MAGIC-TDAS 09-01
090519 / T.Y. Saito & J. Sitarek

Improvement of the θ^2 analysis by using the Random Forest method in the DISP estimation

T.Y. Saito (MPI)
J. Sitarek (MPI)

May 19, 2009

Abstract

We have developed an improved method for the estimation of the DISP parameter. It is based on the use of the Random Forest (RF) method. In this report we compare both the sensitivity and the angular resolution of the new (hereafter RF DISP) and the "classical" parametrized DISP methods.

We tested the method using Crab Nebula data taken in wobble mode. Compared to the classical DISP, the θ^2 analysis using the RF DISP method improves the angular resolution by $\sim 25\%$ at 300 GeV and $\sim 45\%$ at 1 TeV. As a result the sensitivity of the θ^2 analysis improves by $\sim 20\%$ in source position dependent analysis and $\sim 30\%$ in source independent analysis.

Contents

1 Introduction

In the standard analysis DISP is parametrized by using a rather simple formula described in [1]:

$$DISP = A(SIZE) + B(SIZE) \cdot \frac{WIDTH}{LENGTH + \eta(SIZE) \cdot LEAKAGE} \tag{1}$$

The field of view of the MAGIC camera is relatively small (3.5°). For some showers part of the image is outside the camera. This effect can be parametrized by the so called LEAKAGE

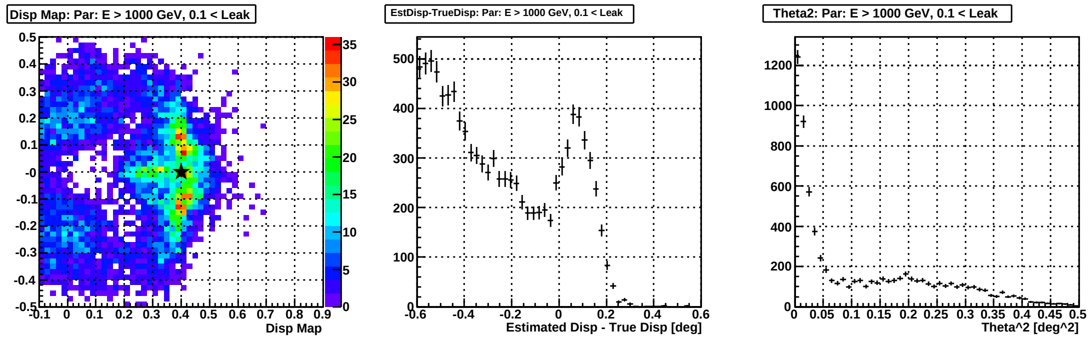

Figure 1: Standard DISP estimation of leaked images (ENERGY > 1 TeV, $0.3 >$ LEAKAGE > 0.1, Wobble HE MC). The estimated DISP position map (left figure), the distribution of the difference between the true DIST and the estimated DISP (center figure) and the θ^2 distribution (right figure). The black star in the left figure indicates the true source position.

parameter (see section 5 for more detailed information), and it is especially pronounced at higher energies. In our analysis we noticed that the DISP estimation is poor for images with a large LEAKAGE. In the Fig. 1 you can see one result of a parametrized DISP method used for wobble mode HE MCs (spectral slope -1) for images with a large value of LEAKAGE parameter. A strong bias in the DISP estimation can be seen in the middle figure. The source position estimation is very inaccurate (left figure), and the θ^2 plot (right figure) has a long tail.

It seems that the proper parameterization of the LEAKAGE is not straightforward, especially in the wobble mode. The Random Forest (RF) method would be more suitable for this rather complex multi-parameter space.

In TDAS [2] it is described that the DIST and the TIME GRADIENT parameters are strongly correlated and can be used for better hadronness estimation. In the standard analysis, the TIME GRADIENT is used for the source dependent hadronness calculation and thus improving the sensitivity. However it has not been used for the source independent analysis. The DIST - TIME GRADIENT correlation can be interpreted as a correlation between the TIME GRADIENT and the distance from the shower arrival direction to the image COG (Center Of Gravity). This means that the DISP and the TIME GRADIENT are strongly correlated and it is not necessary to assume a source position.

So it is clear that the time gradient should improve the DISP estimation, which can be used for both the θ^2 analysis and skymaps. Addition of the timing parameter to the image geometry parameters would make the DISP parameterization more difficult, while it is not the case for the RF method.

Motivated by these two reasons, we studied the improvement of the DISP estimation by using the RF regression method. The same method is commonly used in MAGIC for the energy estimation.

2 Best sensitivity study

To compare the RF DISP method with the classical parametrized DISP and the ALPHA analysis, we apply all these methods to a sample of the Crab Nebula data taken in Jan 2008. This sample consists of $T = 9.1$h of good quality data taken at the low zenith angle range ($< 30°$) in the wobble mode. We've performed 5 different analyses, applying a cut in one of the angular parameters:

- ALPHA
- θ^2, the classical parametrized DISP
- θ^2, RF DISP with geometrical parameters (the same parameters as used in the classical DISP)
- θ^2, RF DISP with geometrical and TIME GRADIENT parameters
- θ^2, RF DISP with all parameters (geometrical, TIME GRADIENT, asymmetry and zenith)

All the parameters used in the analysis can be divided into two groups. The source independent parameters like e.g. SIZE, WIDTH, LENGTH describe the shape of the image without any assumption about the source position. On the other hand source dependent parameters (e.g. DIST) take advantage of apriori knowledge of the source position. When the source position is known, one can calculate the hadronness using all available parameters (both source dependent and source independent), the so called source dependent hadronness. If the real source position is offset from the assumed one, source dependent hadronness value for gamma like events will increase, which will considerably spoil the gamma/hadron separation. Therefore, if the source position is not known at all, or it's precision is poor (e.g. comparable or worse than the telescope PSF for a point like gamma source), one may consider a different approach, the so called source independent hadronness. In this case only the source independent parameters are used in the gamma/hadron separation RF, so the source position assumption is not needed. Since only partial information about each shower is used, the gamma/hadron separation power decreases.

If the source position is not known, one has to perform a scan searching for it. To create a skymap, one projects every event on the sky, producing an ON map. Using one of the available methods in the *celestina* program also the background (OFF map) can be estimated. Only source-independent hadronness can be used for gamma/hadron separation in the case of the skymap.

Another possibility is the so called false source α analysis method. The observed region in the sky is divided into a grid of possible source positions (usually with a grid size of $\sim 0.05°$). For each of those positions a source dependent ALPHA analysis is performed. Technically false source analysis is difficult to perform, because for every assumed source position a set of *melibea* files have to be created, which takes both a lot of time and disc space. In both skymap and false source analysis, the resulting significance has to be corrected for a trial factor. The achieved sensitivity in both methods is estimated to be on a comparable level.

In principle each of the ALPHA/θ^2 analysis can be combined either with a source dependent or a source independent hadronness. Usually ALPHA is associated with the source dependent hadronness, and θ^2 with the source independent hadronness. Refraining from using the source dependent parameters in the hadronness used for the ALPHA analysis weakens it considerably, and does not bring any additional advantage. Therefore the ALPHA analysis with the source independent hadronness is not recommended and it is presented here just for the completeness of the study.

For each case we calculated the sensitivity according to the following formula:

$$\text{sensitivity} = 5/\left(\sqrt{\frac{50\text{h}}{T}}\frac{N_{on} - N_{off}}{\sqrt{N_{off}}}\right)\ \text{C.U.} \tag{2}$$

where C.U. stands for the unit of Crab Nebula flux.

As the previous study has shown (see [2]), the best source dependent sensitivity for a Crab like source is obtained with strong cuts. In our analysis we apply a SIZE cut 400, an ALPHA cut 6 degrees, and θ^2 cut 0.01 degrees^2. We use a cut in the source dependent hadronness $h < 0.05$. In the case of a source independent hadronness, as explained above, the gamma/hadron separation is less powerful. The best sensitivity is obtained with a looser cut. Therefore in this case we choose to apply source independent hadronness cut $h < 0.1$. We have cross-checked that the used values provide optimal significance (see Fig. 2 and 3).

