Use of discriminant analysis in search of magnetic monopoles with AMANDAII

Christian Schwenke*

Department of Physics, Humboldt University of Berlin

(Dated: September 2, 2004)

The search for relativistic magnetic monopoles produced during symmetry breaking phase transitions in the early universe, is, among other objectives, one of the science missions of the AMANDA Cherenkov telescope. We investigate the prospects of separating the monopole signal from background using Fisher's discriminant method.

I. INTRODUCTION

In 1931 Dirac showed [1], that the electric charge must be quantized if magnetic charges were introduced to electromagnetic theory. As Polyakov [3] and 'tHooft [4] 1974 examined the symmetry-breaking-behavior of gauge theories, they found that with each symmetry-breaking where a U(1)-subgroup emerges, magnetic monopoles are produced.
Since the big bang, the universe has passed several phases from grand unification of forces to that what we now know as fundamental forces: gravitation, electromagnetic-, weak- and strong interaction.
A lot of monopoles must have been produced during the phase of grand unification. Some of them should have survived and must be observable in today's universe.

A. Dirac-Monopoles

If there is a magnetic monopole of charge g centered in a coordinate system, then it would produce a radial symmetric magnetic field:

$$\vec{B} = \frac{g}{r^3}\vec{r} \tag{1}$$

Because of the radial symmetry this gives a flux through an enclosed surface:

$$\Phi = 4\pi r^2 |\vec{B}| = 4\pi g \tag{2}$$

Magnetic fields cause a phase change of quantum mechanical objects moving in a electromagnetic field. As Aharonov and Bohm showed, the phase change of an electron on closed circles is coupled to the flux through the circle:

$$\Delta\alpha = \frac{e}{\hbar c}\oint \vec{A}\cdot d\vec{l} = \frac{e}{\hbar c}\Phi \tag{3}$$

even if there is only a flux surrounded and no field lines passed. So the existence of a so-called *Dirac-String* would cause a phase change of an electron. Because of 2π-periodicity of the wavefunction ($\Delta\alpha = 2\pi n$) the electric charge must be quantized:

$$e = \frac{\hbar c}{2g} n \tag{4}$$

B. Monopoles in early universe

Based on the results of 'tHooft and Polyakov, one would expect magnetic monopoles for a unification process, because of the $SU(3)\otimes SU(2)\otimes U(1)$ symmetry of the standard model.
For a SU(5) theory one expects an unification at $10^{15} GeV$. In case of the 'tHooft-Polyakov-Monopoles we can make an estimation of the monopole mass through the mass M_X of the vector boson that causes the break of symmetry [2]:

$$m_M \simeq \frac{M_X}{\alpha_{em}} \tag{5}$$

which means, that the monopole mass is around $10^{17} GeV$. However, a large variety of GUT models do exist and the monopole masses predicted from various models range down to 10^{11} GeV [6].
A GUT-monopole-bubble as shown in FIG.1 conserved the GUT-symmetry. Inside the bubble, the monopole goes over an electroweak unification in a fermion-antifermion condensate. As shown in [10] magnetic monopoles can efficiently be accelerated along magnetic field lines of cosmic objects up to velocities close to the velocity of light.

