Dark Photon Search at BaBar

Ross Greenwood

Office of Science, Science Undergraduate Laboratory Internship Program

Massachusetts Institute of Technology

SLAC National Accelerator Laboratory
Menlo Park, California

July 20, 2011

Prepared in partial fulfillment of the requirement of the Office of Science, Department of Energy's Science Undergraduate Laboratory Internship under the direction of Dr. Arafat Gabareen Mokhtar, SLAC Nation Accelerator Laboratory.

Participant: ____
Signature

Research Adviser: ____
Signature

INTRODUCTION

DISAGREEMENT BETWEEN THE OBSERVATIONS OF THE MOTIONS OF DISTANT GALAXIES AND CALCULATIONS OF THEIR MASS DISTRIBUTIONS BASED ON VISIBLE MATTER SUGGEST THAT LIGHT AND

TABLE OF CONTENTS

1. To
2. Be
3. Included

ABSTRACT

To be included

INTRODUCTION

Based on observational evidence from astronomical surveys of the internal motions of distant galaxy clusters, it is believed that electromagnetically interacting matter contributes to only about 4% of the mass-energy of the universe. Current theory suggests that the remainder is made up of two additional components in a ratio of approximately three to one, respectively: dark energy, which is responsible for accelerating the metric expansion of space, and dark matter – nonbaryonic matter particles with nonzero mass (inferred from its gravitational effects on nearby visible matter) that may experience the weak or strong nuclear forces, or yet undiscovered interactions.

One candidate constituent of dark matter is a theoretical dark photon (symbol: A'). The dark photon is a massive boson that is neutral under the Standard Model forces and whose existence is postulated in several models of supersymmetry. It may serve as the force-carrying gauge boson for particle interactions unique to the dark sector, or even provide a means for interaction between dark matter and Standard Model particles. Theory suggests (and experiment has not excluded) the possibility of virtual dark photon production in electron-positron annihilation events. The electron-positron pair may annihilate directly to a dark photon, which theory suggests would decay to a fermion pair. Alternatively, the e^+e^- pair annihilates to a virtual Standard Model photon, which produces a dark photon via an intermediate virtual fermion loop. This process may account for the discrepancy between the theoretical ratio of number of positrons to number of electrons found in cosmic rays and its value measured by the PAMELA experiment [cite] and others.

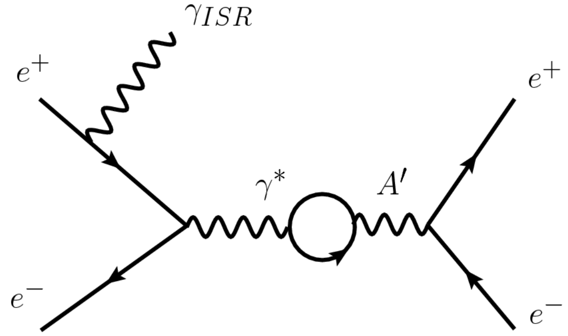
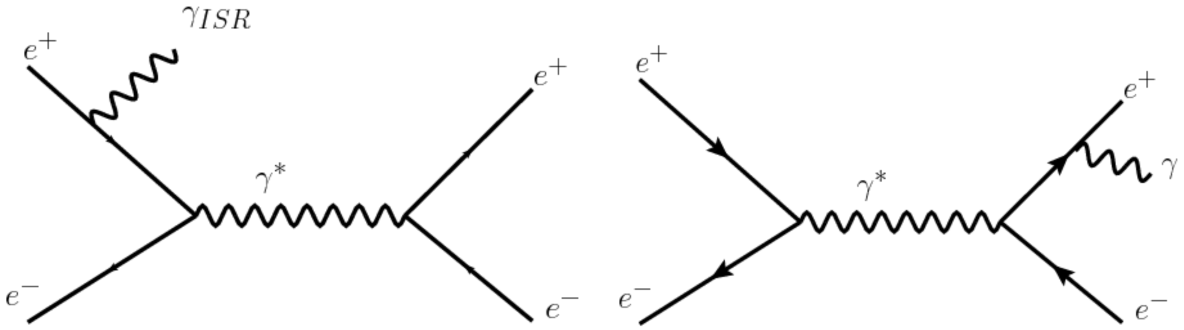


Illustration 1: Feynman diagram illustrating the process $e^+e^-\rightarrow\gamma_{ISR}\gamma^*$, $\gamma^*\rightarrow A'$, $A'\rightarrow e^+e^-$

The BaBar experiment (1999 – 2008) recorded asymmetric collisions of electrons and positrons, supplied by the linear accelerator at the SLAC National Accelerator Laboratory. Activity at BaBar focused on the production and study of B mesons, hadrons containing a quark and an anti-quark, one of which has bottom flavor (b or $\bar{b}$). Collisions took place at center-of-mass energies corresponding to the mass resonances of the upsilon meson ($b\bar{b}$), $m_{\Upsilon(nS)}$ for n = 1,2,3,4.