Values of sensitivity obtained with the Crab data are shown in Table 1 for all types of analyses. Corresponding ALPHA and θ^2 distributions are shown in Fig. 4.

	ALPHA	θ^2 parametrized DISP	θ^2 DISP RF geometrical	θ^2 DISP RF time	θ^2 DISP RF all
Src. dep. hadronness	1.44 ± 0.06	1.81 ± 0.12	1.60 ± 0.10	1.41 ± 0.08	1.40 ± 0.08
Src. indep. hadronness	3.52 ± 0.14	2.37 ± 0.15	2.17 ± 0.13	1.64 ± 0.09	1.63 ± 0.09

Table 1: The sensitivity (in % C.U.) obtained with a wobble Crab sample, for different types of analyses, SIZE> 400, ALPHA < 6, $\theta^2 < 0.01$; hadronness < 0.05 (source dependent) < 0.1 (source independent)

The best sensitivity value for an ALPHA analysis $1.44 \pm 0.06\%$ C.U. seems to be somewhat better than the value ($1.59 \pm 0.07\%$ C.U.) given in [2]. This improvement is partially due to the usage of the absolute starguider calibration, which makes ALPHA and θ^2 distributions narrower.

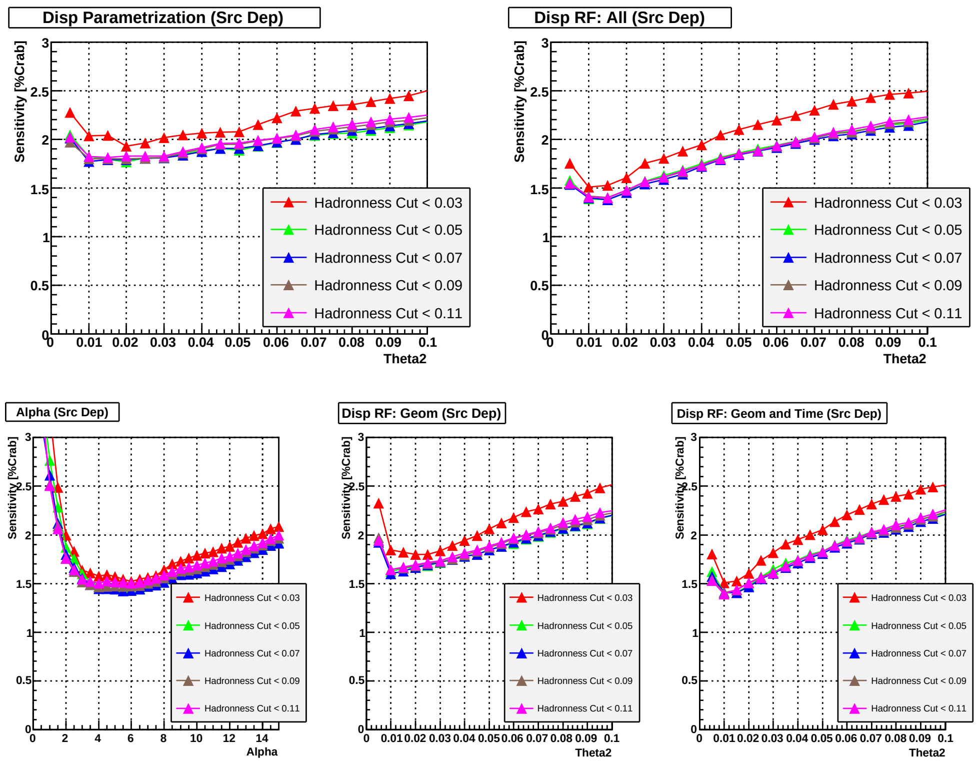

Figure 2: The sensitivity (in percentage of Crab) obtained with the source dependent hadronness for $SIZE > 400$. Parametrized DISP (top left), RF DISP with all parameters (top right), ALPHA(bottom left), RF DISP with geometrical parameters (bottom middle) and RF DISP with geometrical and time gradient parameters (bottom right).

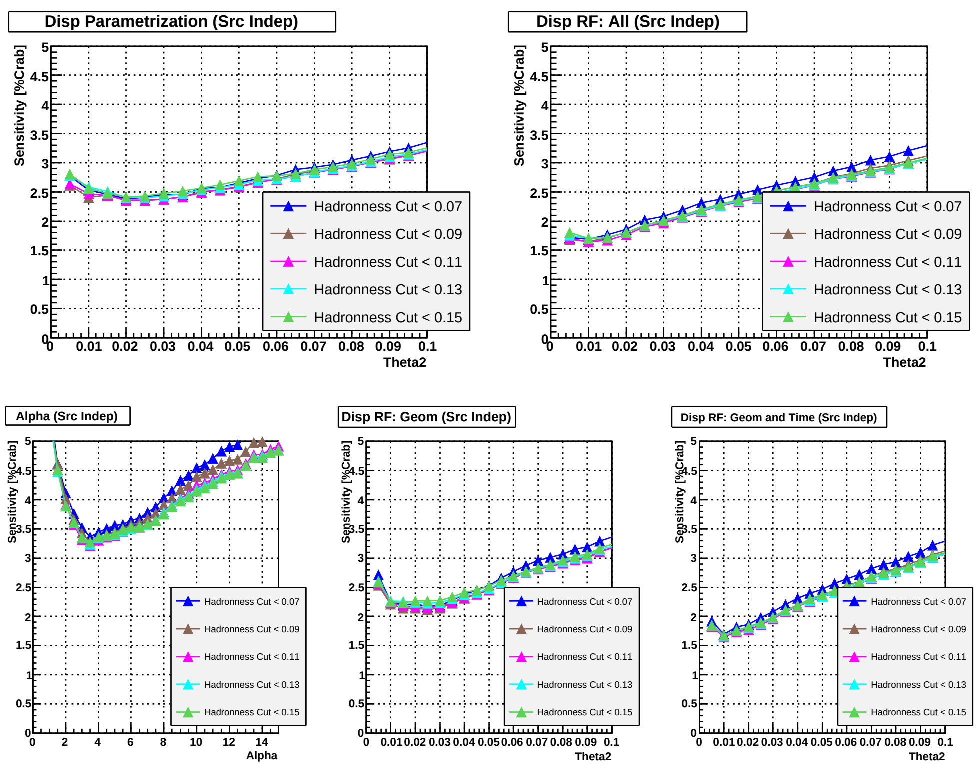

Figure 3: The same as in Fig. 2 but for the source independent hadronness.

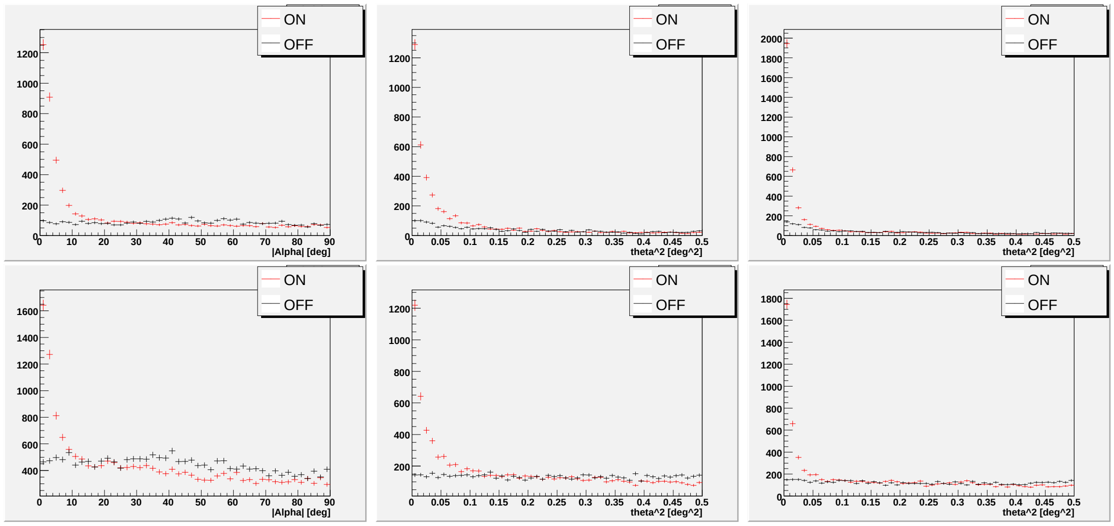

Figure 4: Distributions for ON and OFF: ALPHA (left), parametrized θ^2 (center), RF θ^2 (right), SIZE > 400. Upper figures are for a source dependent analysis (hadronness < 0.05), lower are for a source independent (hadronness < 0.1). Crab Nebula data taken in the wobble mode.

