Search for $B_s \rightarrow K_s \pi^0$ in Belle Data

Tao Luo
Vladimir Savinov
University of Pittsburgh
2nd Annual Belle II Summer School, Aug. 19 2016

Motivation

- Has not been observed yet, not in PDG
- Belle is more sensitive to this decay than LHCb
- Rare decays are sensitive to BSM physics
- Test the theoretical prediction:
  - Br ($B_s \rightarrow \overline{K}^0 \pi^0$) = $(2.25 \pm 0.33) \times 10^{-6}$
    Hai-Yang Cheng, Cheng-Wei Chiang, and An-Li Kuo; PRD **91**, 014011 (2015) (Updating B→PP,VP decays in the framework of flavor symmetry)

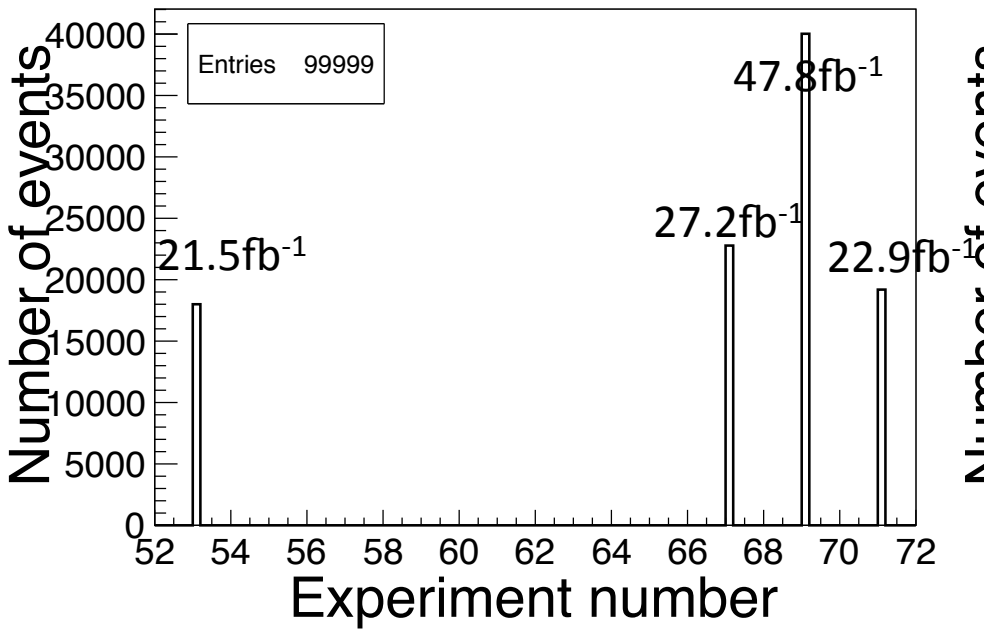
Signal Monte Carlo generation

- Generated (using mcproduzh scripts) 100,000 → events in proportion to cross sections for $B_s^* \bar{B}_s^*$, $B_s^* \bar{B}_s$, $B_s \bar{B}_s$ production at $\Upsilon(5S)$ and distributed over BELLE data samples (53, 67, 69, 71) according to data statistics
- Signal decay: $\pi^0 \rightarrow \gamma\gamma$, K_s is decayed by Geant ($Br(K_s \rightarrow \pi^+\pi^-)$ = 68.61% in Geant)

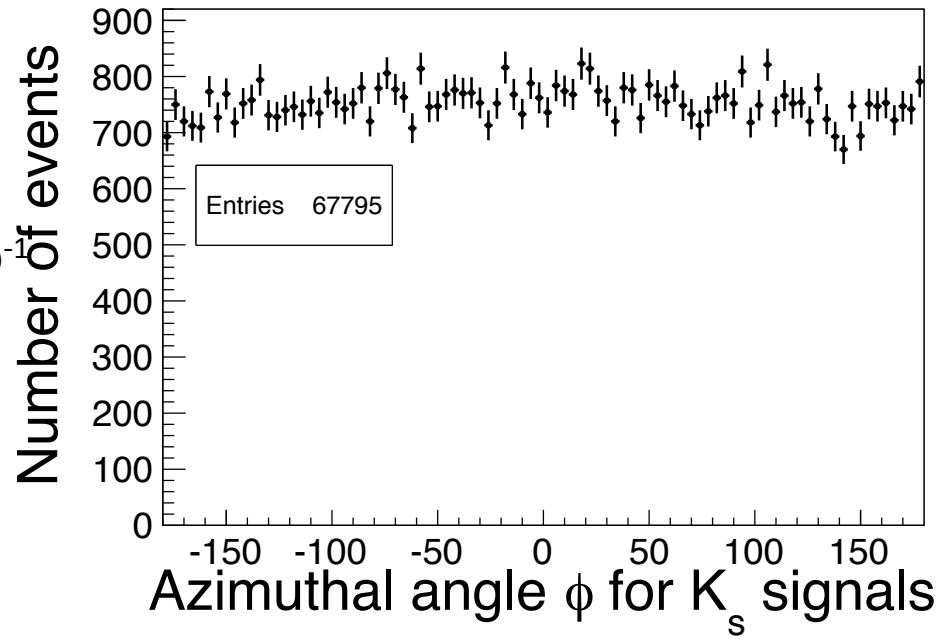
```
Decay My-B_s0
1.000          K_S0          My-pi0          PHSP;
Enddecay

Decay My-anti-B_s0
1.000          K_S0          My-pi0          PHSP;
Enddecay

Decay My-pi0
1.000          gamma          gamma          PHSP;
Enddecay          End
```