We searched the experimental data collected from the BaBar detector for evidence of a dark photon or similar new particle, probing a mass range from 100 MeV to 9.5 GeV. Events featuring electron-positron annihilation preceded by emission of an energetic photon ($E_\gamma \geq 1$ GeV) from either lepton. By requiring initial state radiation (ISR), we access a range of values for the dielectron mass m_{ee}, the relativistic mass of the electron-positron system at the time of annihilation. The chief irreducible backgrounds with the same final state as the sample are radiative Bhabha scattering ($e^+e^-\rightarrow e^+e^-$) and two-photon production ($e^+e^-\rightarrow\gamma\gamma^*$, $\gamma^*\rightarrow e^+e^-$).



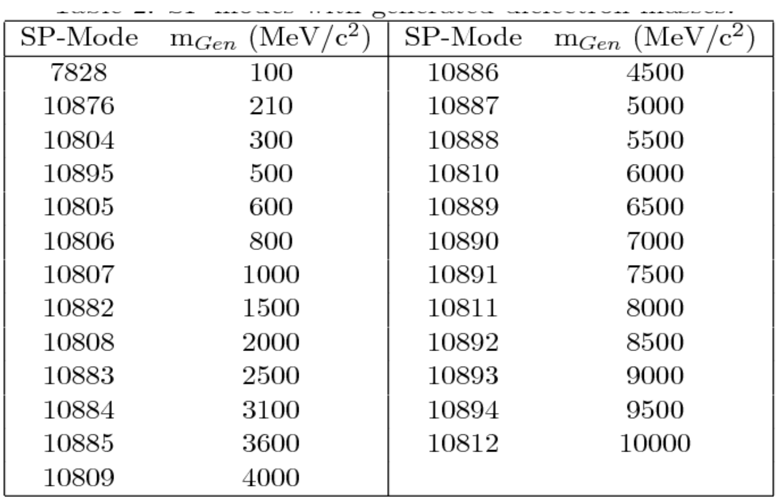
If a dark photon exists and we implement the appropriate selection criteria, the particle will manifest in the data as a resonance peak in the candidate event count plotted against the dielectron mass. The mean of this peak corresponds to the mass of the new particle. The goal of this analysis is to optimize the significance $s(B^{1/2} + 3/2)^{-1}$ for candidate events showing evidence of dark photon production, where s is the number of signal events and B is the number of background events. We begin with a blind analysis using a small portion of the available data, the Run 6 data set, composed of events with a dielectron mass corresponding to the Υ(4S) resonance (10.58 MeV/c^2). After perfecting our optimization techniques, we will unblind the analysis to include all compatible data from Runs 1–7.

MATERIALS AND METHODS

Two terms are used to describe mass in this analysis: *generated mass* and *reconstructed mass*. The generated mass is a single parameter submitted to the event generator as simulated rest mass of the dark photon with a resonance of zero width. The reconstructed mass is the dielectron energy calculated from the kinematics of a generated or experimental event.

We began with a data set containing about 2.3% of Run 6, which we initially reduced using selection cuts. We demand exactly two lepton tracks and at least one photon with energy greater than 1 GeV in the final state. Up to ten additional photons are allowed with energies less than 100 MeV. In case of multiple candidate pairs of lepton tracks are recorded, the event with the best vertex fit (best χ^2) is used, such that there is only one candidate per event. Our analysis is focused on dark photon events with a dielectron final state, though $\mu^+\mu^-$, $\pi^+\pi^-$, and $e^{+/-}\mu^{-/+}$ final states are also considered for the purpose of understanding the background. Using the ROOT software developed at CERN and its Toolkit for Multivariate Analysis (TMVA) package, we trained a program to recognize signal events based on Monte Carlo simulations of the target signal generated using the “GfiGeTUP” simulator. The analysis thus far can be summarized in the following five steps:

1. Monte Carlo
2. Resolution
3. TMVA training
4. Signal efficiency
5. Sensitivity and branching ratio

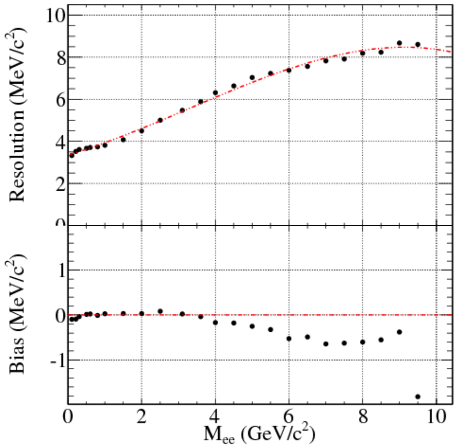


SP-Mode	m_{Gen} (MeV/c^2)	SP-Mode	m_{Gen} (MeV/c^2)
7828	100	10886	4500
10876	210	10887	5000
10804	300	10888	5500
10895	500	10810	6000
10805	600	10889	6500
10806	800	10890	7000
10807	1000	10891	7500
10882	1500	10811	8000
10808	2000	10892	8500
10883	2500	10893	9000
10884	3100	10894	9500
10885	3600	10812	10000
10809	4000		

Table 1: Monte Carlo Simulation Product (SP) Modes

Twenty-five sets of Monte Carlo with discrete dark photon generated mass values ranging from 100 MeV to 9.5 GeV were produced. Bhabha scattering and two-photon production in which one photon decays to e^+e^- were also simulated, corresponding to SP modes 2400 and 1074, respectively. The resolution quantifies the spread of the reconstructed dark photon mass produced by a Monte Carlo simulation. During TMVA training, a program is trained to recognize signal events based those generated by Monte Carlo simulations and assigns a set of values characterizing the probability of an event being part of the signal. The signal efficiency ε(m) is the ratio of the number of reconstructed events to the number of generated events. The sensitivity defines, in theory, to what extent we will be able to place limits on physical parameters like the A' mass.

The reconstructed mass histogram for each SP mode was fit the sum of two Crystal Ball functions with equal widths and means but opposite tails (free to vary). The width was taken as the resolution of each SP mode. The twenty-five resolution values were plotted with the generated mass of each SP mode and fit to a second-degree polynomial, resulting in a continuous resolution function $\sigma(m)$.