As one can see from Table 1 using the RF method for the estimation of DISP increases the sensitivity for both source dependent and source independent θ^2 analyses. Further improvement is achieved by using the time gradient in DISP RF. Usage of additional parameters of the image (asymmetry, M3Long) does not seem to improve the sensitivity. The total relative improvement is $20 - 30\%$.

As it is explained earlier ALPHA combined with a source independent hadronness provides a very weak sensitivity. Unlike ALPHA, θ^2 analysis in that case is preferred, and the resulting sensitivity is comparable to the sensitivity of the ALPHA analysis with the source dependent hadronness. It is due to a fact that the θ^2 cut is constraining in two spatial directions in the camera (along and perpendicular to the main axis of the image), while ALPHA does it only in one (perpendicular). An implicit constrain in the second dimension in the case of the ALPHA analysis is obtained by using the usage of the source dependent parameters (DIST and the TIME GRADIENT with sign). This is why the θ^2 analysis does not depend so strongly on the source dependent parameters used in the hadronness calculation as the ALPHA analysis.

In the ALPHA analysis one can see a hint on a broad bump in the OFF region at $\sim 50^\circ$ (Fig. 4 left). It is a feature of the wobble mode observations of a strong source and it is due to the real signal events calculated with respect to the OFF position. The bump is more pronounced in the case of the source independent hadronness, due to a fact that a given event has the same hadronness calculated with respect to ON and OFF position. Thus, the integral of the ON-OFF excess in the signal region (e.g. ALPHA < 12) is equal to the OFF-ON bump in the background region (e.g. $12 <$ ALPHA < 90). On the other hand, the source dependent hadronness of a real signal event calculated with respect to the OFF position is much larger than for the ON position. Therefore, in the case of the source dependent hadronness, this bump is partially suppressed. In the case of the θ^2 analysis (with only one OFF region) a similar bump is expected at $(0.8^\circ)^2 = 0.64$, which is far away from the signal region.

One of the most important source dependent parameters used in the gamma/hadron separation is DIST. For events with a small value of the source dependent hadronness, the estimated DISP value will be close to the true value of DIST. Thus θ^2 and the source dependent hadronness are strongly correlated. This results in the background region also peaking around 0 (upper center and right figures in Fig. 4). The disadvantage of the method is that the signal from a weak source would present itself as a small peak on top of this broad peak. Due to various systematic errors the broadness of the latter can vary thus mimicking a genuine signal. For source independent hadronness combined with the θ^2 analysis this effect does not take place, so the background is flat.

3 Sensitivity in different SIZE bins

In the previous chapter we considered the best sensitivity for the well defined showers (SIZE > 400 phe). We decided to check the 5 analysis methods by applying them also to 3 SIZE bins ($100 < SIZE < 300$, $300 < SIZE < 1000$, and $1000 < SIZE$). The resulting sensitivities as a function of the hadronness and the θ^2 / ALPHA cuts are shown in figures 5 to 10. The achieved best sensitivity for the above mentioned SIZE bins are compared in tables 2 and 3.

The behavior of the best sensitivity of the source-dependent hadronness analysis is similar to the case described in chapter 2.:

- DISP RF improves the sensitivity,
- TIME GRADIENT improves it further
- additional parameters (M3LONG, asymmetry) don't help.

However, the overall improvement is lower ($\sim 10\%$) in the SIZE range of $100 < SIZE < 300$ than for $SIZE > 300$ ($\sim 20\%$). At high sizes images are well defined and the θ^2 analysis can show its full potential being as good as ALPHA. For lower sizes ALPHA analysis is still more powerful (by $\sim 15\%$), even when compared to the best RF θ^2.

The general trends seen on the source-dependent hadronness analysis are true also for source-independent hadronness analysis for SIZE ¿ 300. Below 300, due to the larger errors the situation is not conclusive. The overall improvements are $\sim 10\%, \sim 20\%$, and $\sim 35\%$ for the SIZE range $100 < SIZE < 300$, $300 < SIZE < 1000$, and $1000 < SIZE$ respectively. Even for the low sizes θ^2 provides a better sensitivity than ALPHA.

SIZE Range	source dependent				
	Alpha	θ^2 param	θ^2 geom.	θ^2 time	θ^2 all
100-300	10.31 ± 0.85	13.48 ± 1.42	13.57 ± 1.44	12.01 ± 1.14	11.92 ± 1.12
300-1000	2.15 ± 0.09	2.85 ± 0.14	2.67 ± 0.13	2.33 ± 0.10	2.28 ± 0.10
> 1000	1.34 ± 0.14	1.61 ± 0.17	1.50 ± 0.19	1.32 ± 0.15	1.18 ± 0.19

Table 2: Sensitivity of the analysis for the source dependent hadronness cut in % C.U.. Hadronness cut is applied at 0.1 for SIZE 100 to 300, at 0.07 for SIZE 300 to 1000, and at 0.05 for SIZE above 1000. Alpha and θ^2 cut is determined by the best sensitivity shown in figures 5 to 7. If the error of the sensitivity estimation is more than 20% due to the lack of background, the cut was loosened.

SIZE Range	source independent				
	Alpha	θ^2 param	θ^2 geom.	θ^2 time	θ^2 all
100-300	27.46 ± 5.61	19.04 ± 3.27	22.43 ± 4.34	19.36 ± 3.37	17.33 ± 2.79
300-1000	5.57 ± 0.30	3.82 ± 0.22	3.74 ± 0.21	3.06 ± 0.20	2.99 ± 0.19
> 1000	3.56 ± 0.24	2.20 ± 0.21	2.02 ± 0.18	1.15 ± 0.23	1.38 ± 0.17

Table 3: Sensitivity of the analysis for the source independent hadronness cut in % C.U.. Hadronness cut is 0.1 for SIZE below 1000 and at 0.07 for SIZE above 1000. Alpha and θ^2 cuts are determined by the best sensitivity from figures 8 to 10. If the error of the sensitivity estimation is more than 20% due to the lack of background, the cut was loosened. In the case of θ^2 analysis with DISP RF with geometrical parameters (4th column), SIZE range 100-300, θ^2 cut 0.02 is used.