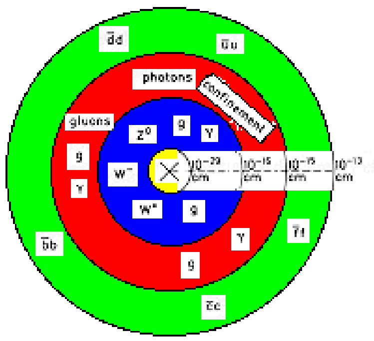

FIG. 1: SU(5) monopole-bubble

*schwenke@physik.hu-berlin.de

C. Monopole detection

Several experiments are sensitive to magnetic monopoles. Induction experiments make direct use of magnetic charge. A super-conducting loop compares the magnetic flux with those of elementary flux quantum. Ionisation experiments, as MACRO, that want to detect monopole tracks by measuring the speed of the candidate particle of its ionisation rate. Accelerator experiments perform searches for monopole-antimonopole-production-processes or virtual monopole loops, that produce two final-state-photons.
AMANDA, Baikal and in future ICECUBE are Cherenkov telescopes installed under ice or under water. Because of the strong electro magnetic coupling, the Cherenkov light emitted by a magnetic charge is enhanced by a factor $\left(\frac{g}{e\,n}\right)^2 = 8300$, compared to an electric charge, once the speed of the magnetic monopole exceeds the Cherenkov threshold ($\beta \approx 0.75$ in ice). The number of photons emitted per track length dx and wavelength interval $d\lambda$ can be calculated [7]:

$$\frac{d^2N}{dxd\lambda} = \frac{2\pi\alpha_{em}8300}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2}\right) \quad (6)$$

This will result in a light output of about 10^6 photons per cm for visible light.
As AMANDAII is placed on south pole under ice, it would be sensitive to those monopoles with $\beta > 0.75$. The high intensity of light will be the main feature of the monopole signature and can be used to discriminate against the vast background of downgoing muon bundles resulting from cosmic ray interactions in the atmosphere.

D. Upper limits of monopole-fluxes

At the moment no monopoles were found. As no events have been observed, an upper limit on the flux can be stated [5]:

$$\Phi \leq \frac{N}{A \cdot T \cdot \eta} \quad (7)$$

Here, N is the upper bound on the number of signal events present in the datasample, A represents the effective area, which is a measure for the acceptance of the detector, T is the actual observation time and η corrects for the deadtime of the detector. The flux limit is velocity-dependent, since the total amount of light slightly depends on β. Upper limits for several experiments are shown in FIG.2

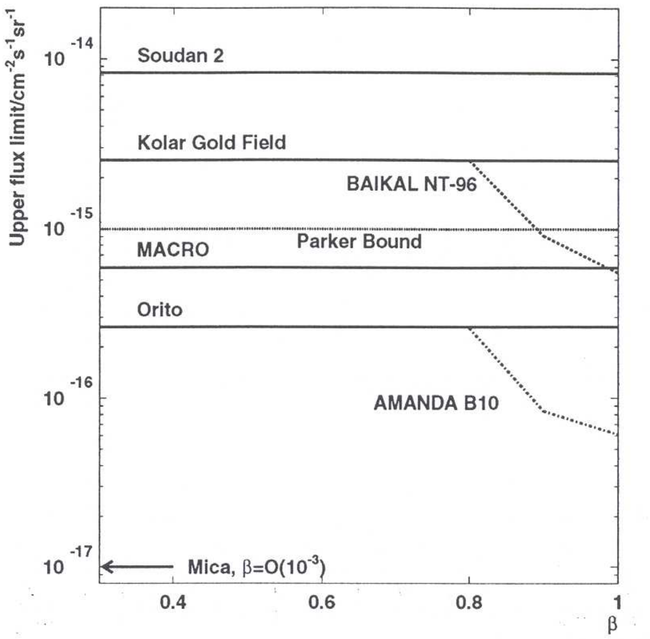

FIG. 2: Flux limits (90% C.L.) for relativistic monopoles [5]

II. DISCRIMINANT ANALYSIS

The event observables which depend on the amount of light emission are correlated with each other. One strategy to deal with linearly dependent variables was developed by the biologist Fisher. In order to obtain a good separation between classes of events (e.g. signal and background), a separating surface is searched for in feature space.
Each event would be characterized by n features $x_1, \cdots, x_n$. If each event is described e.g. by $n = 2$ features and one plots x_1 and x_2 for each class for several events, than one gets several dots in the $x_1 - x_2$-plane. To find the discriminating plane, one needs to find the linear combination of the features which has a minimal variance within each event class and a maximal difference between these event classes. [8]
The overall covariance matrix element between the variables x_k and x_l for all event classes is given by