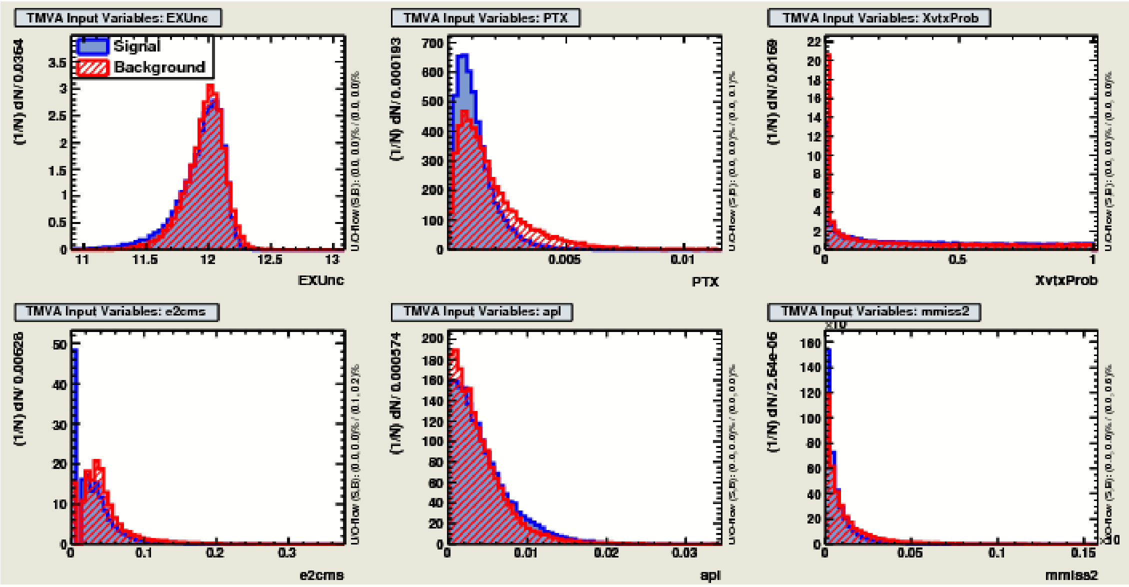


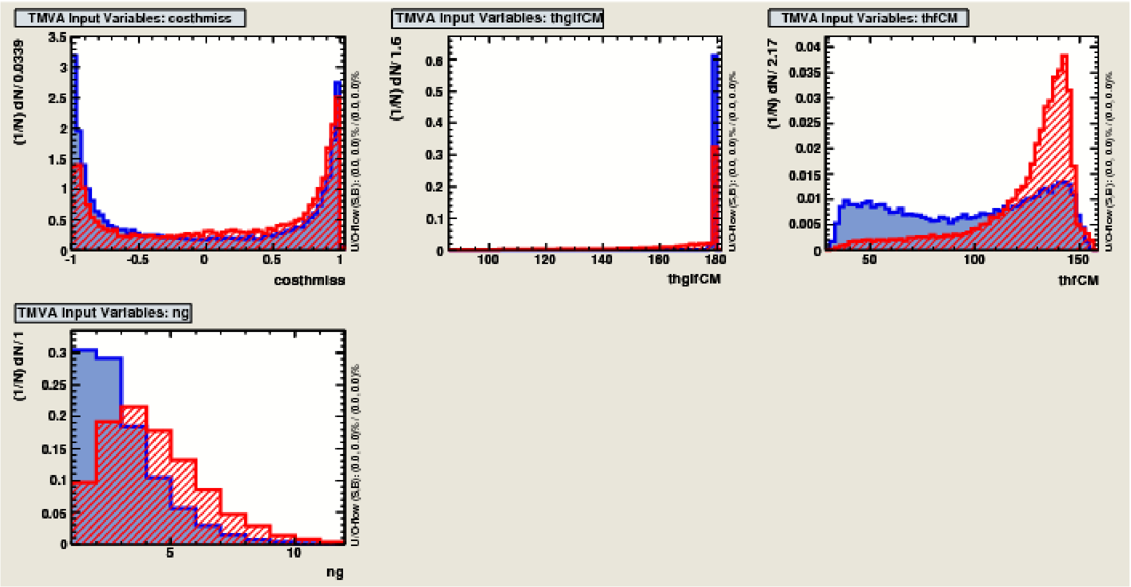
(Top) Resolution and second-degree polynomial fit σ(m). *(Bottom) Bias -- mean of Crystal Ball fit function.*

TMVA is a programming package that runs with ROOT; it was created to simplify the implementation of machine learning algorithms for classification and regression analysis of physical data. We used TMVA to characterize signal events based on ten variables:

[Unless otherwise stated, kinematic calculations were performed in the center-of-momentum frame of the dielectron.]

- *EXUnc* : Unconstrained total energy of $\gamma_{ISR}e^+e^-$ calculated in the laboratory frame
- *PTX* : Transverse momentum of $\gamma_{ISR}e^+e^-$
- *XvtxProb* : Vertex probability
- *e2cms* : Energy of the 2nd highest-energy photon
- *apl* : Positive angle (cosine) between the γ_{ISR} and the plane containing the lepton tracks
- *mmiss2* : Squared missing mass
- *costhmiss* : Polar angle (cosine) of the missing momentum vector
- *thglfCM* : Angle between γ_{ISR} and the fastest lepton track
- *thfCM* : Polar angle of the negative track [CM]