4 Angular resolution

The θ^2 distribution for the ON-OFF excess is shown in Fig. 11. The DISP RF method gives much more peaked distribution than the parametrized DISP. When using the time parameters in the DISP RF, the distribution becomes even narrower. The angular resolution of the MAGIC telescope can be estimated using the cumulative θ (or θ^2) distributions of the ON-OFF excess. A part of the showers have a wrongly reconstructed sign of DISP (the head-tail discrimination gives a wrong result). This leads to very large values of θ^2. Showers with a reconstructed source position further than 0.4° from the assumed source position are treated as a background and not taken

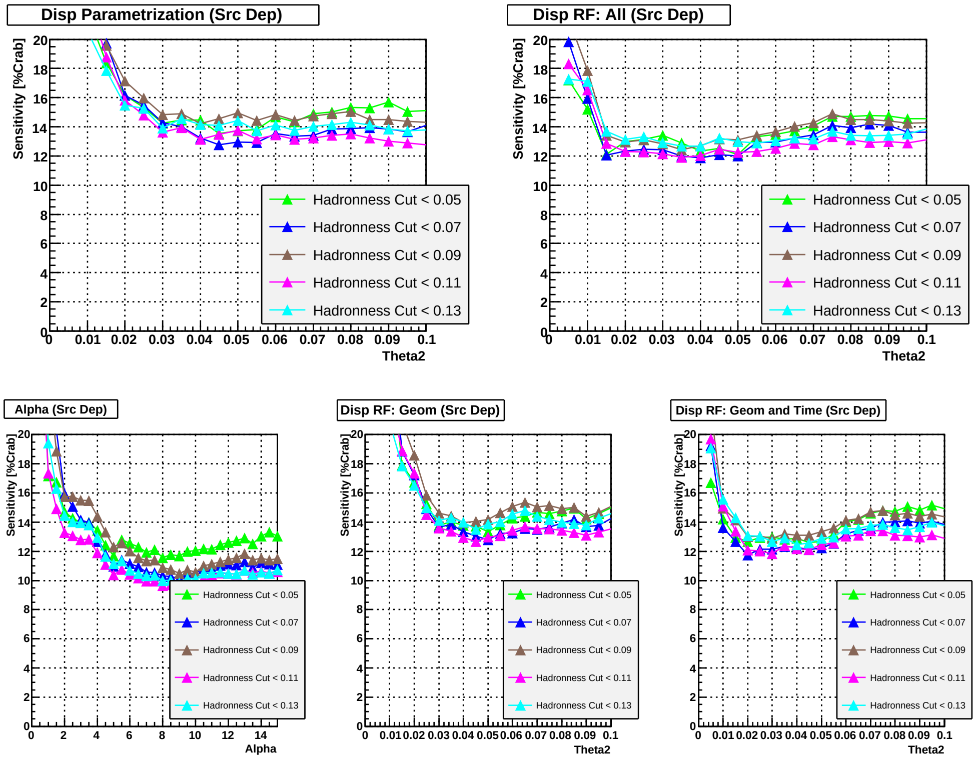

Figure 5: Analysis with the source-dependent hadronness in SIZE Range $100 < SIZE < 300$. Parametrized DISP (top left), RF DISP with all parameters (top right), ALPHA(bottom left), RF DISP with geometrical parameters (bottom middle) and RF DISP with geometrical and time gradient parameters (bottom right).

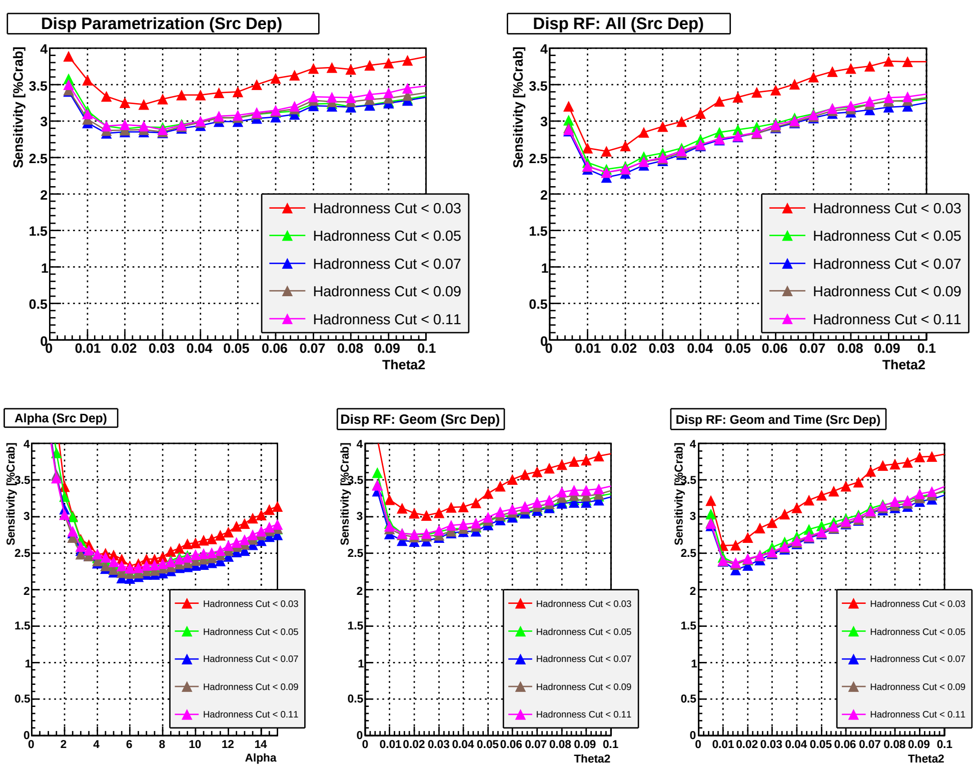

Figure 6: The same as in Fig. 5 but for the SIZE range $300 < SIZE < 1000$.

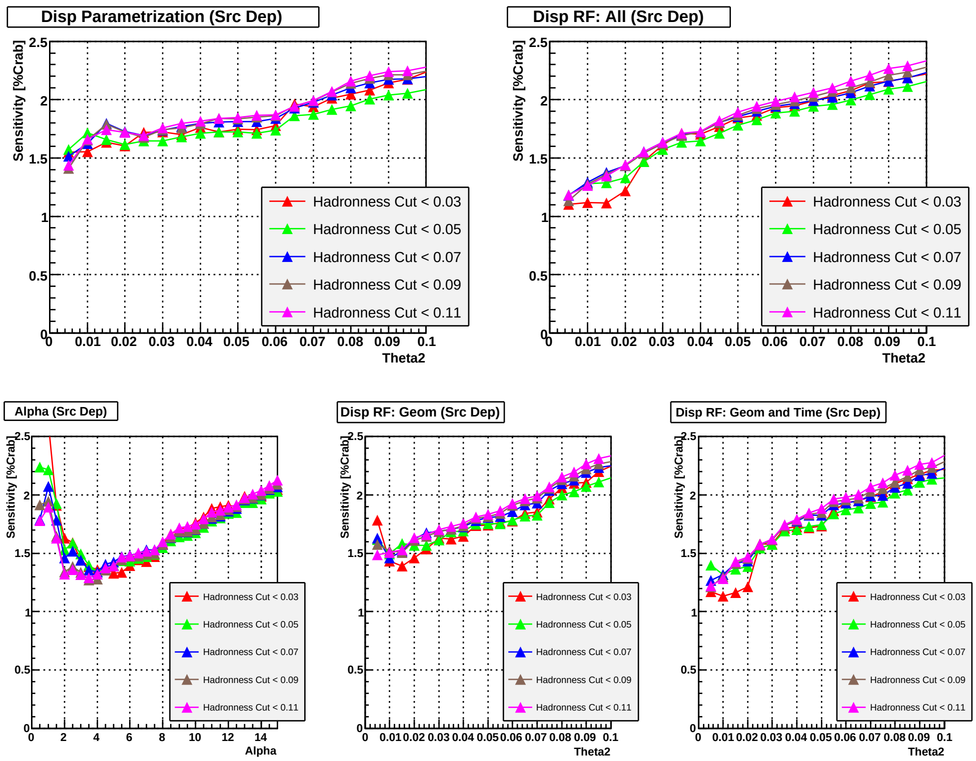

Figure 7: The same as in Fig. 5 but for the SIZE range $1000 < SIZE$.

