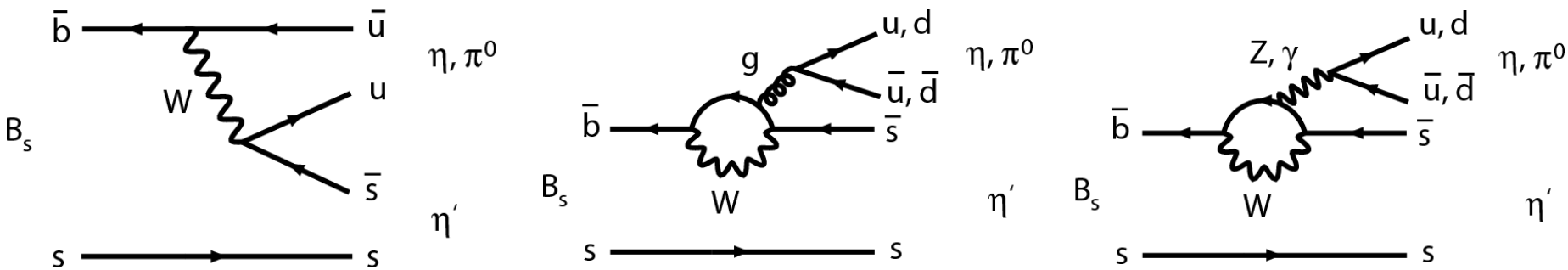
# Search for $B_s \to \eta\eta'$ and $B_s \to \pi^0\eta'$ in BELLE data

Anthony Zummo
Vladimir Savinov
University of Pittsburgh
August 19, 2016

# Motivation

- Proceeds primarily through b → u transitions or b → s transitions, which are sensitive to new physics
- Updating $B \rightarrow PP, VP$ decays in the framework of flavor symmetry
  - Hai-Yang Cheng, Cheng-Wei Chiang, and An-Li Kuo, Phys. Rev. **D** 91, 014011 (2015)
  - Predicts a branching fraction for $B_s \rightarrow \eta\eta'$ of $(33.47 \pm 3.64) \times 10^{-6}$
- A global SU(3)/U(3) flavor symmetry analysis for $B \rightarrow PP$ with $\eta - \eta'$ Mixing
  - Yu-Kuo Hsiao, Chia-Feng Chang and Xiao-Gang He, arXiv:1512.09223 (2016)
  - Shows how to use SU(3)/U(3) flavor symmetry framework to extract CP-violating parameters from data on $B_{s,d} \rightarrow \eta\eta, \eta\eta', \text{and } \eta'\eta'$ decays. Only $B_s \rightarrow \eta'\eta'$ has been measured so far.



# Signal Monte Carlo Generation

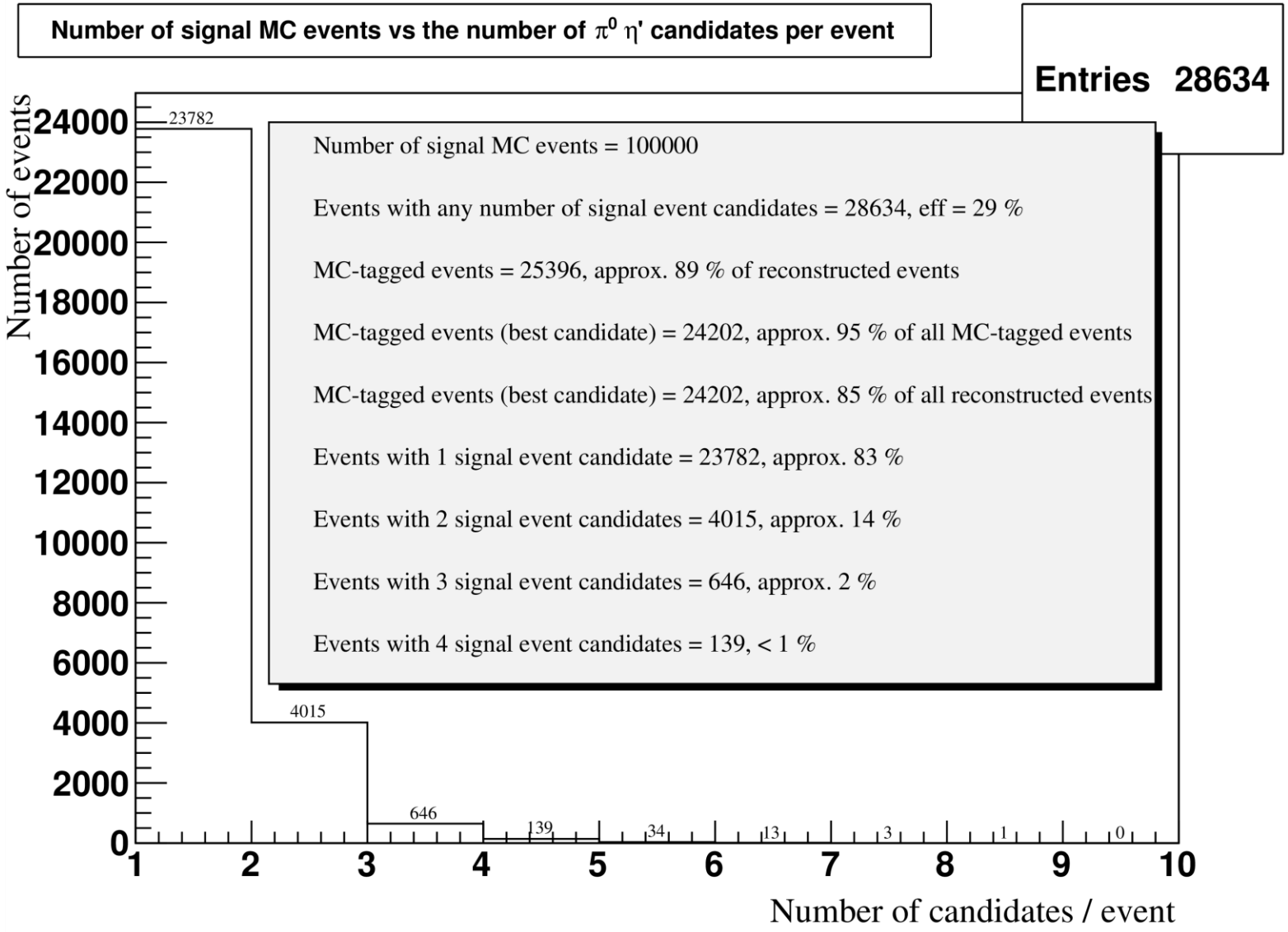
- Generated (using **mcproduzh** scripts) 100,000 $B_s \to \eta\eta'$ events for several decay channels of $\eta$ and $\eta'$ and 100,000 $B_s \to \pi^0\eta'$ events
- Generated in proportion to cross sections for $B_s^*\overline{B_s^*}$, $B_s^*\overline{B_s}$, $B_s\overline{B_s}$ production at $\Upsilon(5S)$ and distributed over BELLE data samples 53, 67, 69, and 71 according to data statistics
- We have done MC studies of $B_s \to \pi^0\eta'$ and several channels of $B_s \to \eta\eta'$
- Currently focusing on $\eta' \to \rho^0\gamma,\ \eta \to \gamma\gamma$

```
Decay My-eta'
1.000     My-rho0  gamma                    SVP_HELAMP  1.0 0.0 1.0 0.0;
#1.000     pi+ pi- My-eta2                   PHOTOS PHSP;
#1.000     My-pi0 My-pi0 My-eta2             PHSP;
Enddecay

Decay My-eta1
1.000          gamma gamma             PHSP;
#1.000         My-pi0 My-pi0 My-pi0     PHSP;
#1.000         pi- pi+ My-pi0           PHOTOS ETA_DALITZ;
Enddecay

Decay My-eta2
1.000          gamma gamma             PHSP;
#1.000         My-pi0 My-pi0 My-pi0     PHSP;
#1.000         pi- pi+ My-pi0           PHOTOS ETA_DALITZ;
Enddecay

Decay My-pi0
1.000 gamma            gamma        PHSP;
Enddecay

Decay My-rho0
0.989353   pi+ pi-                       PHOTOS VSS;
0.009900   pi+ pi- gamma                 PHOTOS PHSP;
#0.000610   pi0 gamma                     VSP_PWAVE;
##0.00038    eta gamma                    VSP_PWAVE;
#0.000045   pi0 pi0 gamma                 PHSP;
#0.000046   mu+ mu-                       PHOTOS     VLL;
#0.000046   e+ e-                         PHOTOS     VLL;
##0.000018   pi+ pi+ pi- pi-               PHSP;
Enddecay
```

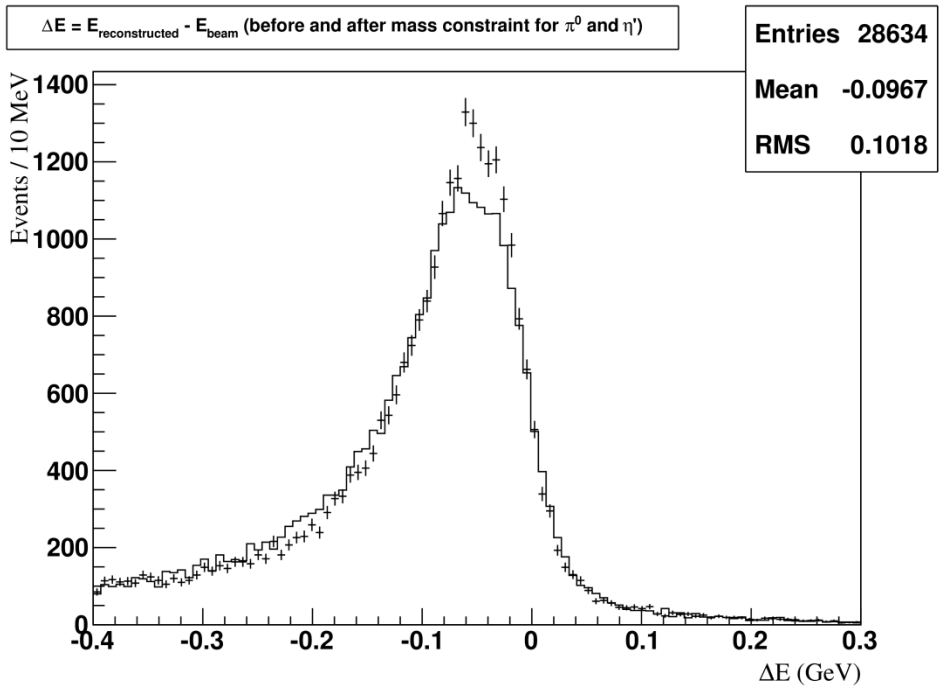
# $B_s \to \pi^0 \eta'$ selection criteria for signal MC