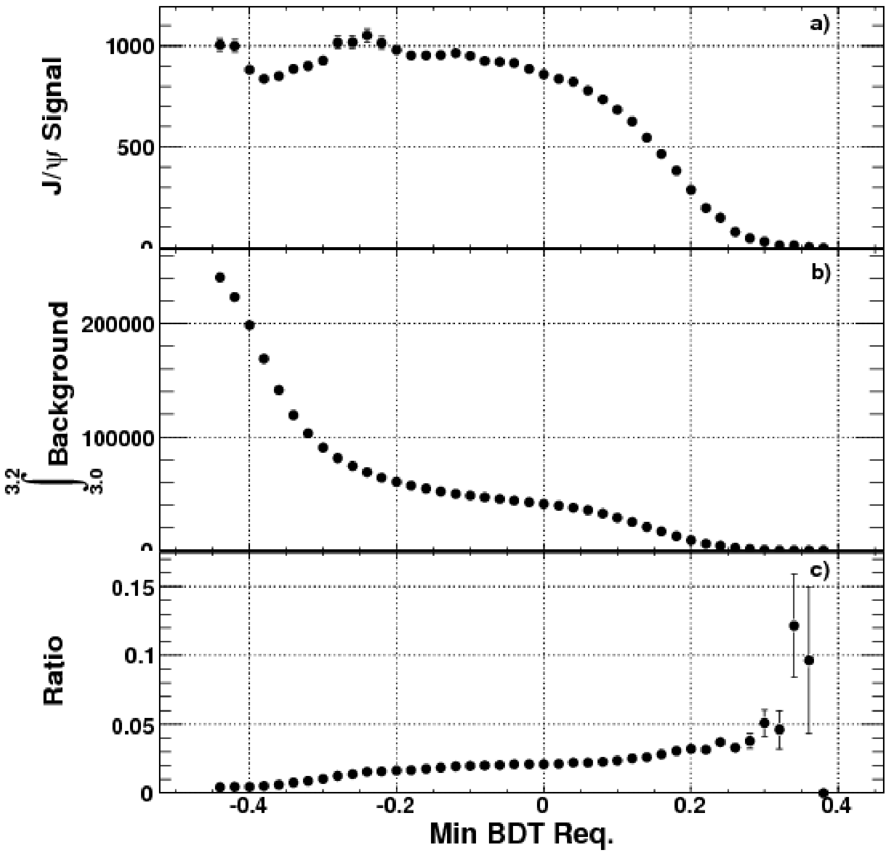




Performance of TMVA variables for Monte Carlo signal (blue) and empirical background (red) with generated mass of 100 MeV

A TMVA::Factory object is trained to recognize signal event candidates through comparison of sample target events generated in Monte Carlo to a background of the Run 6 data. A TMVA::Reader object is then instantiated to read real data and use the training results to characterize the probability that the data shows evidence of a dark photon candidate with a particular mass. The TMVA classification methods employed in this analysis are Cuts, Fisher, Likelihood and Boosted Decision Tree (BDT).

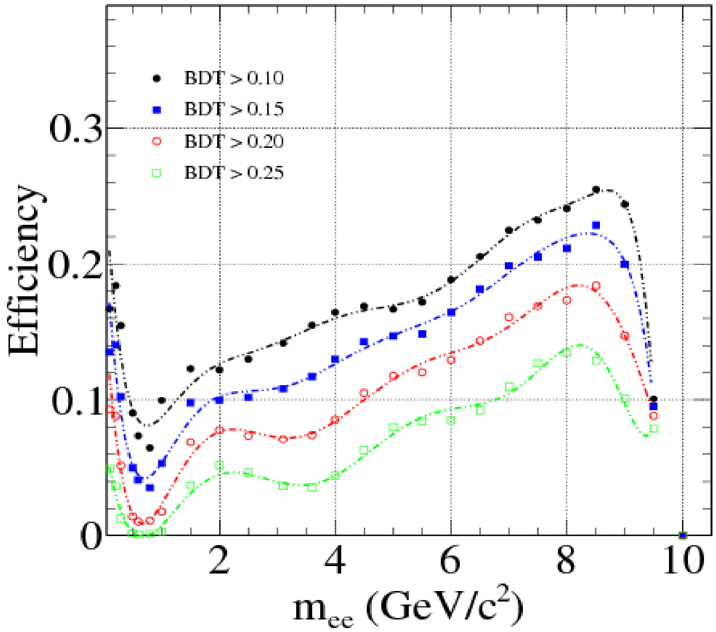
We ran a separate TMVA analysis for each Monte Carlo SP mode, and interpolated the results to produce a BDT probability density function for the full mass range. We then determined the most appropriate BDT-based selection cut using the J/ψ resonance peak as a reference. We varied the maximum accepted value of the BDT parameter in our selection cut of the portion of the Run 6 data set, and calculated the signal-to-background ratio in the vicinity of the J/ψ peak. We found the best combined effect of background reduction and signal efficiency when requiring a BDT value greater than 0.2.