MC truth

Experiment Number



Azimuthal angle phi distribution for signal Ks's (MC truth)



$Br(K_s \rightarrow \pi^+\pi^-)$ = 67.95% for this MC sample

Selection criteria for signal event candidates

- K_s :
  - nisKsFinder (standard selection, BN#1253)
- π^0:
  - They are from standard Mdst_pi0 bank
  - Mass constrained fit
  - Additional criteria in next page
- B_s:
  - $M_{bc} \geq 5.2$ GeV / c^2
  - -0.4 GeV $\leq \Delta E \leq 0.3$GeV

Selection of daughter photons for signal π^0 candidates and selection of signal π^0 candidates

- Selection criterion applied to photon energy:
  - Detected in the barrel calorimeter: $E_\gamma > 50$MeV
  - Detected in the endcap calorimeters: $E_\gamma > 100$MeV
- Selection criteria applied on signal π^0 candidates
  - Momentum of π^0 : $P(\pi^0) > 100$ MeV
  - χ^2 of π^0 mass constraint fit : $\chi^2(\pi^0) < 50$
  - π^0 invariant mass before mass constraint fit: $M(\pi^0) > 120$ MeV and $M(\pi^0) < 150$ MeV

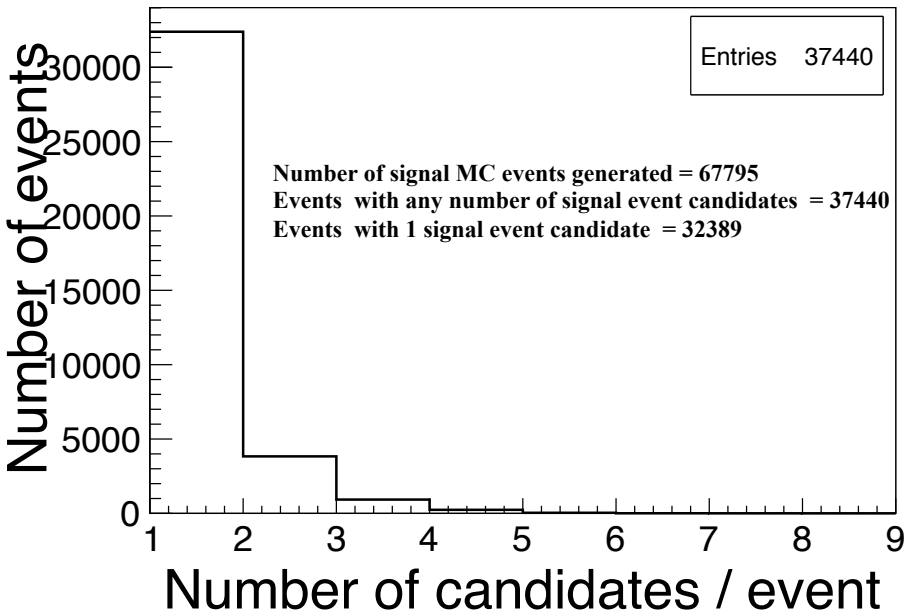
Efficiency and number of candidates per event in signal MC