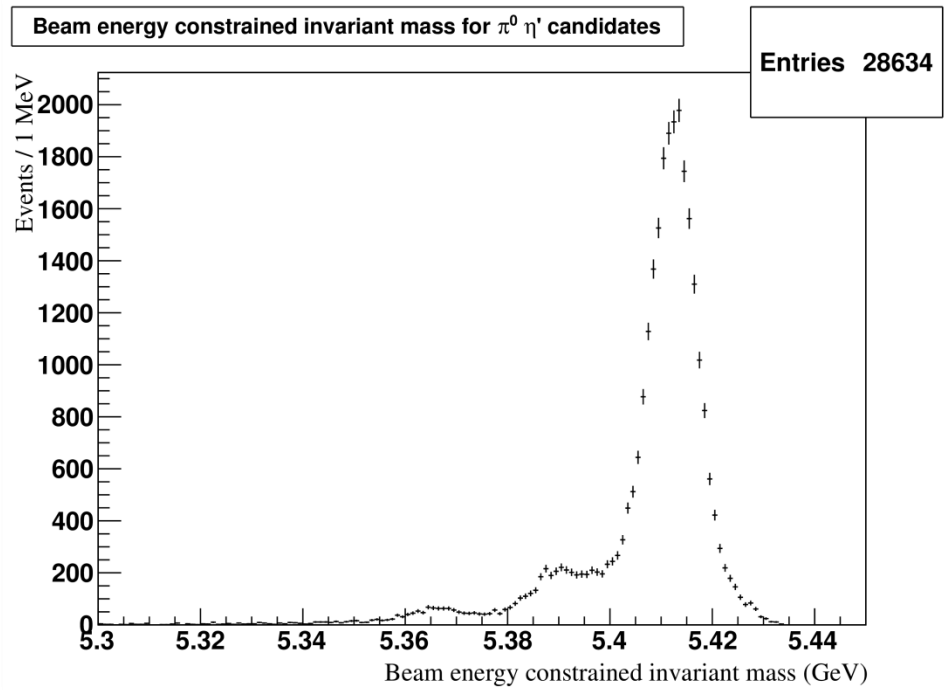
| | |
|---|---|
| $\gamma$ | $E > 100\ MeV$ |
| $\pi^0$ | $.11\ GeV/c^2 \leq M_{\gamma\gamma} \leq .16\ GeV/c^2$ |
| $\pi^\pm$ | $dr \leq .2\ cm$<br>$\lvert dz \rvert \leq 4\ cm$<br>$p_T \geq\ .1\ GeV/c$<br>$R_{K,\pi} < .6$ |
| $\rho^0$ | $.435\ GeV/c^2 \leq M_{\pi^+\pi^-} \leq 1.005\ GeV/c^2$ |
| $\eta'$ | $.92\ GeV/c^2 \leq M_{\rho^0\gamma} \leq .98\ GeV/c^2$ |
| $B_s$ | $M_{bc} \geq 5.3\ GeV/c^2$<br>$-0.4\ GeV \leq \Delta E \leq 0.3\ GeV$ |

# Signal candidates per event for $B_s \rightarrow \pi^0 \eta'$ signal MC

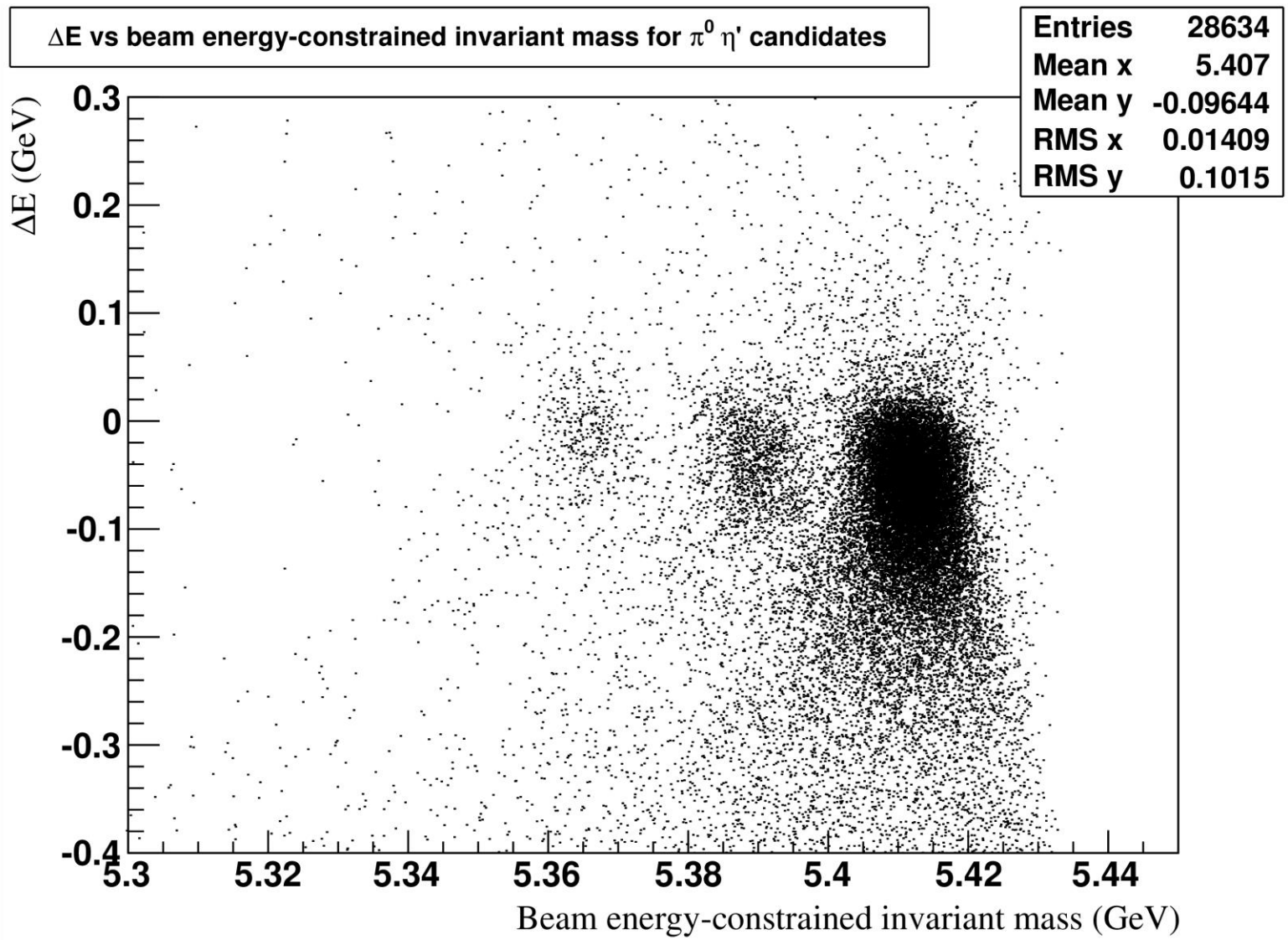


# $B_s \rightarrow \pi^0 \eta'$ signal MC: no MC tagging, best $\chi^2$ candidate only



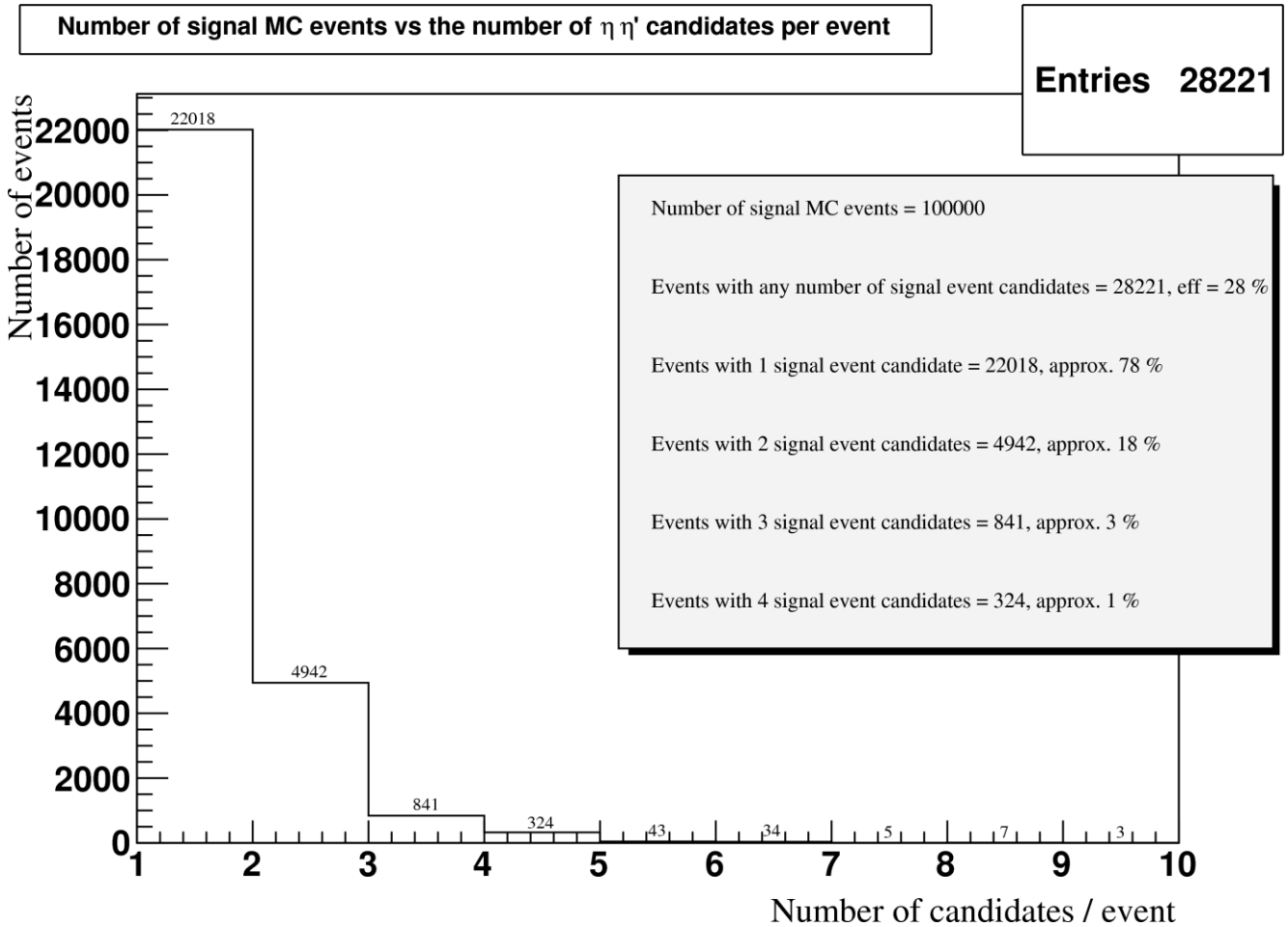
# $B_s \to \pi^0 \eta'$ signal MC: no MC tagging, best $\chi^2$ candidate only