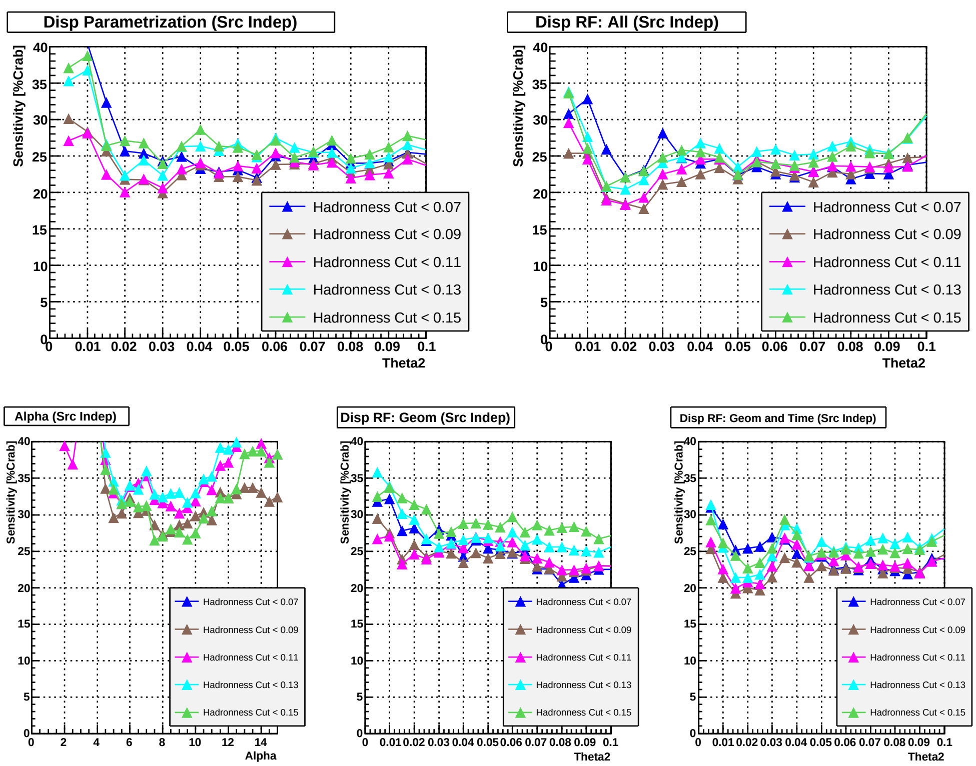

Figure 8: Analysis with the source-independent hadronness in SIZE range $100 < SIZE < 300$. Parametrized DISP (top left), RF DISP with all parameters (top right), ALPHA(bottom left), RF DISP with geometrical parameters (bottom middle) and RF DISP with geometrical and time gradient parameters (bottom right).

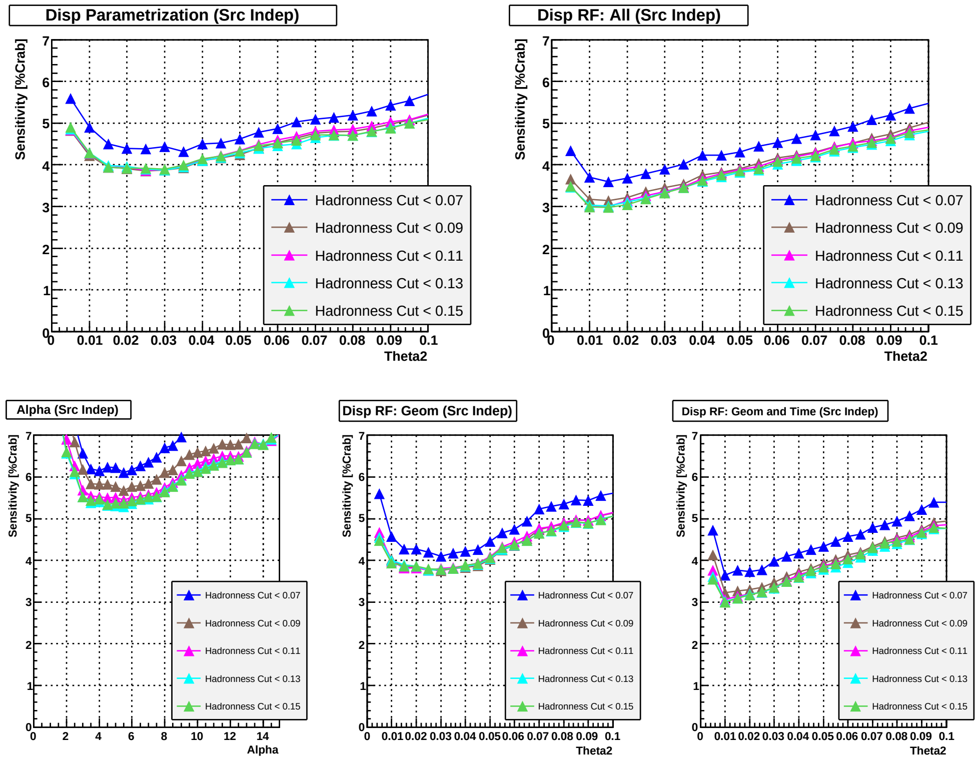

Figure 9: The same as in Fig. 8 but for SIZE range $300 < SIZE < 1000$.

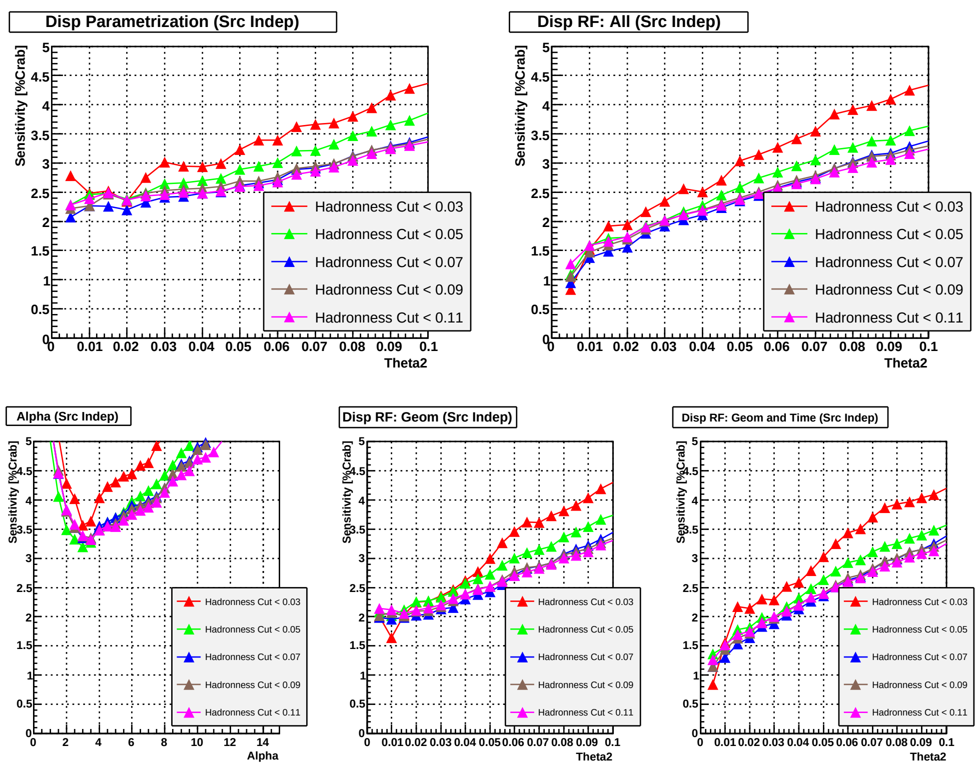

Figure 10: The same as in Fig. 8 but for SIZE range $1000 < SIZE$.