$$T_{kl} = E[(x_k - E[x_k])(x_l - E[x_l])] \quad (8)$$

where $E[X]$ means the expectation value of some statistical variable X. This can be splitted into two parts: $T_{kl} = W_{kl} + B_{kl}$ where W describes the variance of one parameter within the event classes and B the differences in between them. For a linear optimization of the problem, one has to look at the variance of the linear combination of the input parameters

$$u(n) = \mathbf{y^T x}(n) \quad (9)$$

for both classes, depending on event number n:

$$var(u) = \mathbf{y^T T y} = \mathbf{y^T W y} + \mathbf{y^T B y} \quad (10)$$

To find the optimum, one has to find the extremum of the discriminating power

$$d = \frac{\mathbf{y^T B y}}{\mathbf{y^T T y}} \quad (11)$$

which leads to an eigenvalue problem:

$$\mathbf{T^{-1}By} = d\mathbf{y} \tag{12}$$

For two classes it can be reduced to a simple straight forward analysis [9]. For N_i events of class i and by defining

$$v_i = \frac{\sqrt{N_1 N_2}}{N_1 + N_2}(E[x_i^1] - E[x_i^2]) \tag{13}$$

one can find a solution for the weights

$$\mathbf{y} = \mathbf{T^{-1}v} \tag{14}$$

and for the discriminating power

$$d = \mathbf{v^T y} \tag{15}$$

The discriminating power can be used as a value, that represents the quality of how well the datasets have been discriminated. d can have values between 0 and 1, where 0 means no separation and 1 means perfect separation.

III. DISCRIMINANT ANALYSIS OF SIMULATED AMANDA DATA

We have tested the capabilities of a discriminant analysis using two sets of Monte Carlo simulated events in the AMANDAII detector. The speed of the simulated monopoles in the signal sample was close to c ($\beta \approx 1.0$). The background sample contained simulated downgoing air showers induced by cosmic ray interactions in the atmosphere above the detector, those being the dominant background to the magnetic monopole search.

A. Selection of variables

The data used in the analysis were produced in Monte Carlo Simulations. For latest tests we used 6400 monopole events and the same number of background events for the AMANDAII detector. The events were available in ROOT trees, which contain around 100 features (as possible observables) per event.
However for a first test, we have used the most generic observables *Nhits* , the total number of photomultiplier pulses (hits) recorded during the event, and *Nch* , the number of optical modules (channels) in which these pulses occured. Both provide a measure of the event *brightness.*
For the final analysis more variables will be used that are correlated with the light output of an event.

B. Discriminant analysis in monopole search

To implement a routine, that uses the Fisher algorithm, I used C++ with GSL [13] and wrote a class called *DA2* . The heart of the class are two methodes that calculate the weightvector and the discriminating power as described in section II. To find a suitable cut value I have implemented a routine that performs signal over square root background ratio optimization.
To test the class I wrote a class, that connected the DA2 class with the ROOT data trees. With this class it is possible to plot the events in the Nch-Nhits-plane and, by using the calculated cut and weights, to plot the cut line, that separate both event classes (FIG.3).

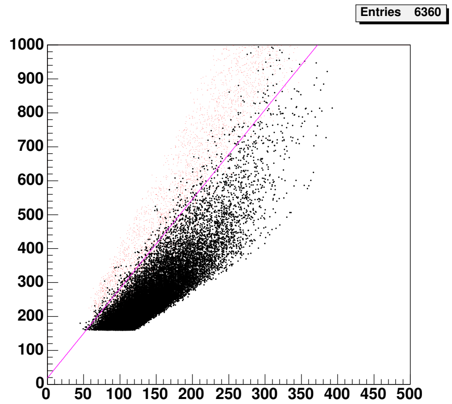

FIG. 3: Signal events (small dots) and background events(bigger dots) in Nch-Nhits-plane separated by a cut line.