# Selection criteria for signal event candidates

| | |
|---|---|
| $\gamma$ | $E > 100\ MeV$ |
| $\eta$ | $.52\ GeV/c^2 \leq M_{\gamma\gamma} \leq .57\ GeV/c^2$ |
| $\pi^{\pm}$ | $dr \leq .2\ cm$<br>$\lvert dz \rvert \leq 4\ cm$<br>$p_T \geq\ .1\ GeV/c$<br>$R_{K,\pi} < .6$ |
| $\rho^0$ | $.435\ GeV/c^2 \leq M_{\pi^+\pi^-} \leq 1.005\ GeV/c^2$ |
| $\eta'$ | $.92\ GeV/c^2 \leq M_{\rho^0\gamma} \leq .98\ GeV/c^2$ |
| $B_s$ | $M_{bc} \geq 5.3\ GeV/c^2$<br>$-0.4\ GeV \leq \Delta E \leq 0.3\ GeV$ |

# Efficiency and Number of Candidates per Event in Signal MC




- Candidates are sorted by the lowest $\chi^2$ of $\eta$ and $\eta'$

# Best Candidate Selection based on $\chi^2$ of mass constraints

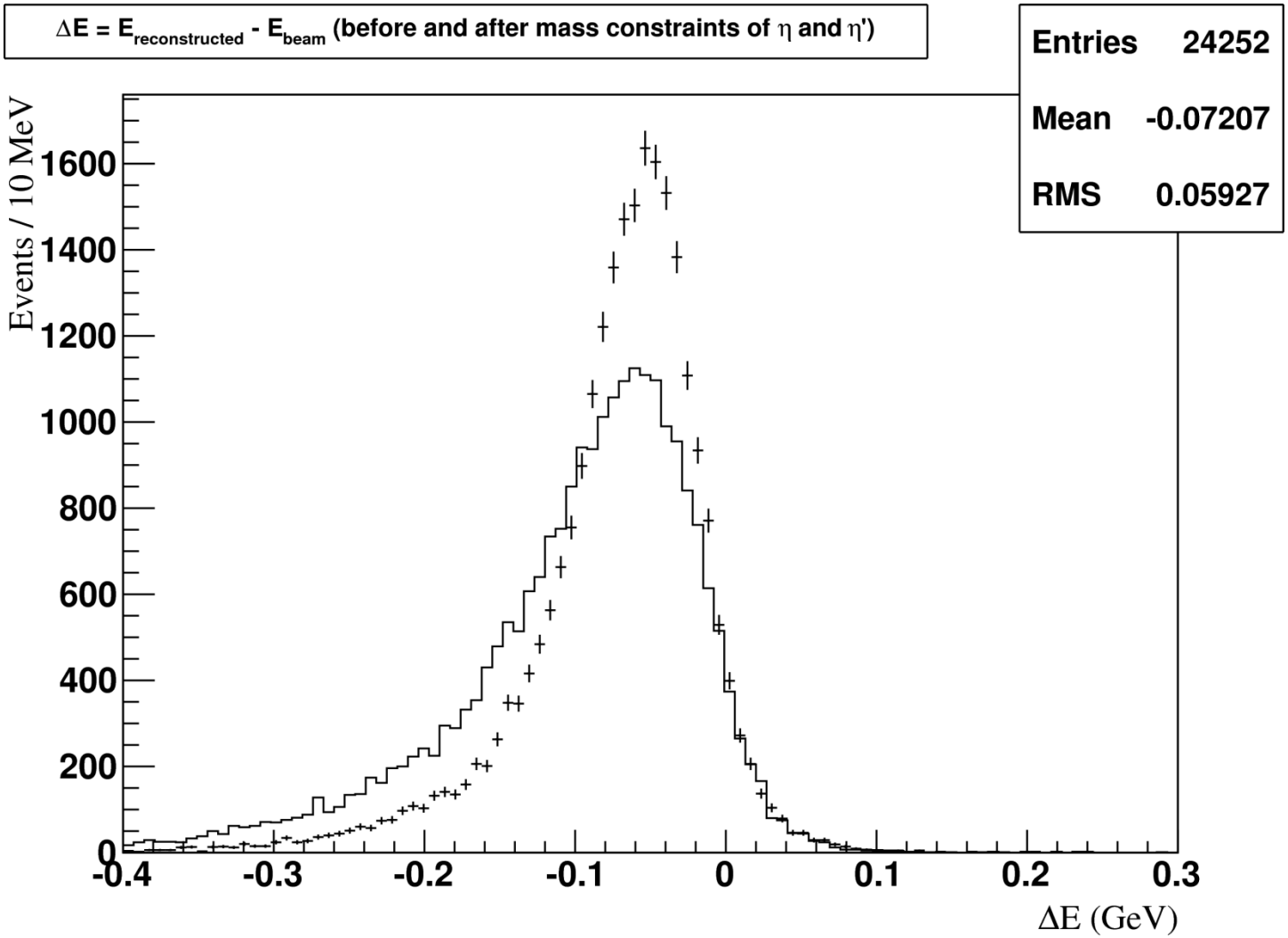
- Previously we chose our best candidate based on the best $M_{bc}$
- Now we choose the best candidate using (smallest) $\chi^2$ from fitting $\eta$ and $\eta'$
- We compared the results obtained using the two methods using fully MC-tagged candidates in signal MC before background suppression:

| | $M_{bc}$ | $\chi^2$ |
|---|---|---|
| % of MC Tagged candidates that are the best candidate | 94.9% | 93.6% |
| % of best candidates that are MC Tagged | 81.7% | 80.6% |

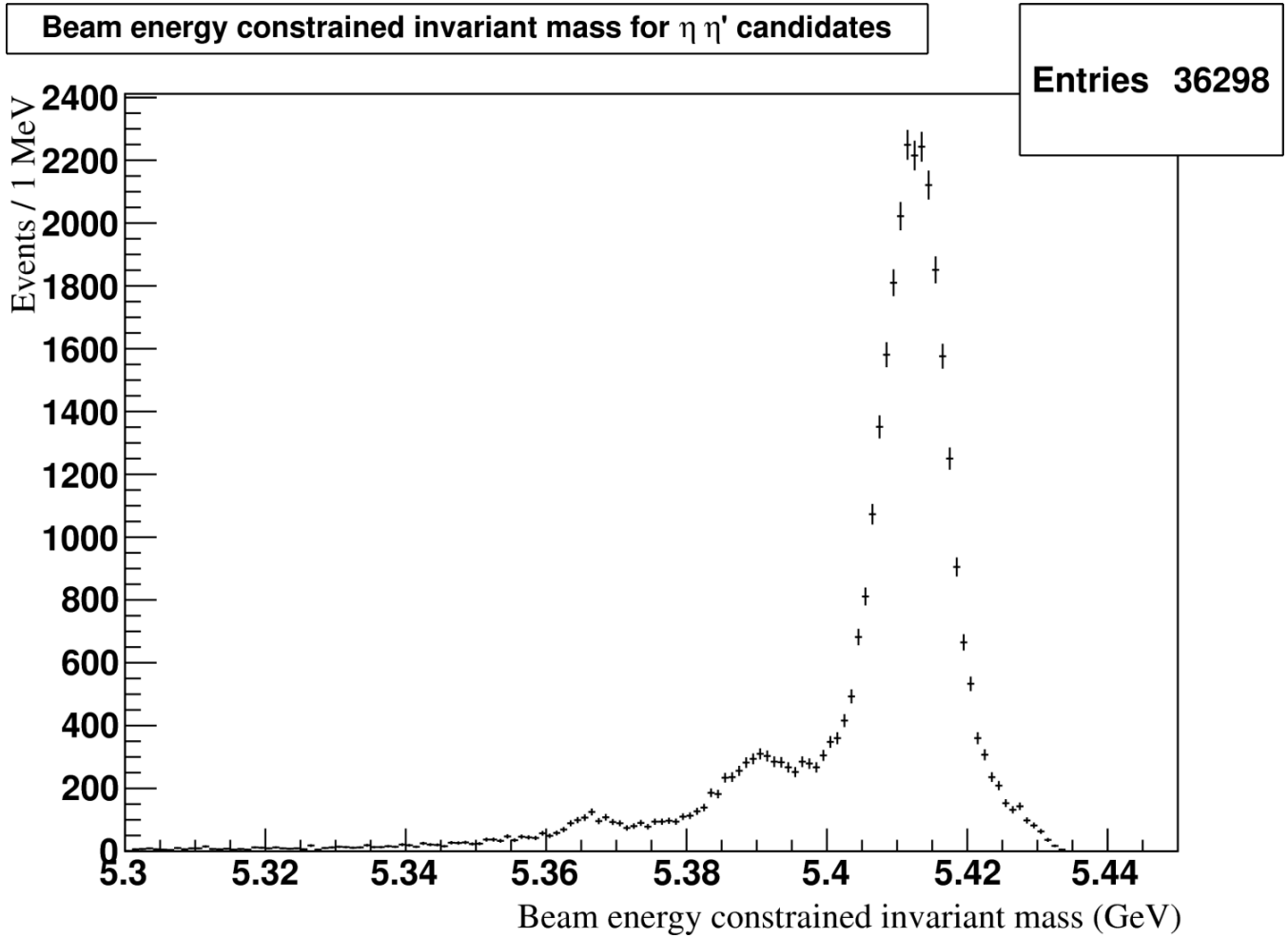
- Sorting by $M_{bc}$ gives slightly better efficiency of choosing fully MC-tagged candidates, however we choose to sort by $\chi^2$ to allow $M_{bc}$ to be used in our multidimensional likelihood fit
- Later we will compare the results obtained using the two methods to estimate systematics associated with the best candidate selection algorithm