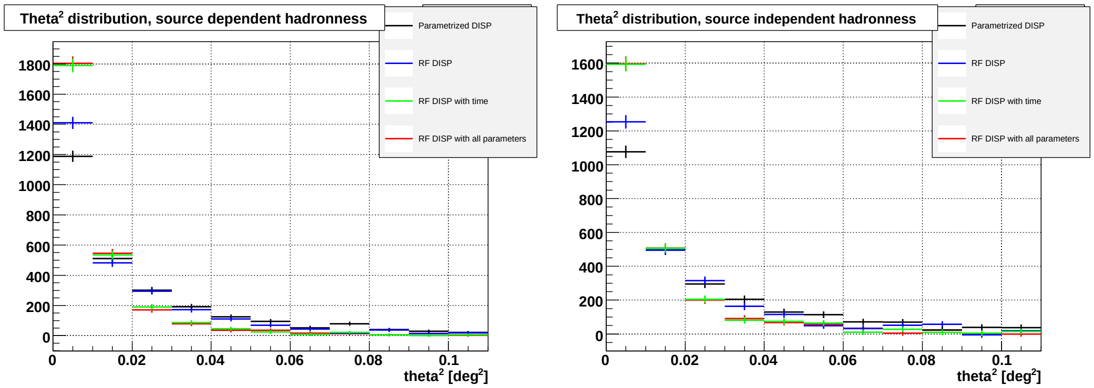

Figure 11: Comparison of the θ^2 distribution of the excess obtained with different methods: parametrized DISP (black), DISP RF with only geometrical parameters (blue), with geometrical and time parameters (green), with all parameters (red). Left figure is for the source dependent hadronness < 0.05, right for the source independent hadronness < 0.1.

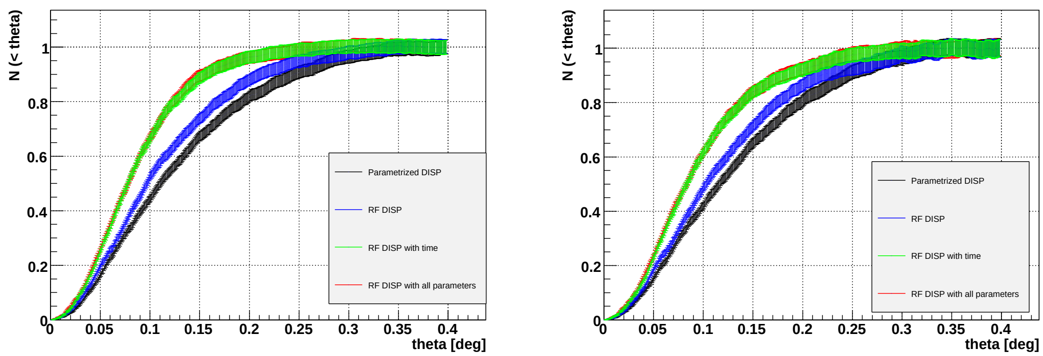

Figure 12: Cumulative θ^2 distribution of the excess from Crab Nebula data, obtained with different DISP methods. Left figure is for the cut in the source dependent hadronness < 0.05, right for the cut in the source independent hadronness < 0.1, SIZE > 400. Thickness of the curve represent statistical error.

into account while calculating the angular resolution. Depending on the SIZE of the image, the percentage of the misclassified showers is around $10 - 20\%$.

In Fig. 12 cumulative θ^2 distributions for all the used methods are shown. Selecting a given fraction of events, one can read the radius which contains the corresponding excess. Table 4 shows the radii corresponding to a few widely used definitions commonly used in the estimation of the angular resolution.

Usage of the RF method in the DISP estimation gives an improvement of the order of 10-15% in the angular resolution. Adding time parameters boosts again the angular resolution for additional 10-20%, giving a total improvement of 20-30%.

With the Monte Carlo simulations the angular resolution can be calculated for energy bins. Results of those calculations are shown in Fig. 13. As one can see there, above 1 TeV the 68% (41%) containment angular resolution is $\sim 0.08°$ ($\sim 0.05°$). Below 1 TeV the angular resolution is becoming comparable to MAGIC II stereo reconstruction.

5 Images with leakage

Showers for which the images are truncated (a part of the image is outside of the camera) are especially difficult to reconstruct. To determine how strong is this effect, for every shower the value of the so-called LEAKAGE1 is calculated. It is defined as a fraction of the SIZE of the image contained in the last ring of the camera. Images with a large leakage (> 0.2) are usually excluded from analysis. In Fig. 14 a comparison of θ^2 distributions produced with various DISP reconstruction methods for different values of leakage is presented.

fraction	source dependent				source independent			
	θ^2 param.	θ^2 geom.	θ^2 time	θ^2 all	θ^2 param.	θ^2 geom.	θ^2 time	θ^2 all
0.41 (1D gauss)	0.09	0.08	0.07	0.07	0.10	0.09	0.07	0.07
0.5	0.11	0.10	0.08	0.08	0.12	0.10	0.08	0.08
0.68 (2D gauss)	0.15	0.13	0.10	0.10	0.16	0.14	0.11	0.11

Table 4: The angular resolution (the radius, in units of degree, containing the given excess fraction), obtained by using the Crab sample, for different types of analysis. $\mathrm{SIZE} > 400$, $\mathrm{hadronness} < 0.05$ (source dependent analysis) < 0.1 (source independent)

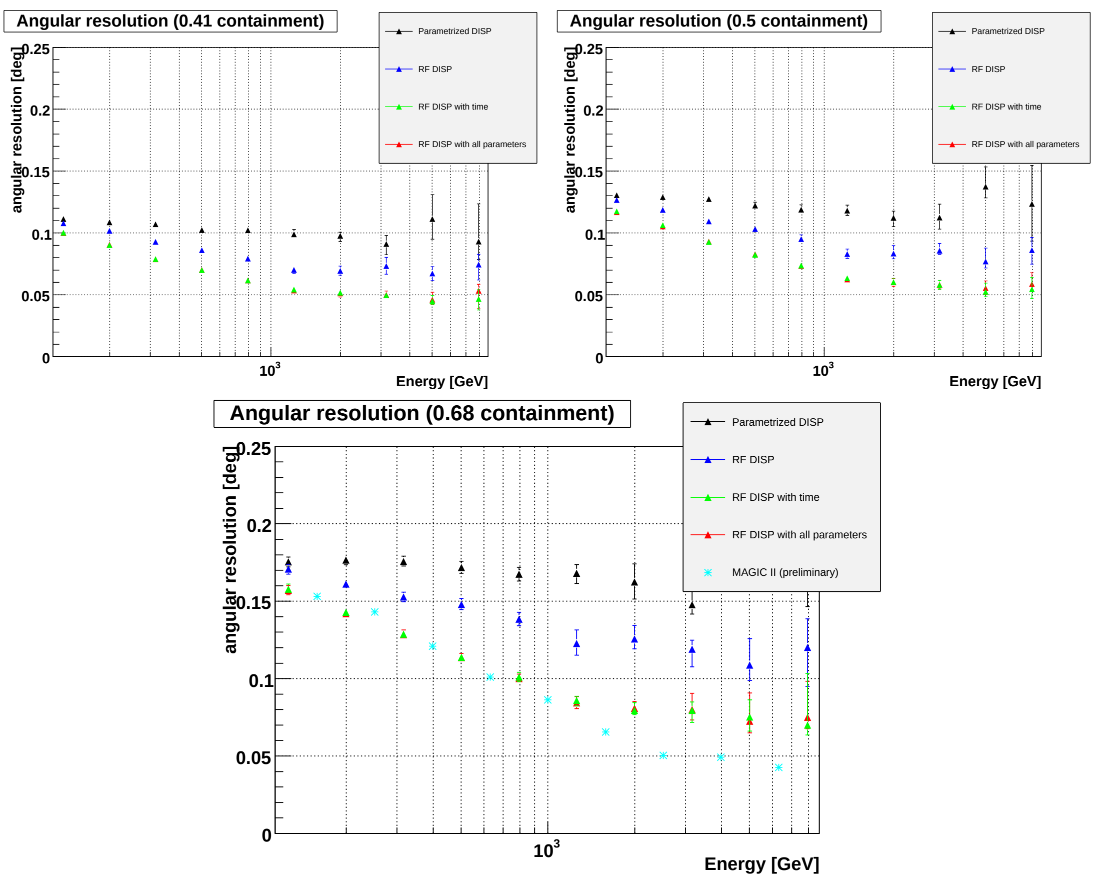

Figure 13: Monte Carlo calculations of the energy dependence of the angular resolution for different DISP methods. Different figures show the confinement radii for different fractions of gamma-like events: 0.41 (top left), 0.5 (top right) and 0.68(bottom). $\mathrm{hadronness} < 0.2$, $\mathrm{SIZE} > 100$. Cyan stars represent the preliminary angular resolution of the MAGIC II calculated with the standard stereo reconstruction method (P. Colin at DPG 09).