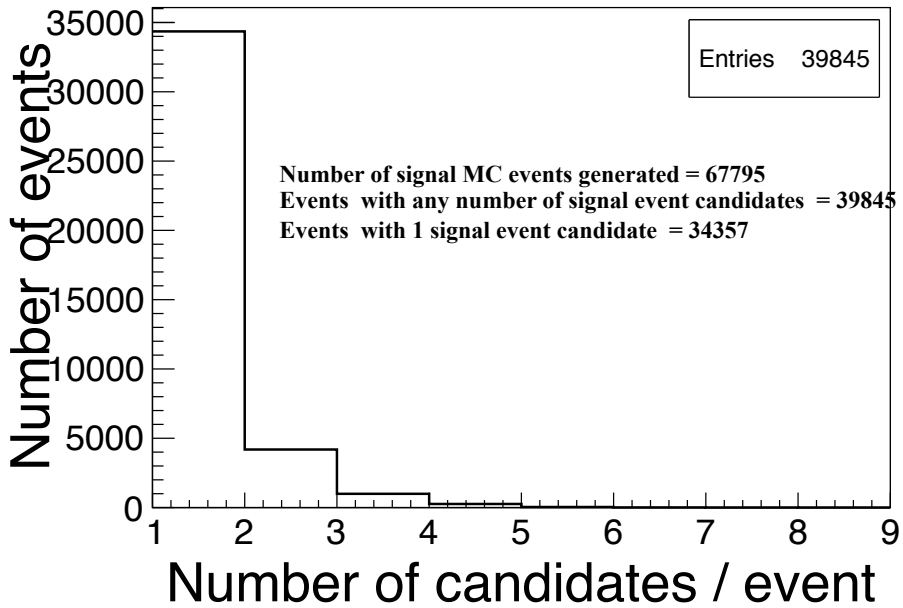
GoodKs selection

- 86.5% of signal MC events have only one signal candidate
- ε = 55.2% (37,440/67,795) for signal MC events which have at least one signal candidate

nisKsFinder selection

- 86.2% of signal MC events have only one signal candidate
- ε = 58.8% (39,845/67,795) for signal MC events which have at least one signal candidate





The efficiency improves by an number of ~6.5% (relative, i.e. (58.8-55.2)/55.2)

Best Candidate Selection based on χ^2 of mass constraints

- $\chi^2_{fit} = \chi^2_{\pi 0 mass} + \chi^2_{Ks\ vertex}$: best B_s candidate has the smallest χ^2_{fit}

	χ^2_{fit}
% of MC Tagged candidates that are the best candidate	96.9%
% of best candidates that are MC Tagged	93.1%

Skimming of $\Upsilon(5S)$ data and generic MC

- Same selection criteria are used to do the skimming
- We have produced skims of six streams of generic $\Upsilon(5S)$ MC, and four $\Upsilon(5S)$ Belle data samples

MC backgrounds

- We used 6 streams of generic $\Upsilon(5S)$ MC
  - S0-5
- Number of events satisfying skimming selection criteria:

Stream Number	Events
S0	119,912
S1	119,812
S2	119,973
S3	120,056
S4	120,344
S5	120,438
Data	155,300

Background predictions according to generic MC (before background suppression)

Stream #	uds	charm	qq total	bsbs	Non-bsbs	bb total	Total
S0	79,394	40,517	119,911	2	66	68	119,979
S1	79,247	40,562	119,809	8	102	110	119,919
S2	79,215	40,758	119,973	3	74	77	120,050
S3	79,292	40,760	120,052	8	68	76	120,128
S4	79,760	40,577	120,337	4	78	82	120,419
S5	79,849	40,587	120,436	4	70	74	120,510
Data							155,300

Almost all background comes from continuum production of light quarks (uds and charm)

Toolkit for **M**ult**iv**ariate **A**nalysis (TMVA)

- TMVA is used to perform the multivariate classification
  - Cuts: Rectangular Cut Optimization
  - Likelihood: Projective Likelihood (PDE Approach)
  - LD: Linear Discriminant identical to Fisher
  - Fisher: Fisher discrimination
  - MLP: Artificial Neural Networks (Non-Linear Discriminant Analysis)
  - MLPBFGS: ANN with optional training method
  - MLPBNN: ANN with BFGS training method and Bayesian regulator
  - CFMlpANN: Depreciated ANN from ALEPH
  - BDT: uses Adaptive Boost
  - BDTG: uses Gradient Boost

Input variables for TMVA

- vr2: the ratio of the second and zeroth Fox-Wolfram moments (https://kds.kek.jp/indico/event/3808/contribution/25/material/slides/0.pdf)
- vthrust (i.e. $|\cos\theta_{th}|$) : θ_{th} is the angle between the thrust axis of the Bs (candidate) and the thrust axis of the rest of reconstructed particles (charged tracks and photons of energies above 50 MeV)
- cosbt (i.e. $|\cos\theta_{bt}|$): θ_{bt} is the angle between the thrust axis of the Bs (candidate) and the beam axis