# $\Delta E$ Distribution

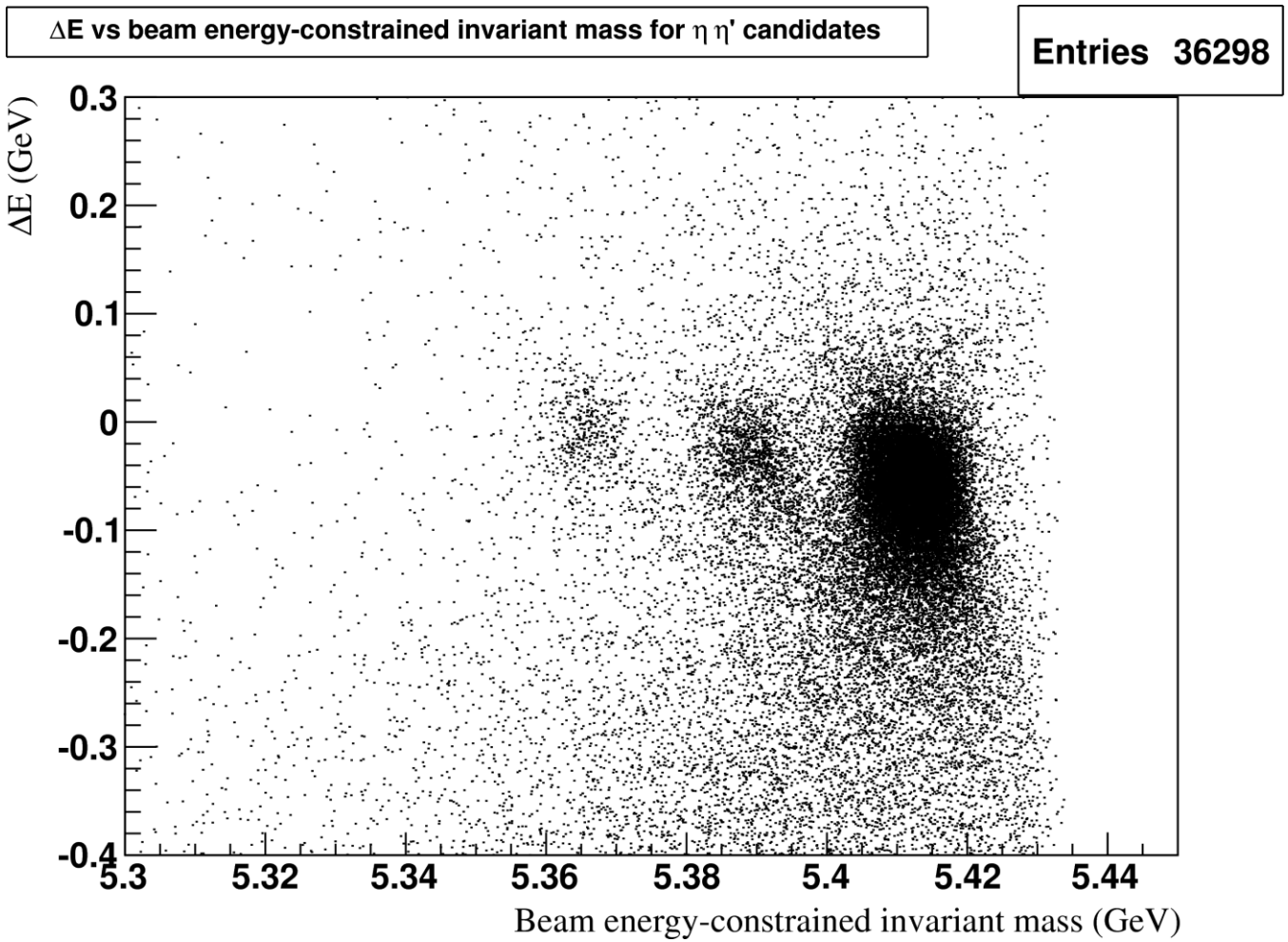
- Effect of mass constraints to $\eta$ and $\eta'$ on $\Delta E$ in MC



# $M_{bc}$ Distribution



# Distribution of $\Delta$E vs $M_{bc}$ for $\eta\eta'$ Candidates in signal MC



# Estimated $B_s \to \eta\eta'$ Signal Yield in BELLE data

- Using predicted branching fraction $(33.47 \pm 3.64) \times 10^{-6}$
- $(6.53 \pm 0.66) \times 10^6$ $B_s\overline{B_s}$ events
- x2 for possible signal $B_s$ on either side
- Expected number of signal decays with BELLE statistics: 437 events
- B( $\eta' \to \rho^0\gamma$) = .291
- B($\rho^0 \to \pi^+\pi^-$) = .989
- B($\eta \to \gamma\gamma$) = .3941
- Expected signal decays in this channel: 50
- $\varepsilon = .24$
- We expect to observe approximately 12 events in this channel

# MC Backgrounds

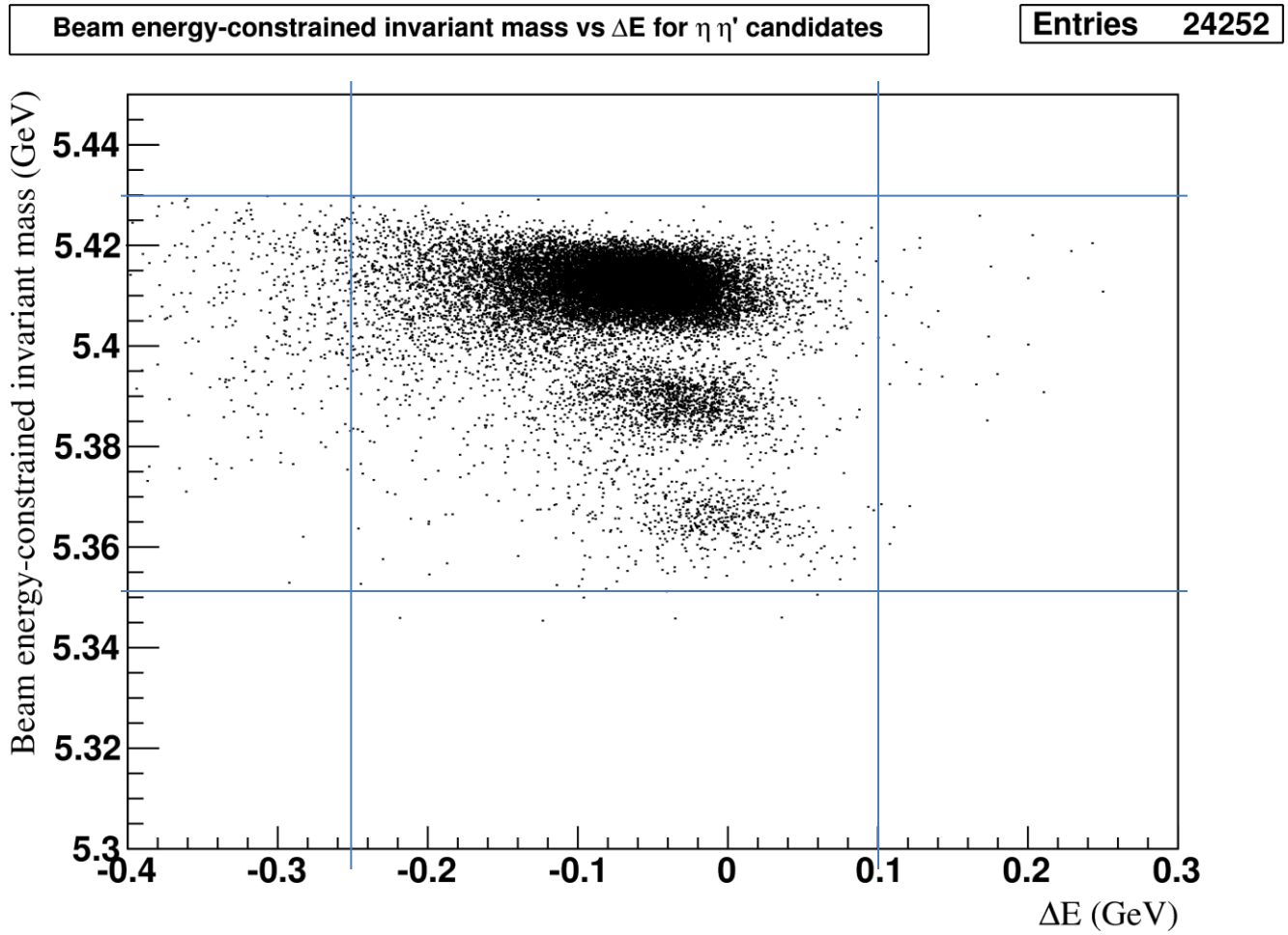
- We have 4 streams of generic $\Upsilon(5S)$ MC
  - S0+10, S30, S31, S32
  - S30 had several jobs crash, explaining a slightly smaller number of events
- Events passing through each stream:

| Stream # | Events |
|---|---|
| S0+10 | 290,367 |
| S30 | 279,540 |
| S31 | 287,728 |
| S32 | 287,605 |
| DATA | 335,693 |

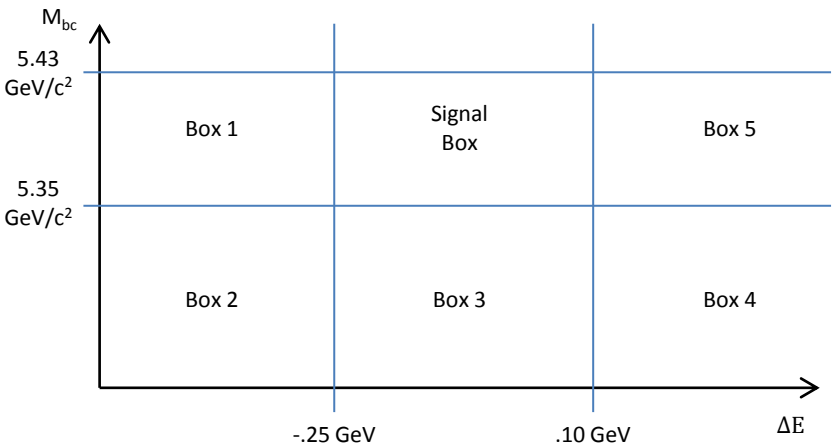
- Majority of background comes from continuum production (uds and charm)
- Data has significantly more events than MC

# Selection Criteria for Sidebands and Signal Region

| | |
|---|---|
| $B_s$ | $5.35\ GeV/c^2 \leq M_{bc} \leq 5.43\ GeV/c^2$<br>$-0.25 GeV \leq \Delta E \leq 0.1\ GeV$ |