Determination of BDT cutoff using J/ψ

The efficiency was calculated for each Monte Carlo simulation product as

$$\varepsilon(m) = \frac{N_{rec}}{N_{gen}}$$



We extracted a continuous efficiency function of mass by fitting to a third order polynomial.

The sensitivity was calculated as the expected error on the branching ratio σ_{BR}.

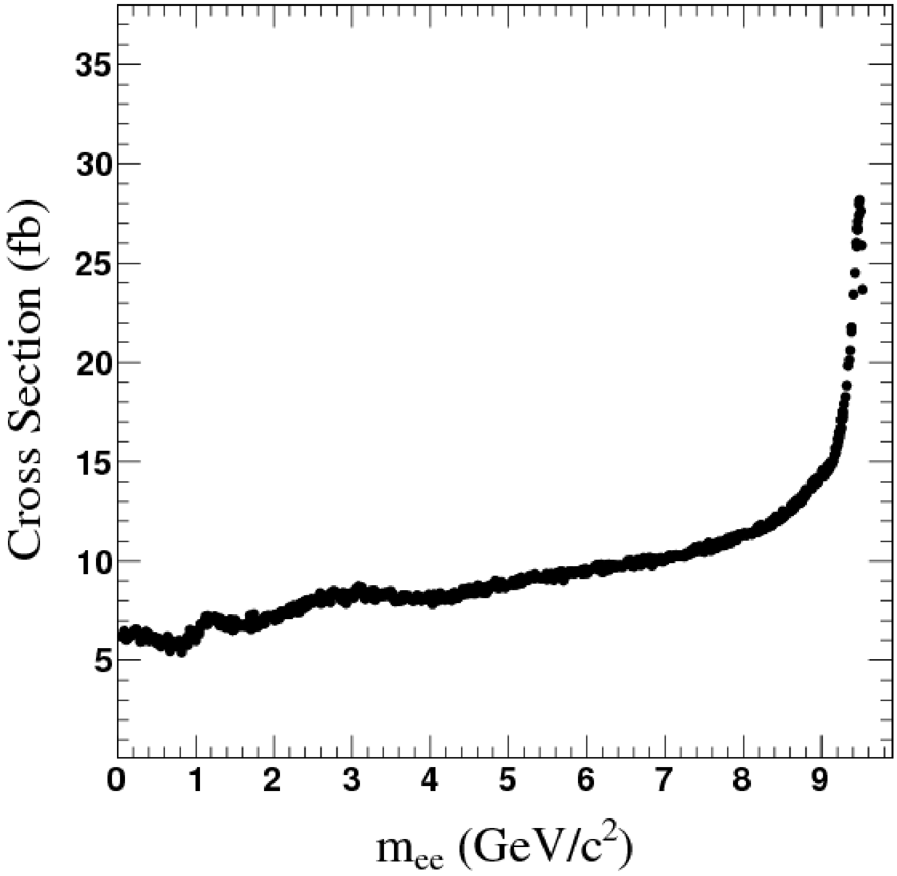
$$\sigma_{BR}(m) = \frac{N_{bkg}}{2\sqrt{\varepsilon(m) \cdot N_{total}}}$$