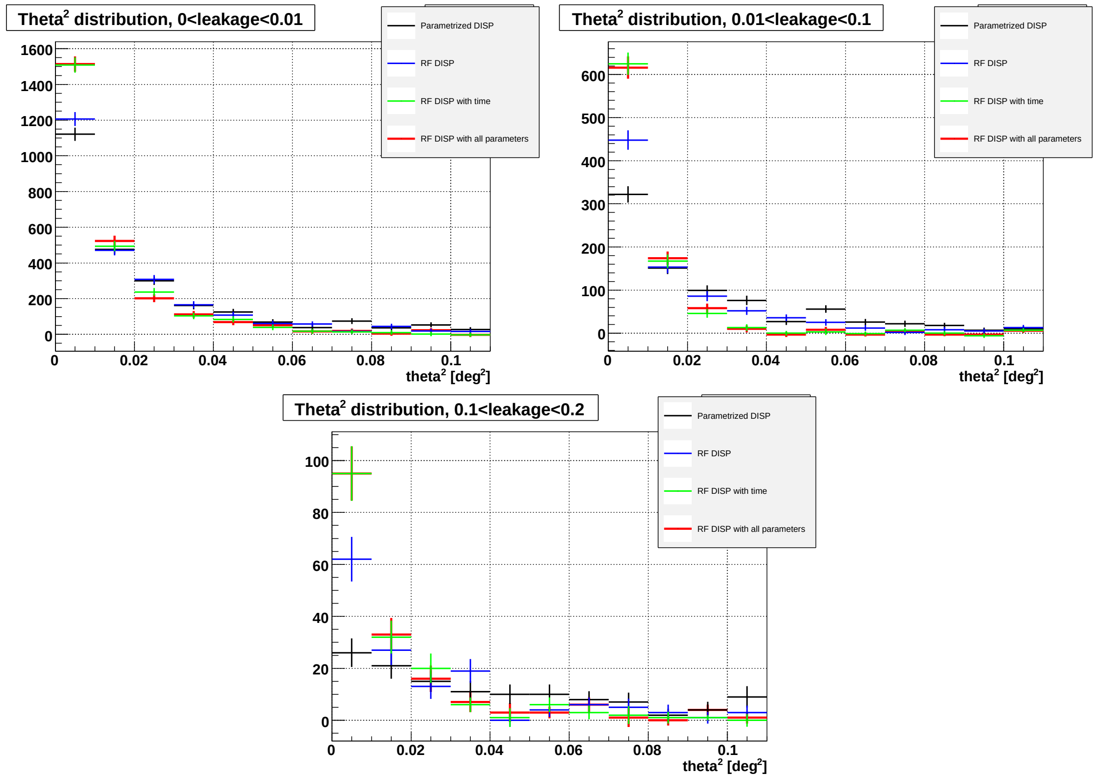

Figure 14: Comparison of the θ^2 distribution of the excess obtained with different DISP methods for images without leakage (top left), with small LEAKAGE1 (top right) and with moderate value of LEAKAGE1 (bottom). Source dependent hadronness < 0.1, SIZE> 400

The DISP RF method is the best in all the cases. For images with $0.1 <$ LEAKAGE1 < 0.2 the distribution of parametrized θ^2 is very broad. Even in this case the RF DISP (especially with the usage of time parameters) provides narrow distribution.

6 Ghostbusting

In the standard analysis estimation of the source position is a two stage process. In the first step the absolute value of DISP is calculated. For the given value of DISP and the direction of main axis of the image there are still two possible solutions for the estimated source position: on opposite sides of the COG. For selecting the correct one the asymmetry of the image is used. This second step is called GhostBusting (GB), or more commonly head-tail discrimination.

Hillas Parameters used for the DISP estimation can be categorized into unsigned parameters (LENGTH, WIDTH, SIZE, LEAKAGE etc. – 0th or 2nd moments) and signed parameters (TIME GRADIENT, M3Long, ASYM – odd moments). One may hope that by using the signed parameters and training the Random Forest with "Signed DIST", GB would become unnecessary.

In a simple example we will see that it is not the case. Let's consider a type of showers for which the absolute value of DISP is determined with a good precision around the true value X, but GB gives a wrong sign in a small fraction ($\sim 10\%$) of cases. In this case 90% of the RF trees gives DISP $= X$, and 10% DISP $= -X$. The final value of the RF for this type of the showers is just the average value of all the values from all the trees. Thus the obtained value will be $0.9 \cdot X + 0.1 \cdot (-X) = 0.8X$. This value largely deviates from both X and $-X$. The resulting absolute value of DISP will be biased by $\sim 20\%$ in the direction of lower values (see Fig. 15). Therefore, the uncertainty of the sign spoils the precision of the absolute value of DISP.

If only a single RF tree were used, this problem would automatically disappear. But in this case the performance of the RF method is much worse.

Therefore, the best way to proceed would be using 2 separate RFs: one for the absolute value of the DISP and one for its sign. In the previous sections we discussed the estimation of the absolute value of DISP combined with the standard GB by ASYM.

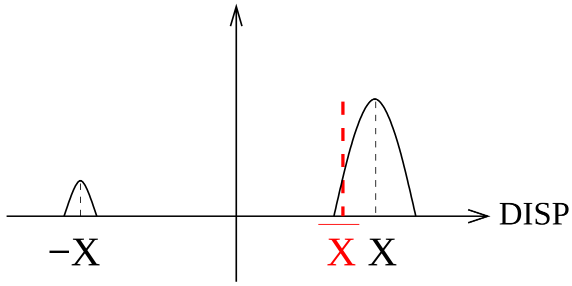

Figure 15: The reason of the spoiling of the DISP estimation by the wrong sign assignment in a fraction of RF trees.

SIZE Range	source dependent with DISP RF	
	Asym GB	RF GB
100-300	11.00 ± 0.86	10.65 ± 0.80
300-1000	2.34 ± 0.10	2.44 ± 0.10
> 1000	1.27 ± 0.12	1.39 ± 0.13

Table 5: Comparison of the sensitivity (in % C.U.) of different GB methods. Source dependent hadronness cut of 0.1 is applied for SIZE < 300 and 0.05 for the SIZE > 300. DISP is calculated by using RF with all the parameters and θ^2 cut is determined by Fig 16.

Here we describe briefly the technical details of the GB RF method, which unfortunately is not successful so far. The sign of DISP indicates whether the arrival direction of the shower is on the right side (+) or on the left side (−) of the image. So the training is performed for `sgn(MSrcPosCam.fX - MHillas.fMeanX)` value. As RF training parameters, we used `TimeGrad*sgn(M3Long)`, `M3Long`, `Asym`, and `LENGTH/WIDTH`. Training was done on ON MC to minimize the bias (In wobble MC the source stays in 0.4° in X direction and thus give more positive than negative values.)

RF GB is done as follows. Output value of the RF ranges from -1 to 1. If the value is positive, arrival direction of the shower is assigned to be on the right side of the image and the other way around. For this study slightly more data (11.3 hours, wobble mode) are used than in previous chapters. DISP RF with all parameters and source dependent hadronness were applied. The resulting sensitivity scans are shown in Fig. 16. The comparison of standard GB and RF GB is presented in Table 5.

RF GB may improve the sensitivity at low SIZEs, but it worsen a bit for larger SIZEs. If this method is applied to wobble MC, it gives a slight improvement (efficiency 65.3% → 67.9% for $SIZE < 300$ and 84.2% → 86.9% for $SIZE > 300$). However, the existing differences between the MCs and the data could mask this small effect.