# Comparison of Signal MC and Data Sidebands

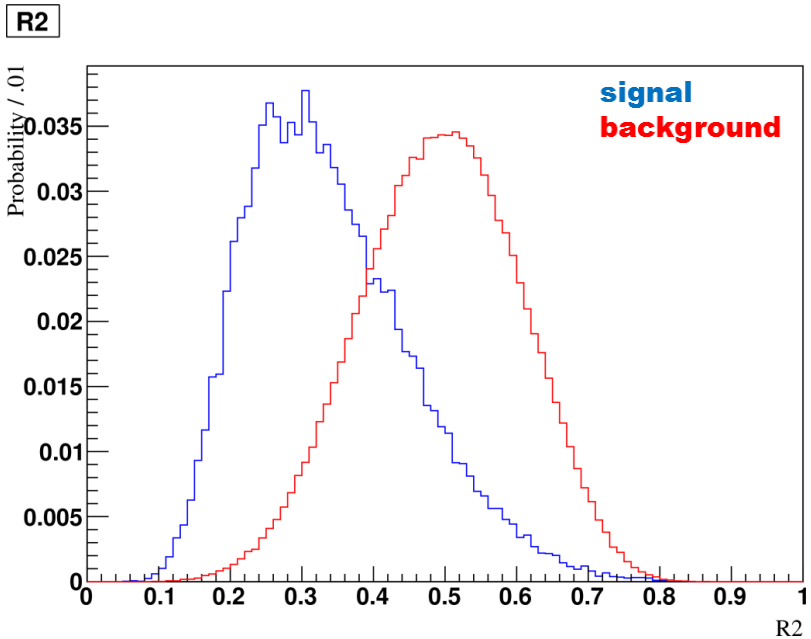


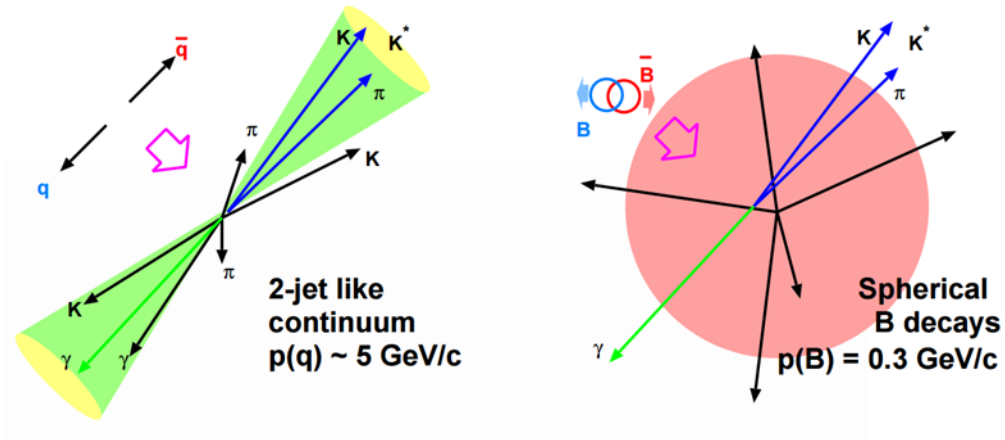

| | BOX 1 | | | SIGNAL BOX | | | BOX 5 |
|---|---|---|---|---|---|---|---|
| **S010** | 43,163 | | **S010** | 63,442 | | **S010** | 21,267 |
| **S30** | 41,323 | | **S30** | 61,104 | | **S30** | 20,709 |
| **S31** | 42,480 | | **S31** | 62,898 | | **S31** | 21,196 |
| **S32** | 42,439 | | **S32** | 63,213 | | **S32** | 20,966 |
| | | | | | | | |
| **AVERAGE** | 42,694 | | **AVERAGE** | 63,184 | | **AVERAGE** | 21,143 |
| **DATA** | 45,654 | | **DATA** | 69,637 | | **DATA** | 24,651 |
| | | | | | | | |
| | **BOX 2** | | | **BOX 3** | | | **BOX 4** |
| **S010** | 54,673 | | **S010** | 79,505 | | **S010** | 26,182 |
| **S30** | 52,619 | | **S30** | 76,395 | | **S30** | 25,271 |
| **S31** | 54,416 | | **S31** | 78,210 | | **S31** | 26,395 |
| **S32** | 53,884 | | **S32** | 79,019 | | **S32** | 25,930 |
| | | | | | | | |
| **AVERAGE** | 54,324 | | **AVERAGE** | 78,911 | | **AVERAGE** | 26,169 |
| **DATA** | 63,633 | | **DATA** | 97,625 | | **DATA** | 34,493 |

- Every box has more events in data than background MC
- uds has the largest background contribution in all boxes
- bsbs and nonbsbs contributions are very small in all boxes

# Background Suppression

- Continuum background events from light quark pair production carry large initial momenta resulting in a "jet-like" event shape
- $B_s^{(*)}\overline{B_s^{(*)}}$ events are distributed isotropically
- Jet-like $q\bar{q}$ events can be separated from spherical $B_s^{(*)}\overline{B_s^{(*)}}$ events using event shape variables
- Example: R2



# Background Suppression Strategy

- Choose best candidate based on least overall $\chi^2$ of kinematic fits for $\eta$ and $\eta'$

- Use MC-tagged signal MC candidates and the best candidates from background MC to optimize Likelihood Ratio (LR) and then NeuroBayes output (at a later time we will try to use signal sideband in data for this optimization)

- Use likelihood ratio, $|\cos\theta_{thrust}|$, $|\cos\theta_{bt}|$, and R2 as inputs to NeuroBayes

- Cut on NeuroBayes ($NN$) output ($NN \geq NN_{cut}$ = 0.6), prepare NB' and include it in our multidimensional fit

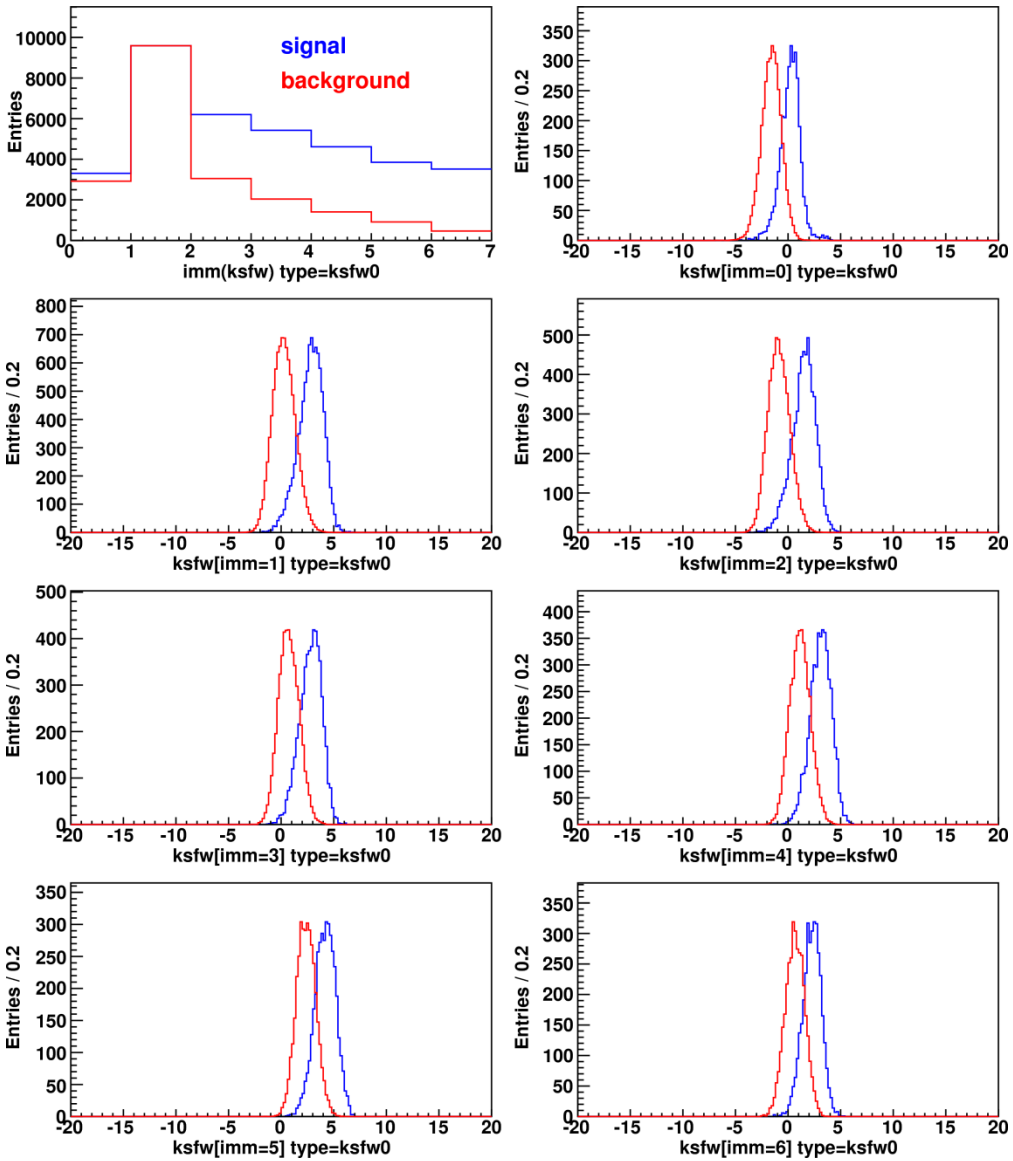
$$NB' = \log\frac{NN - NN_{cut}}{NN_{max} - NN}$$

# Output of RooKSFW

- Prepared 18 KSFW moments
- Used Nakao-san's rooksfw program to optimize coefficients of Fisher Discriminant
- Prepared likelihood ratio to use as input for NeuroBayes:

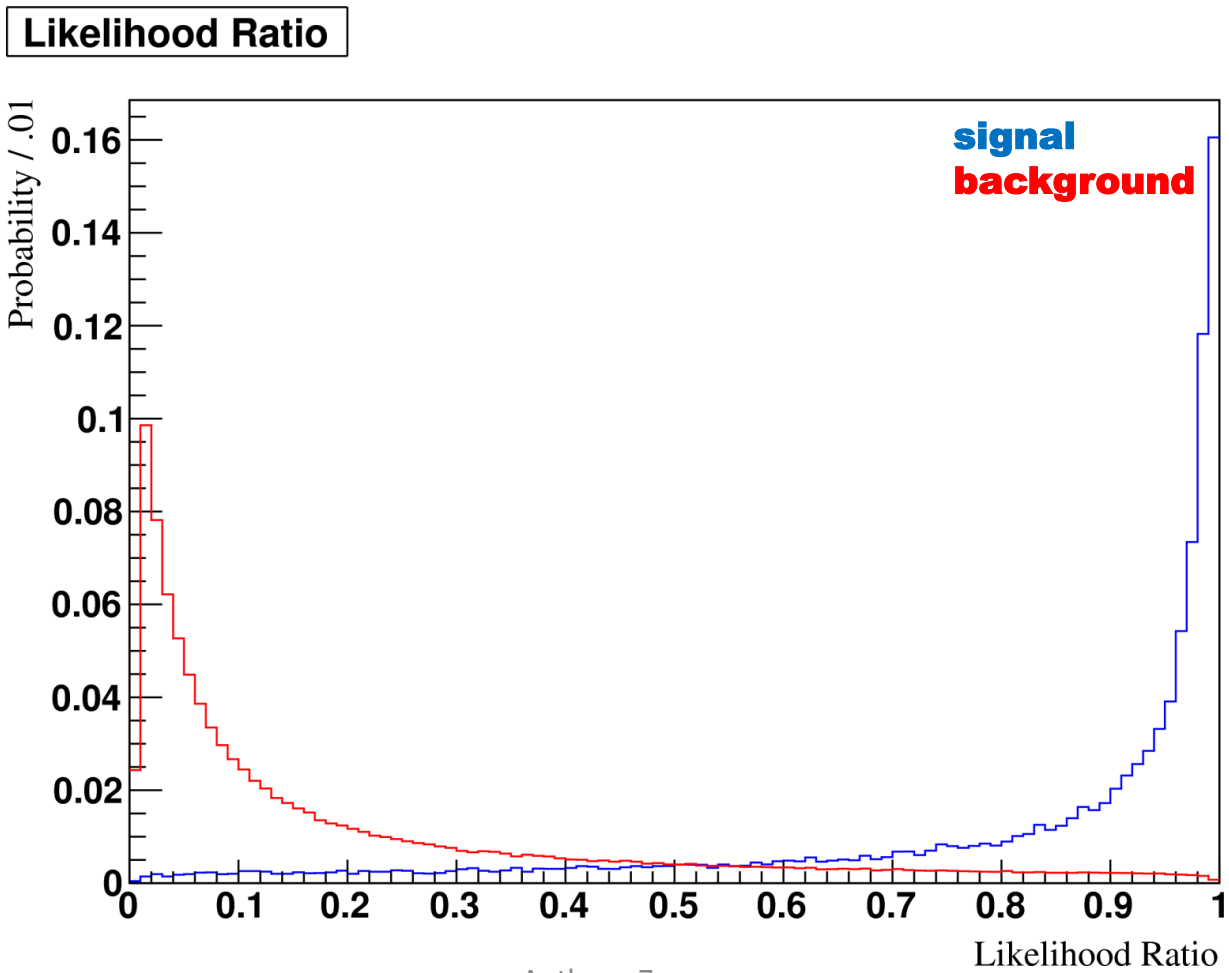
$$LR = P_s/(P_s + P_b)$$

- Likelihood ratio peaks at 1 for signal and 0 for background (shown on the next slide)
- Note that we can feed all KSFW moments "as is" to NeuroBayes, however, we chose to separate these two steps, so we can estimate how much we gain by using NeuroBayes (as compared to LR alone)
- For final analysis we plan to compare the results obtained using NeuroBayes with all KSFW moments, NeuroBayes with LR and LR alone. The comparison could then be used to estimate systematic uncertainty associated with background suppression