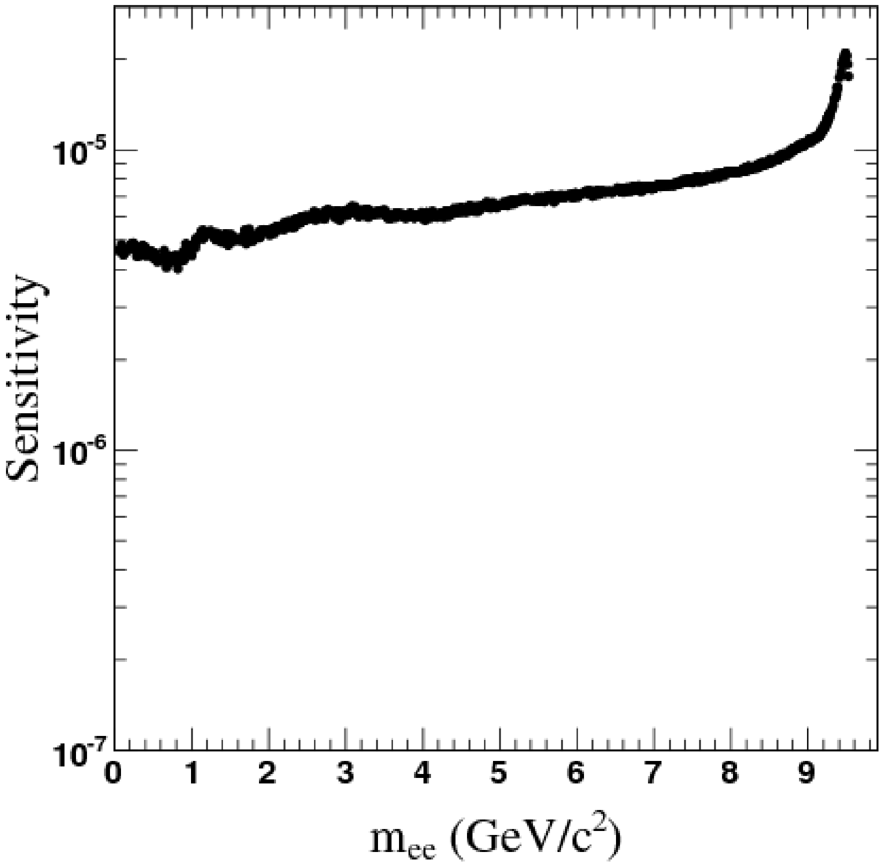
N_{bkg} is the number of background events with reconstructed masses close to m, within $1.5 \cdot \sigma(m)$.

N_{total} is the total number of electron-positron annihilation events.

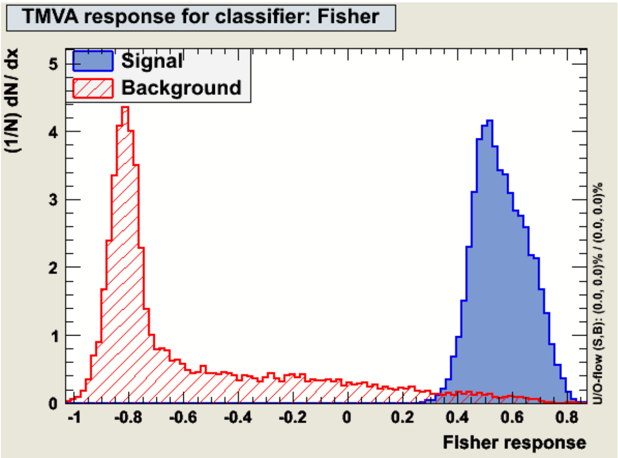
RESULTS

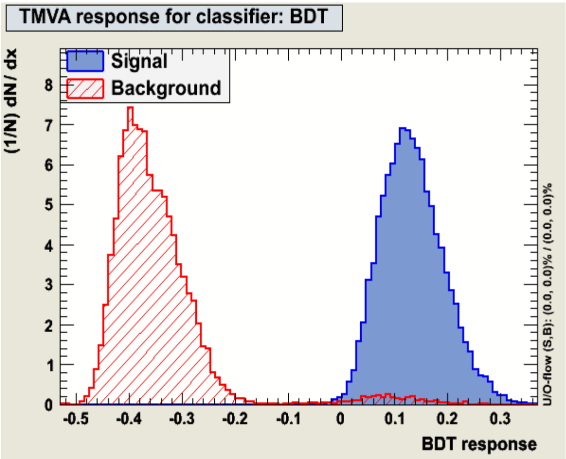
Our results for the sensitivity and branching ratio are reproduced below.





We were able to achieve good separation between signal and background using the cut of 0.2 as the minimum allowed BDT value.





DISCUSSION AND CONCLUSIONS

The analysis is still underway and has yet to produce conclusive results. The next step in this analysis is to fit the reconstructed A' mass histogram. We will first fit the background, including recognized features (i.e. J/ψ) and assuming that no other signal is present. We will then iterate through the range of possible reconstructed A' masses, attempting to fit the residuals to a Gaussian function at each mass value and producing a measure of the likelihood that a discovery resonance peak sits atop the background at that mass value.

[what to cover here: problems we encountered, sources of error, significance of results, broader impact, how conclusions will be made once analysis is finished]

ACKNOWLEDGMENTS

To be included

REFERENCES

To be included

APPENDIX A: DETERMINATION OF BDT CUTOFF USING J/PSI