As seen in the histogram FIG.4 the signal and background events can be plotted over the linear combination (9) of the input parameters, where the cut represents a "threshold" in the scalar discriminant output.

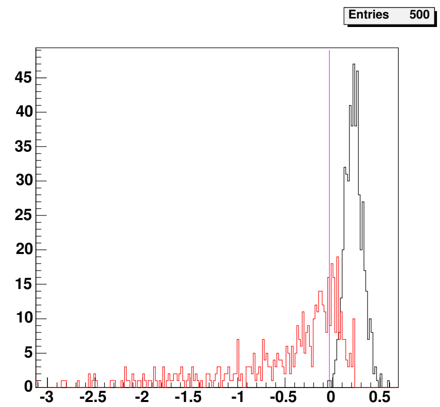

FIG. 4: Linear combinations of the input parameters (dotted line is signal, solid line is background) with calculated weights and suitable cut (thin vertical line).

If more than two features are used, the separation of the distributions tend to be better and a little increase of the discriminating power can be seen. Instead of a cutline we have a hyperplane in the feature space and is not so easy to visualize.

C. Nonlinear approach

As a possible improvement I implemented a nonlinear extension of the DA2 class, where additionally to the original input variables, also second order combinations are used. That means for the case of two variables x_1, x_2, the products x_1^2, x_2^2 and $x_1 x_2$ are included as event features. Since the separating surface is not limited to being a straight line (or a hyperplane, in case of feature spaces with higher dimensions) one expects a more efficient separation. (As stated in [11] no further improvement is expected from including higher order terms.)

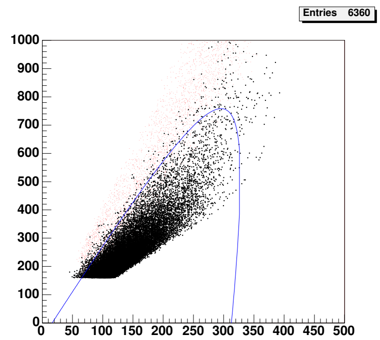

FIG. 5: Signal events (small dots) and background events(bigger dots) in Nch-Nhits-plane separated by a cut hyperbola.

Here the cutline is a hyperbola with two solutions. The one cutting the background is plottet in FIG.5. Even for the two trivial parameters Nch and Nhits the distributions of the scalar discriminant output show a much better separation of the two distributions (FIG.6) than for the linear approach.

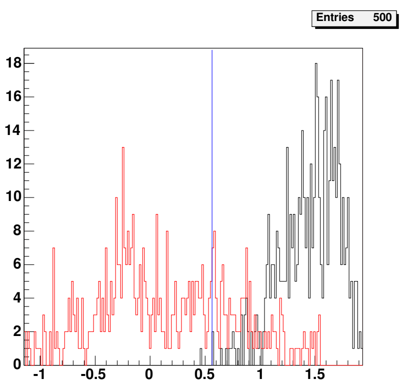

FIG. 6: Quadratic combinations of the input parameters (dotted line is signal, solid line is background) with calculated weights and suitable cut (thin vertical line).

D. Testing discriminant analysis

For testing I wrote a program that calculated the cut and weights with the DA2 class for training samples of several sizes ranging from 10 to 1000 events and used them to operate on a different Monte Carlo sample of 6400 signal and background events (FIG.7).

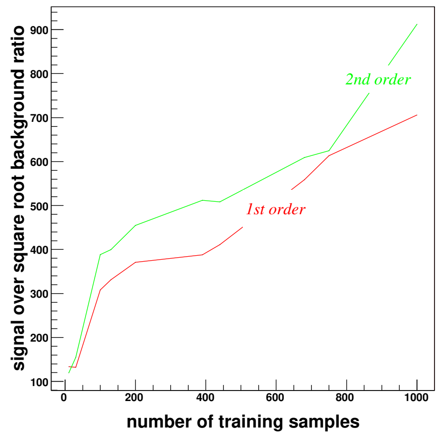

FIG. 7: Linear vs. quadratic combinations of 6400 events over several training samples

The upper line shows the behavior of the quadratic approach, the other the behavior of the first order calculations. As one can also see in FIG.8/9 the background rejection for the quadratic case is much better than for the linear approach. The signal passing is a little worse for the quadratic samples compared to the linear ones, but with an increasing signal over square root background ratio.