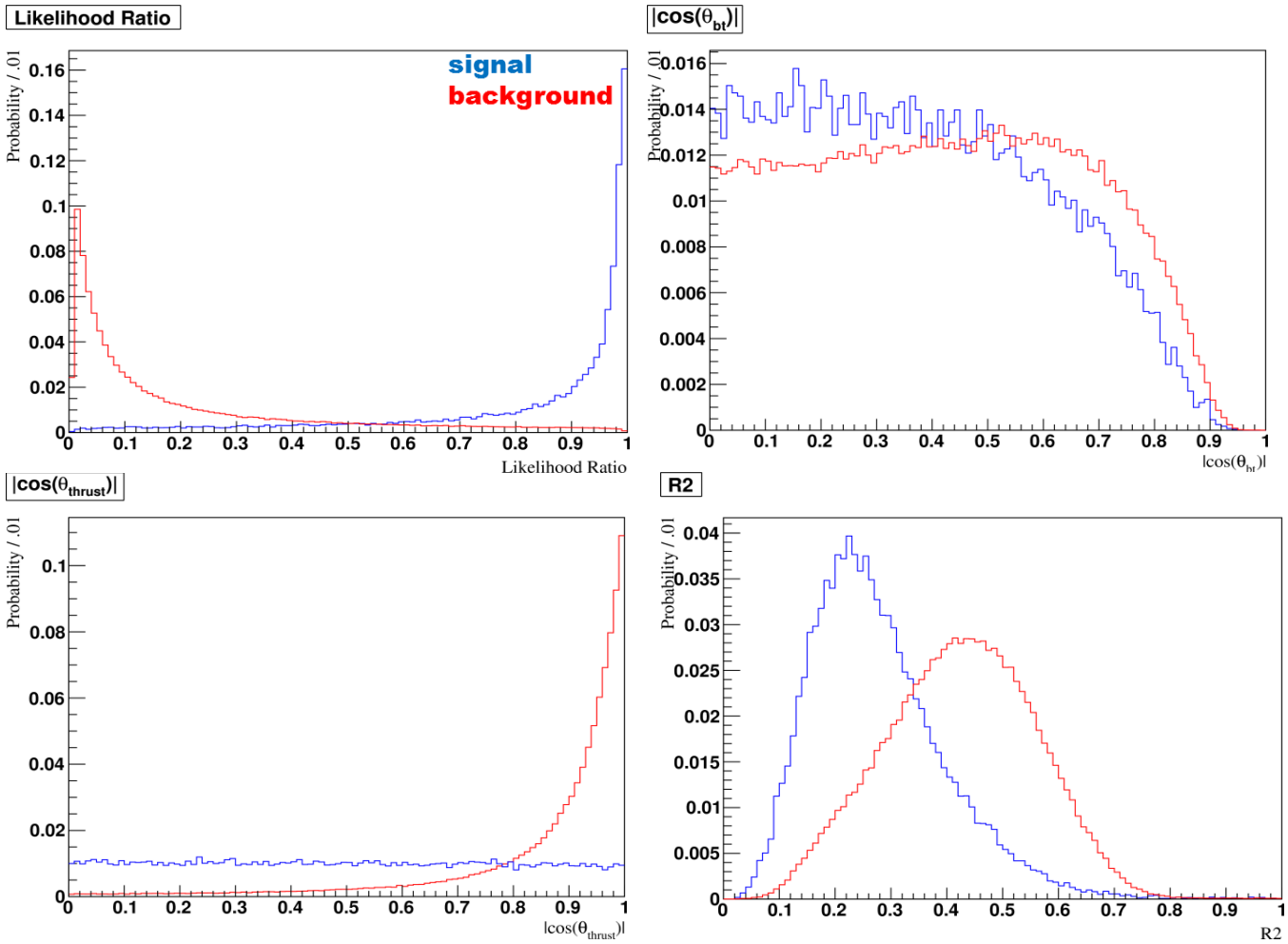
# Likelihood Ratio

- We use all MC tagged signal candidates (~24k) and the best entry of all MC background events (~230k) from one stream to optimize our Likelihood Ratio (in figure below the distributions are scaled to the same area)



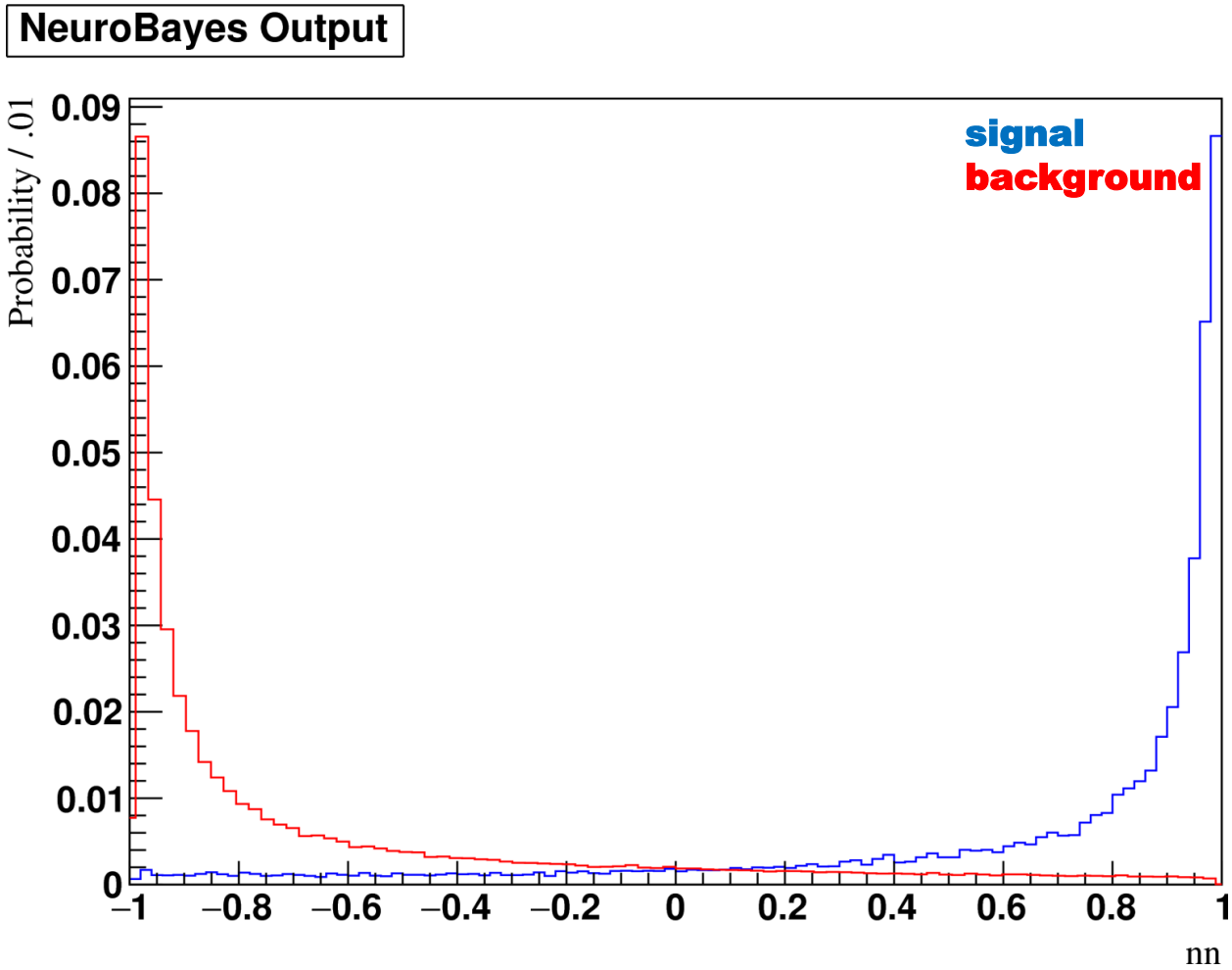
# NeuroBayes Input Variables

- We use likelihood ratio, $|\cos\theta_{bt}|$, $|\cos\theta_{thrust}|$, and R2 as inputs to NeuroBayes to suppress continuum background. Similar numbers of signal MC and generic background MC events from one stream of generic MC are used for NeuroBayes training



