MAGIC or VERITAS. In this approach, the event number in a signal (“on-source”) region is statistically compared to that of an assumedly source-free background (“off-source”) region. The size of the signal region is defined through a so-called θ^2 -cut, with θ being the opening angle between reconstructed gamma-ray direction and source position. Taking into account the instrument point spread function (PSF), the cut is usually either set to a canonical value of the order of the 68% containment PSF radius (e.g. in VERITAS [1]), or a canonical value of a fixed efficiency cut (75% in MAGIC [2]) or optimised case-by-case using Monte Carlo (MC) simulations (in H.E.S.S. [3]). A possible source detection is evaluated using the formula introduced in Eq. (17) of the famous Li & Ma paper [4].

Although an ad hoc, experience-based choice or a full MC optimisation of the θ^2 problem can lead to good results, this paper argues for a simple mathematical solution, which is much easier and more flexible to apply without computing efforts, and at full transparency of procedures and trials.

The work in this paper focuses on point-like sources (or sources with known extension). It neglects the aspects of systematic un-

certainties and Poissonian count numbers, which require other or additional constraints that can easily be adopted as needed.

2. Nomenclature

The main parameters that have influence on the number of events in the on- and off-regions for a given θ^2 -cut are the background event density n , the number of photons provided by the source N_{src} , and the gamma-ray point spread function, which in the Gaussian approximation is determined by the parameter $\sigma_{\text{PSF}39}$. This Gaussian sigma in two dimensions contains about 39% of the signal events.¹

In case of a source with known extension, $\sigma_{\text{PSF}39}$ can simply be replaced by the source size $\sigma_{\text{SRC}39}$. If the PSF (or source extension) is energy dependent, an effective PSF for the considered energy range has to be computed. For spectral studies, each energy bin might have its own $\sigma_{\text{PSF}39}$, in which case the optimal sensitivity requires one θ^2 -cut per energy bin. The following calculations can thus either be applied to an integral signal or each energy bin of a spectral study separately.

The calculations are simplified considerably defining

$$\begin{aligned} \zeta &= \theta / \sigma_{\text{PSF}39} \\ \tilde{n}_{\text{bkg}} &= n \pi \sigma_{\text{PSF}39}^2 \end{aligned} \quad (1)$$

¹ see Appendix A for how to derive it from a 68% containment radius and more details on the 2D Gaussian calculus used in this paper.

E-mail address: stefan.klepser@desy.de

Table 1
List of precise numerical constants presented in this paper.

Variable	Value
ζ_∞^2	2.51286242
ζ_∞	1.58520106
$\zeta_\infty, 68\%$	1.04621793
p_0	160.607603
p_1	4.28324658
p_2	0.0789513156

with ζ being the PSF-scaled θ and \tilde{n}_{bkg} the number of background events within a circle of radius $\sigma_{\text{PSF}39}$. In this case, the 2D Gaussian signal distribution can be expressed as

$$\frac{dN}{d\zeta^2} = \frac{N_{\text{src}}}{2} \exp(-\zeta^2/2). \quad (2)$$

Locally around the source (within a few $\sigma_{\text{PSF}39}$), the background of an instrument with a field of view $\gg \sigma_{\text{PSF}39}$ is always well-described by an isotropic background density, which can be extracted a-priori from an off-source θ^2 (or ζ^2) histogram or skymap. So the expected numbers of excess events and on- and off-events for a given cut in ζ^2 amount to

$$\begin{aligned} N_{\text{ex}} &= N_{\text{src}} (1 - \exp(-\zeta^2/2)) \\ N_{\text{off}} &= \tilde{n}_{\text{bkg}} \zeta^2 \\ N_{\text{on}} &= N_{\text{ex}} + N_{\text{off}}. \end{aligned} \quad (3)$$

3. Simple case

An important number can be derived considering the simplified significance

$$S_{\text{simple}}(N_{\text{ex}}, N_{\text{off}}) = \frac{N_{\text{ex}}}{\sqrt{2N_{\text{off}}}}, \quad (4)$$

which is Eq. (9) from ref. [4] assuming $\alpha = 1$ and $N_{\text{on}} \approx N_{\text{off}}$ for the denominator. In this case, applying Eq. (3) leads to

$$S_{\text{simple}}(\zeta^2, N_{\text{src}}, \tilde{n}_{\text{bkg}}) = \frac{N_{\text{src}}}{\sqrt{2\tilde{n}_{\text{bkg}}}} \frac{1 - \exp(-\zeta^2/2)}{\sqrt{\zeta^2}}. \quad (5)$$

The shape of this function in dependence of the cut value ζ^2 is shown Fig. 1 (top left). It has a maximum whose position is invariant against background density and signal strength, and which can analytically be determined to be

$$\zeta_\infty^2 = -2W_{-1}\left(\frac{1}{2\sqrt{e}}\right) - 1 \approx 2.51 \quad (6)$$

where $W_{-1}(x)$ is the Lambert-W function. A more precise value of ζ_∞^2 is shown in Table 1.