With the second telescope MAGIC can work in a stereo mode. In the standard stereo analysis images from both telescopes are projected on the camera. The intersection point of the main axes of two images is a good estimation of the source direction and also gives a nearly perfect GB. If the images are close to parallel, the standard stereo reconstruction performance degrades considerably, spoiling the angular resolution. GB is also affected, but still it will be effective down to very low values of the angle between both main axes.

7 Conclusions

- The usage of RF DISP (trained with the TIME GRADIENT) in the θ^2 analysis leads to 20-30% improvement in the sensitivity with respect to the classical parametrized DISP.
- The improvement is present both in the analysis with source dependent or source independent hadronness.
- The new method provides an angular resolution as good as 0.05° for 41% containment for showers with the energy > 1TeV

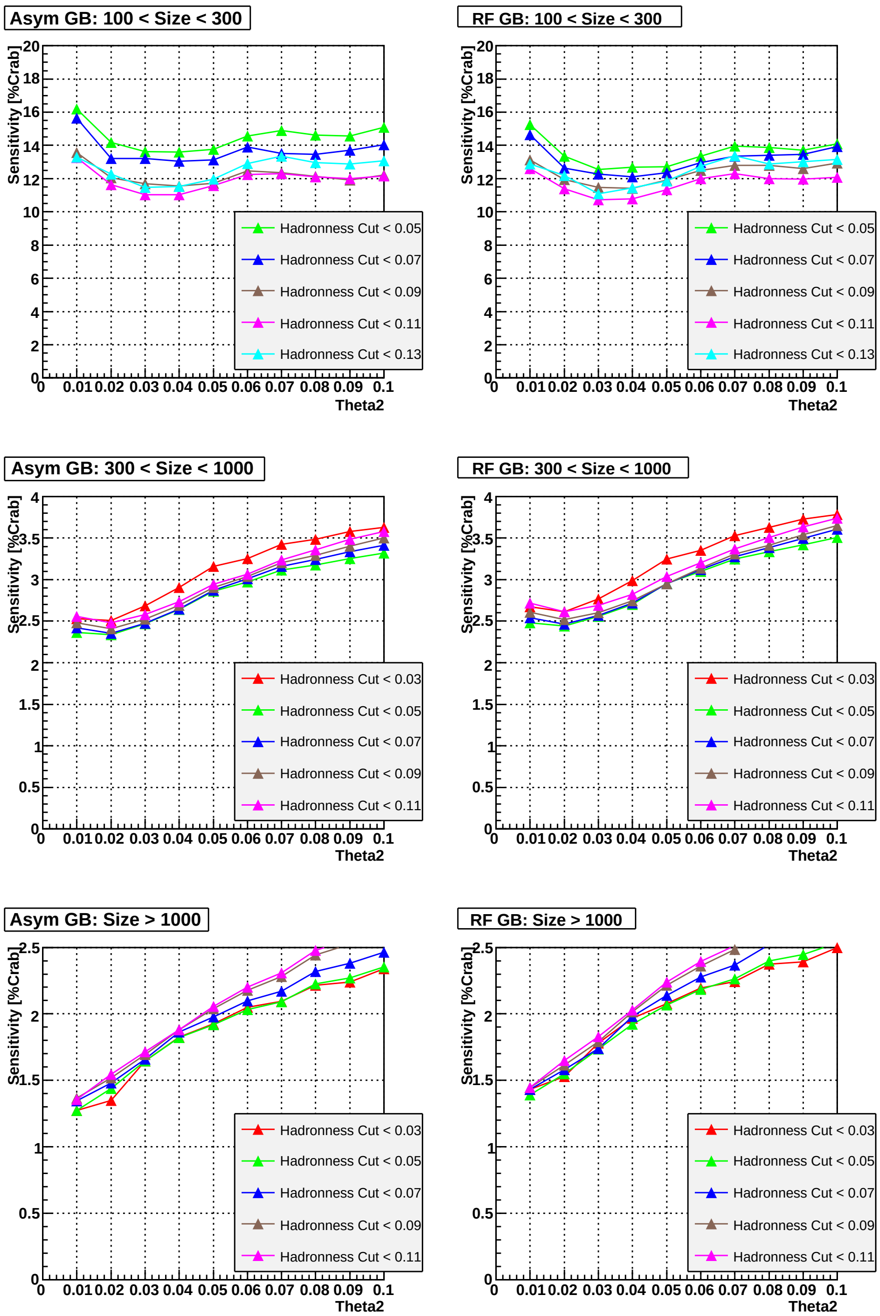

Figure 16: The comparison of the sensitivity of different GB methods ($left$: sgn(ASYM) and $right$: GB RF) in 3 SIZE bins ($top : 100 < SIZE < 300$, $middle : 300 < SIZE < 1000$ and $bottom : 1000 < SIZE$). Source dependent hadoronness and RF DISP trained with all parameters are used.

- Contrary to the classical parameterization of the DISP, the DISP estimation by using the RF for images with leakage is nearly as good as that for images without leakage.
- No further improvement is seen when using additional parameters (asymmetry, M3Long)
- Inclusion of the GB into the DISP RF reconstruction produces a severe bias in the estimation of the absolute value of the DISP, thus degrading the angular resolution.
- A separated RF used for the GB doesn't improve its precision.

8 Future plans

- Adapting DISP RF method for MAGIC II, e.g. averaging estimated DISP position from both telescopes. It can be especially useful for nearly parallel images (events with impact point close to the connection line between the telescopes).

Acknowledgements

We would like to thank Razmik Mirzoyan and Abelardo Moralejo for many useful comments and discussions which helped to improve this TDAS note.

A How to use it ?

The creation of DISP RF is done in the *osteria* program using

```
osteria -f -q \
    --zdmin=<zdmin> --zdmax=<zdmax> \
    --config=<configrc> \
    --dispgammafile=<mc_train_files> \
    --dispout=<out_dir> \
    --disprf \
    --log=<logfile>
```

This will create a file named `DispRF.root` inside directory `<out_dir>`. The option `--disprf` can be combined with all the normal *osteria* options like `--rf`, `--train-enr`, `--disp` etc.

Parameters used for the training of the DISP RF are specified in the `.rc` file e.g. :

```
# DISP RF
OsteriaLoop.MRFDisp.RFDispEstMatrix.Column0: log10(MHillas.fSize)
OsteriaLoop.MRFDisp.RFDispEstMatrix.Column1: MHillas.fWidth
OsteriaLoop.MRFDisp.RFDispEstMatrix.Column2: MHillas.fLength
OsteriaLoop.MRFDisp.RFDispEstMatrix.Column3: MHillas.fLength/MHillas.fWidth
OsteriaLoop.MRFDisp.RFDispEstMatrix.Column4: MNewImagePar.fLeakage2
OsteriaLoop.MRFDisp.RFDispEstMatrix.Column5: MHillasTimeFit.fP1Grad*
   sgn(MHillasExt.fM3Long)
OsteriaLoop.MRFDisp.RFDispEstMatrix.Column6: MPointingPos.fZd
OsteriaLoop.MRFDisp.RFDispEstMatrix.Column7: abs(MHillasExt.fAsym)
#
# True dist in deg (has to come as last column!)
#
OsteriaLoop.MRFDisp.RFDispEstMatrix.Column8: MHillasSrc.fDist*0.0033703
```

As in the case of the energy estimation with RF, last parameter have to be the true value, for which the RF optimization is done. In the case of RF DISP method, we optimize DIST. The factor 0.0033703 is to convert it from mm to degrees.

RF DISP can be applied to the data at the *melibea* level. To do it one have to run *melibea* with two additional options:

```
--calc-disp-rf \
--rfdisptree=path/to/your/DISPRF.root
```

In the output file, DISP calculated with RF method is stored in the usual container (MImageParDISP).

It is possible to calculate both parametrized DISP and RF DISP at the same time in *melibea*. In this case the parametrized DISP result will be stored in MImageParDISP, and the output tree from DISP RF will be called MImageParDispRF.

References

[1] Proceedings of ICRC2005, Pune, August 2005 E.Domingo et al., Vol. 5, 363 - 366

[2] MAGIC-TDAS 07-03