# NeuroBayes Output

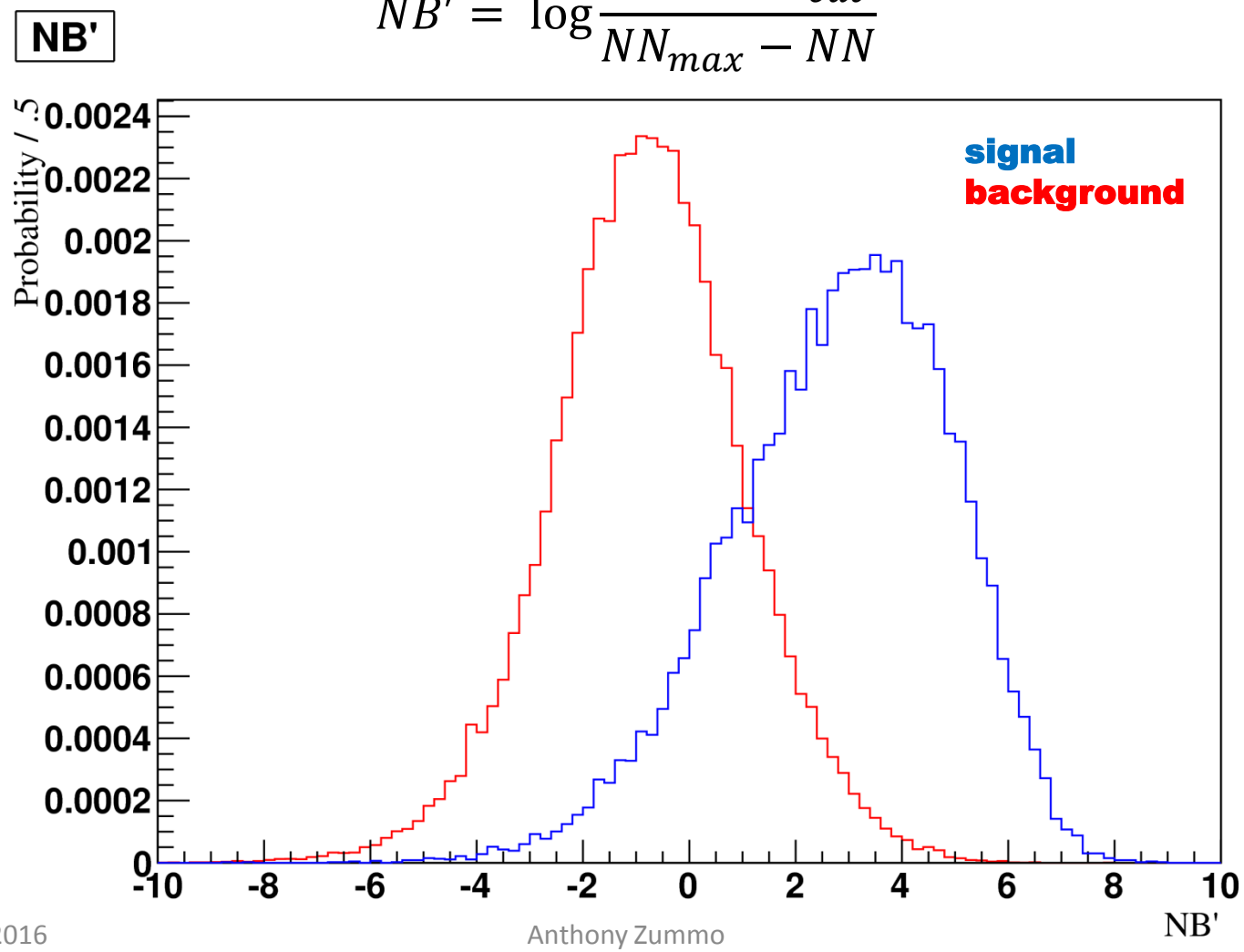
- We use all MC tagged signal candidates (24k) and 1/10 of the best entry of all background events (23k) to optimize our NeuroBayes output



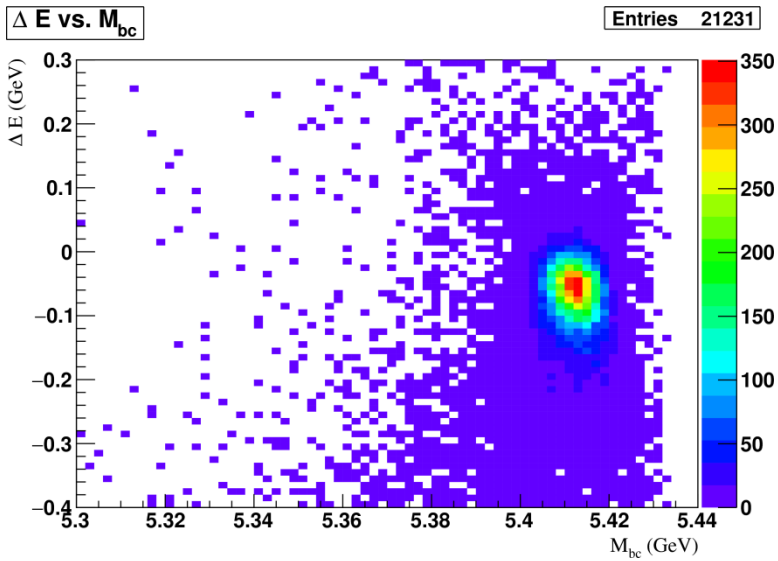
# NB’

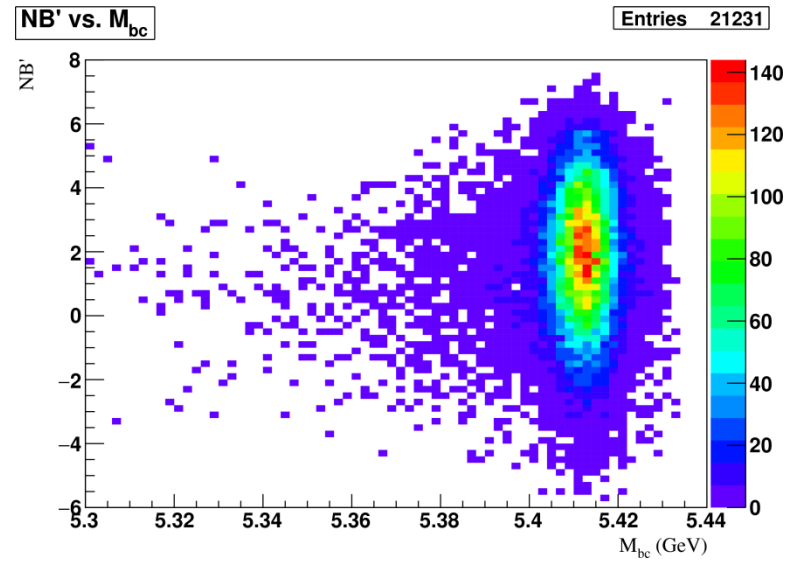
- NB’ distribution shows separation between signal and background after we apply selection criterion on NeuroBayes output variable: $NN \geq NN_{cut}$= 0.6

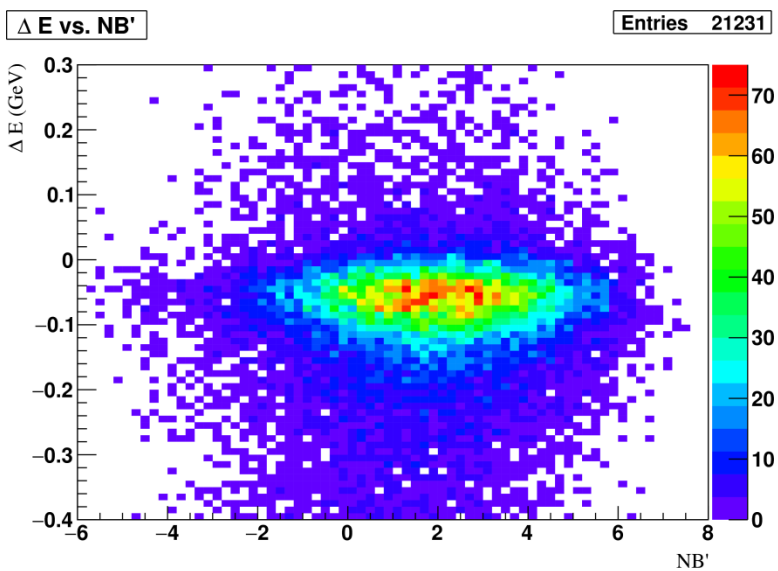
$$NB' = \log \frac{NN - NN_{cut}}{NN_{max} - NN}$$



# Correlations Between $M_{bc}$, $\Delta$E, and $NB'$ in Signal MC?



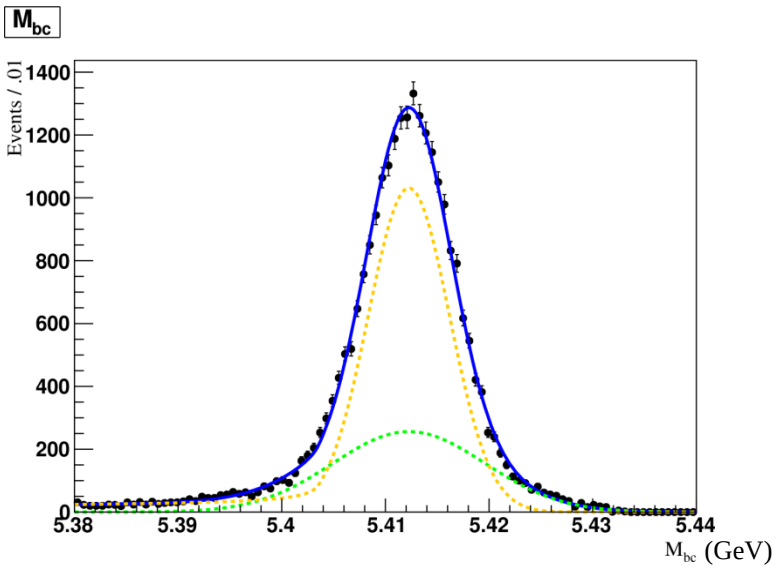

Note: These plots (and fits on the next slides) use only $B_s^* \overline{B_s^*}$ events

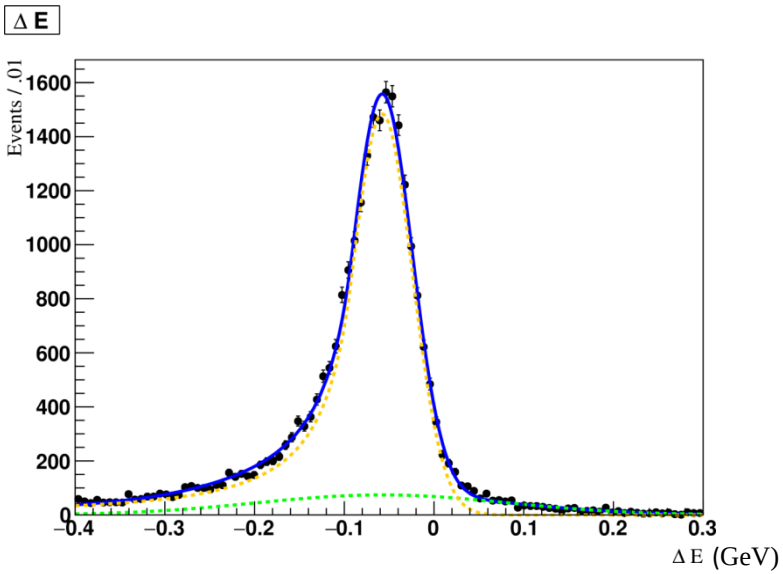
# Modeling Signal MC

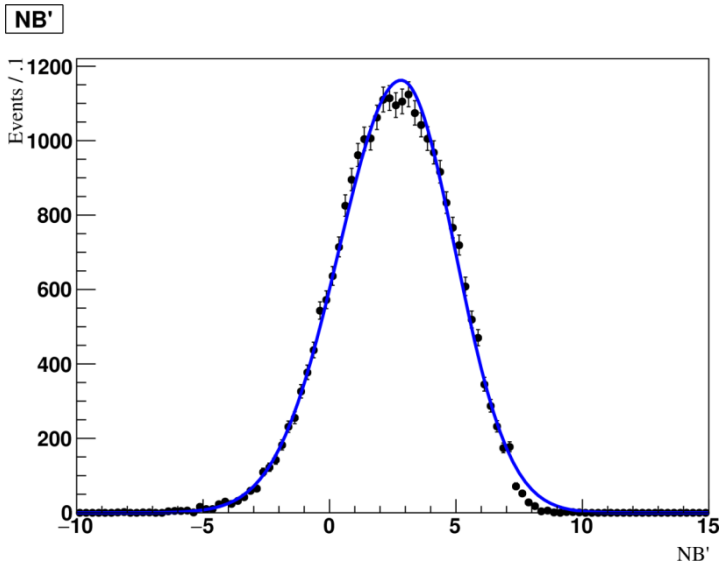
- Tentatively, our model for signal is

| Variable | Function |
|---|---|
| $M_{bc}$ | Crystal Ball<br>Gaussian |
| $\Delta E$ | Crystal Ball<br>Gaussian |
| $NB'$ | Bifurcated Gaussian |