4. Li & Ma case

The Li & Ma significance depends on the background density and the signal strength and is therefore slightly more difficult to evaluate. As can be seen in Fig. 1 (top left), it generally needs a slightly higher number of source events to get to a given significance value, and favours a somewhat larger signal region cut.

The complexity can however be reduced if one considers the fact that for a source detection, only a signal strength $N_{\text{src},5}$ is of interest that can just actually lead to a significant detection (typically the canonical 5σ). Therefore, the calculation of this optimum for a given background density can be done in two dimensions: The maximum of $S_{\text{LiMa}}(\zeta^2)$ is determined numerically for a given \tilde{n}_{bkg} and N_{src} , and the latter is increased until $S_{\text{LiMa}} = 5$, resulting both in $N_{\text{src},5}$ and its respective ζ_{opt}^2 .

Fig. 1 (middle left) shows the dependence of the optimal cut on the background density. Clearly, in the highly Gaussian regime ($\tilde{n}_{\text{bkg}} > 100$), the optimum cut value approaches ζ_∞^2 (therefore the index “ ∞ ”), but cases of low count numbers favour a somewhat larger cut. This in reverse is equivalent to the concept outlined in ref. [3], namely that weak sources should be analysed with tighter cuts (because weak sources require large datasets, i.e. high background number) and stronger sources with looser cuts. It has to be noted that the Li & Ma formula is not valid in the very low-count Poissonian regime, roughly marked by the line labeled “ $N_{\text{off}} > 5$ ”.

The signal event efficiency implied by the cut is plotted in Fig. 2. In the high count number limit an efficiency of 71.5% is approached (proving the 75% in MAGIC ref. [2] likely to be a fair compromise in many cases).

Fig. 1 (bottom left) shows that if the dynamic adjustment of the cut with background density is replaced by a constant cut, a significance loss of the order of 10% can be expected in the low count number case. This is equivalent to a 10% loss of sensitivity or a 20% increase in required observation time. The dashed and dash-dotted curves furthermore show that if a canonical or weakly motivated constant cut radius is more than a factor of 2 away from ζ_∞ , the sensitivity can even be degraded by a factor of 2 or more.

In cases where different amounts of on- and off-exposures are available, the Li & Ma formula offers the application of the parameter α , which is the exposure ratio between the two. Typically, more off- than on-data is available, and α is smaller than 1. The dashed-dotted curve in Fig. 1 (middle left) shows that the ζ_{opt}^2 curve is almost unaltered if the background density is scaled to $\tilde{n}_{\text{bkg}}/\alpha$.

For simplicity of application, the curve in Fig. 1 (top left) can be parametrised with an analytical function of the form

$$\zeta_{\text{opt}}^2(\tilde{n}_{\text{bkg}}, \alpha) = \zeta_\infty^2 - p_0 \ln[1 - \exp(-p_1 (\tilde{n}_{\text{bkg}}/\alpha)^{p_2})] \quad (7)$$

The result is shown as a green dotted line in the figure, and the according parameters p_n are listed in Table 1.

5. Non-Gaussian point spread functions

Although the point spread functions of instruments can usually be approximated by a Gaussian distribution to some level, the exact distributions are sometimes more complex. Misreconstructed events can lead to non-Gaussian tails of the PSF. In order to test the robustness of the above results in these cases, the calculations are repeated using a so-called King profile

$$\frac{dN}{d\zeta'^2} = \frac{N_{\text{src}}}{2} (1 - 1/\gamma) \left(1 + \frac{\zeta'^2}{2\gamma}\right)^{-\gamma}. \quad (8)$$

This distribution has a tail that is small for large γ and gets longer for $\gamma \rightarrow 1$. ζ' is defined as $\theta/\sigma_{\text{King}}$, and is related to ζ like

$$\zeta'^2 = \zeta^2 \times 2\gamma \left[(1 - 0.39347)^{\frac{1}{1-\gamma}} - 1 \right]. \quad (9)$$

The number of excess events in Eq. (3) now changes to

$$N_{\text{ex}} = N_{\text{src}} \left[1 - \left(1 + \frac{\zeta'^2}{2\gamma}\right)^{1-\gamma} \right] \quad (10)$$

In this case, the maximum of the corresponding significance function $S(\zeta)$ depends on γ and cannot be derived analytically, even though the shape is qualitatively similar to the Gaussian case (see Fig. 1, top right, where a PSF with $\gamma = 2$ is taken as an example). The optimum cut in the case of S_{simple} , however, is still invariant against \tilde{n}_{bkg} , and lies somewhat lower (if calculated w.r.t. $\sigma_{\text{PSF}39}$, i.e. converting ζ' to ζ).

In the Li & Ma case, the function of ζ_{opt}^2 vs. \tilde{n}_{bkg} is different from the Gaussian PSF case, but can still be fitted with the

Download English Version:

<https://daneshyari.com/en/article/5486750>

Download Persian Version:

<https://daneshyari.com/article/5486750>

[Daneshyari.com](https://daneshyari.com)