Training sample events	50	100	300	1000
Signal after cut	5561	5990	5941	5332
Background after cut	186	378	388	57
Sig. bg. ratio	407,75	308,09	301,61	706,24
Discriminating power	0,41099	0,36929	0,37212	0,36847

FIG. 8: Results of the linear Discriminant Analysis calculations for several numbers of events in the training sample.

Training sample events	50	100	300	1000
Signal after cut	5891	5724	5636	4828
Background after cut	358	217	148	28
Sig. bg. ratio	311,349	388,57	463,276	912,406
Discriminating power	0,72789	0,69802	0,66470	0,65433

FIG. 9: Results of the quadratic Discriminant Analysis calculations for several numbers of events in the training sample.

IV. POSSIBLE IMPROVEMENTS

Because of the much bigger variances for the quadratic approach, we suspect that the second order calculation has numerical instabilities. These can possibly be overcome by scaling all variables back, so that their variances are equal to 1.0 This would make the inversion of the covariance matrix "saver".
Further, one can make a preselection on the training samples. Including only these events that lie within a 3σ range of around their mean makes the algorithm robust against statistical fluctuations in the training samples.

V. CONCLUSIONS

I investigated, if Discriminant Analysis could be a useful tool for monopole search with AMANDAII. For me the method looks promising for linear and quadratic combinations of observables in simulated data, but it needs further refinement in order to make quantitative statements of the effectiveness of the method.

VI. ACKNOWLEDGEMENTS

I'm greatly indebted to the AMANDA group in Zeuthen, where I got the support needed to write the DA-tool and fight against the frustration by searching bugs. A special thanks to Henrike Wissing, my superviser, for her support and the MC-data, where my studies are based on. Furthermore thanks to all summer students of DESY Zeuthen 2004 to have an unforgetable time together. I probably found new friends and will never forget the parties, the trips to Berlin and Hamburg. And a very special thanks to Edwin, my Mexican friend and colleague at AMANDA group, without him, I would never have been in Wacken, the greatest festival in the world.

[1] P.A.M.Dirac, Proc. Roy. Soc. A 133 (1931) 60
[2] H.V.Klapdor-Kleingrothaus/K.Zuber, *Teilchenastrophysik* (Teubner 1997)
[3] A.Polyakov, jetp Lett. 20 (1974) 194
[4] G.'tHooft, Nucl. Phys. B 79 (1974) 276
[5] P.Niessen, Diss.,*Search for relativistic magnetic monopoles with AMANDA*
[6] Thomas W. Kephart, Thomas J., Weiler Astropart.Phys. 4 (1996) 271-279
Thomas J. Weiler, Thomas W., Kephart Nucl.Phys.Proc.Suppl. 51B (1996)
[7] D.R.Tompkins, Phys. Rev. B 138 (1965) 1
[8] V.Blobel/E.Lohrmann, *Statistische und numerische Methoden der Datenanalyse* (Teubner 1998)
[9] T.Schweizer, Ph.D. Thesis, *Analysis of Large Gamma Ray Flares of Mkn421*
[10] S.D.Wick/T.W.Kephart/T.J.Weiler/P.L.Biermann, Astropart. Phys. **18 (2003) 663-687, astro-ph/0001233**
[11] T.G.M.Malmgren/K.E.Johansson, Nucl. Phys. A 401 (1997) 409-413
[12] L.Lonnblad/C.Peterson/T.Rognvaldsson, Comput. Phys. Commun. 70 (1992) 167
[13] R.Stallman, *Open Sources* (1996) www.gnu.org/software/gsl