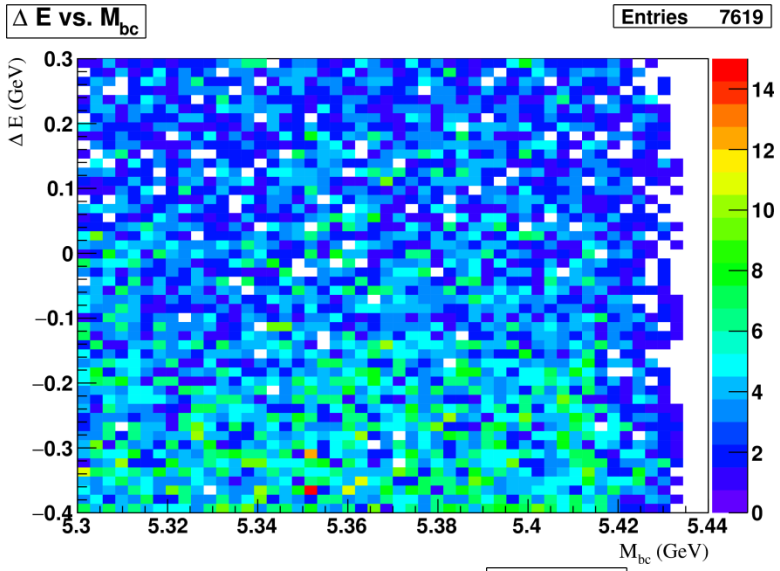
# Modeling Signal MC

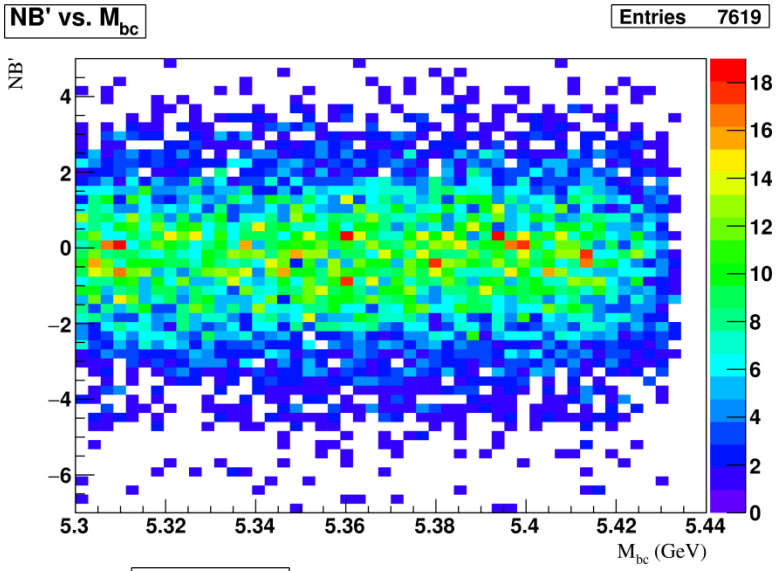


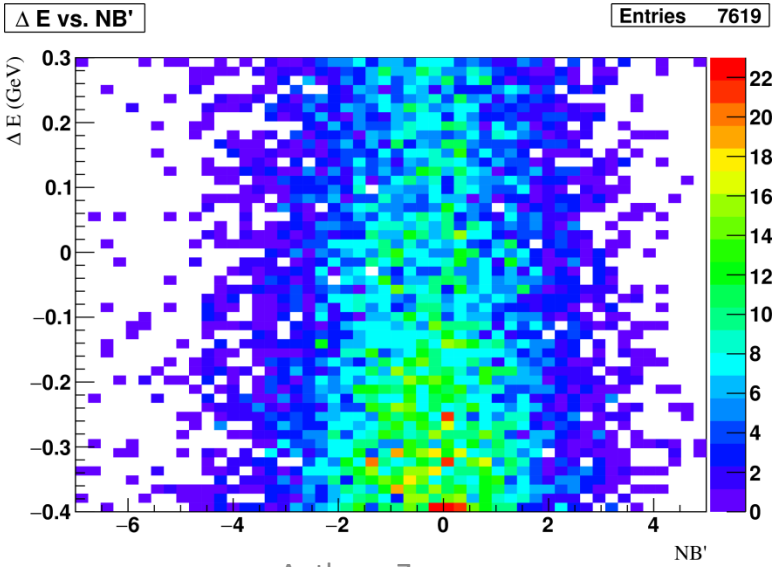



# Correlations Between $M_{bc}$, $\Delta$E, and $NB'$ in Background MC?



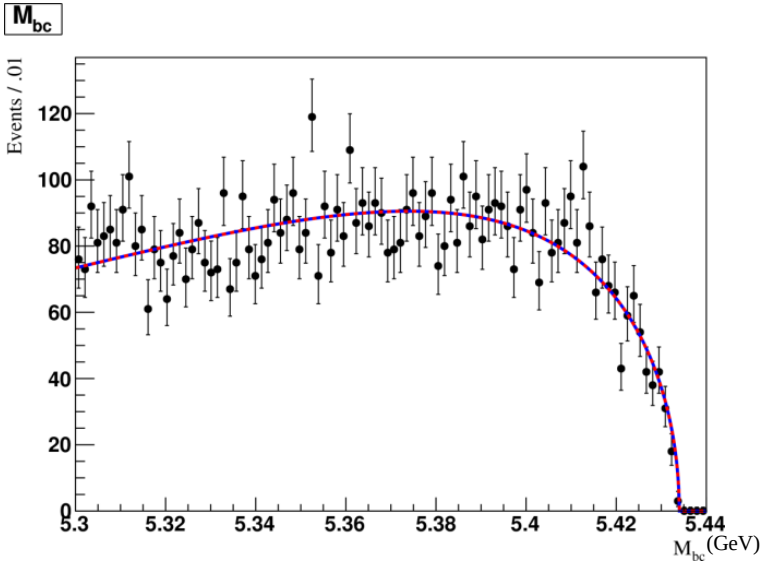



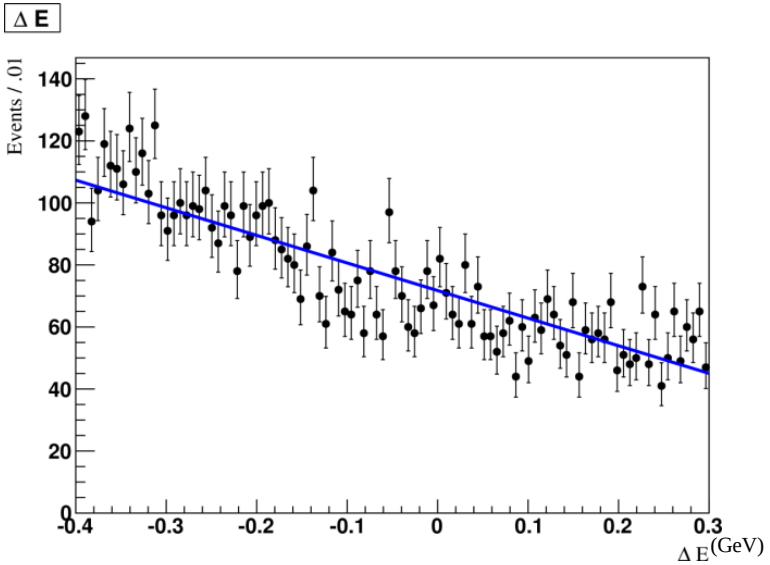
# Modeling Background MC

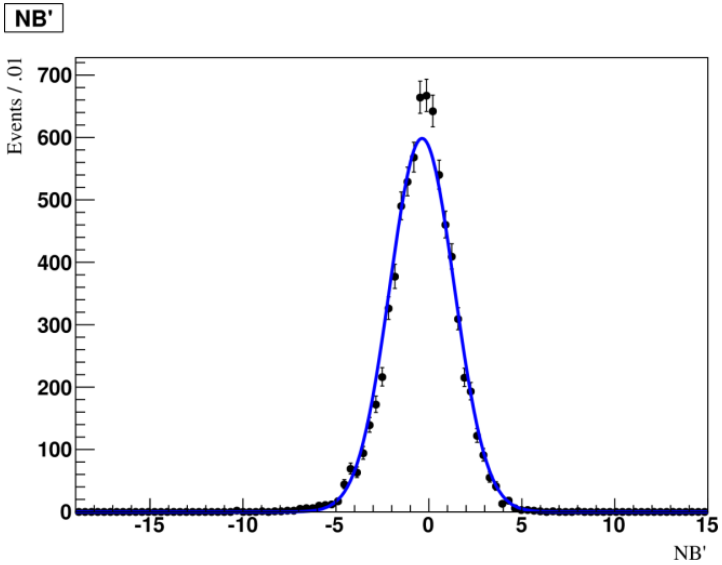
- Tentatively, our model for background is

| Variable | Function |
|---|---|
| $M_{bc}$ | ARGUS |
| $\Delta E$ | Chebychev Polynomial |
| $NB'$ | Gaussian |

# Modeling Background MC



# Summary of Progress

- Focusing on $\eta' \to \rho^0\gamma, \eta \to \gamma\gamma$
- Used mass fit to $\eta$ and $\eta'$ improving $\Delta E$ resolution
- Estimated signal yield of approximately 14 events
- Ran on 4 streams of generic $\Upsilon(5S)$ MC background
- Compared background streams with sidebands in real data
- Prepared 18 KSFW moments which were used to calculate Likelihood Ratio
- Input LR, $|\cos\theta_{bt}|$, $|\cos\theta_{thrust}|$, and R2 to NeuroBayes for background suppression
- Prepared NB' variable to be used in fit
- Began modeling signal and background distributions

# Outlook

- Continue modeling of signal and background pdf's
- Toy MC studies and linearity/ensemble tests

Thank You!