- 16 Kakuno-Super-Fox-Wolfram (KSFW) moments

See Nakao's talk on 2009 Belle Analysis school (https://kds.kek.jp/indico/event/3808/contribution/25/material/slides/0.pdf)

19 variables are used for TMVA: vr2, vthrust, cosbt and 16 KSFW moments

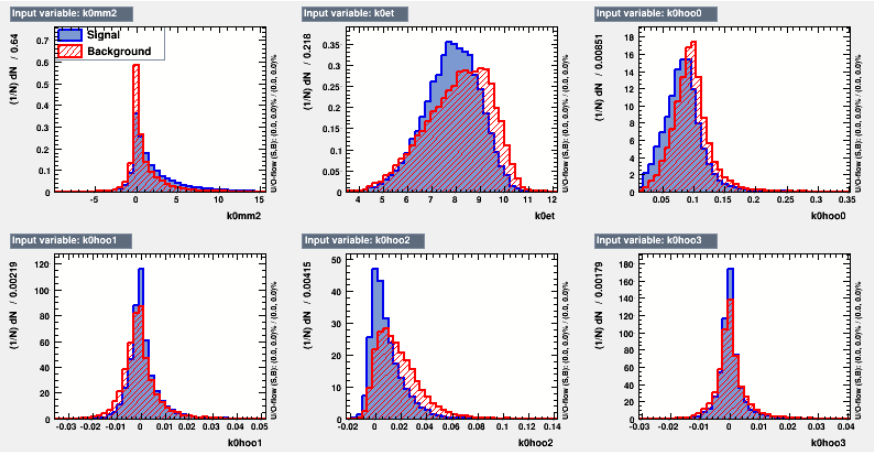
Training and validation MC samples

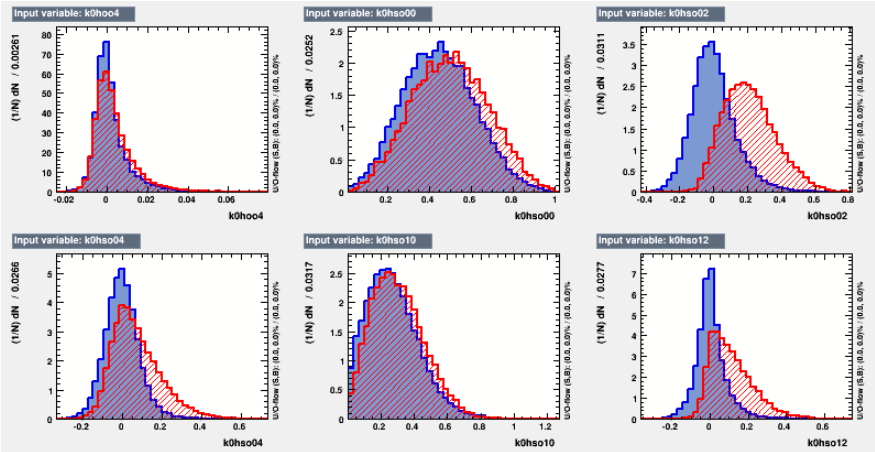
- Signal: only best MC-tagged candidates are used

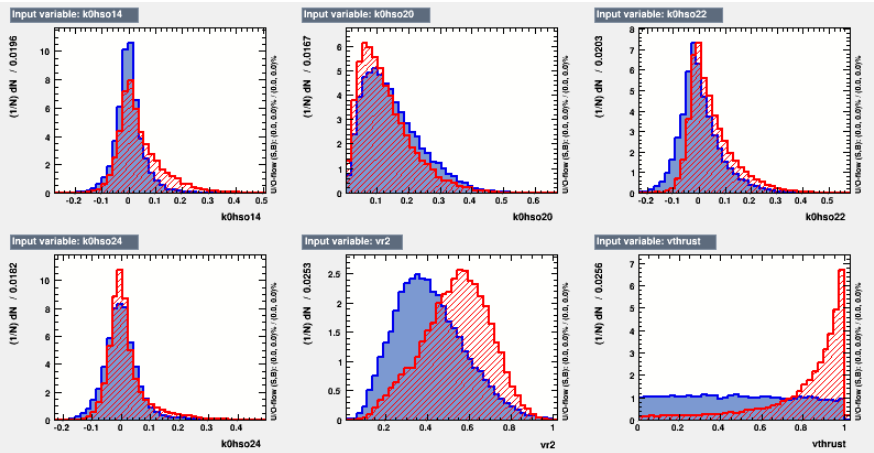
- Background: part of one stream of generic MC
  - Only best candidates are used

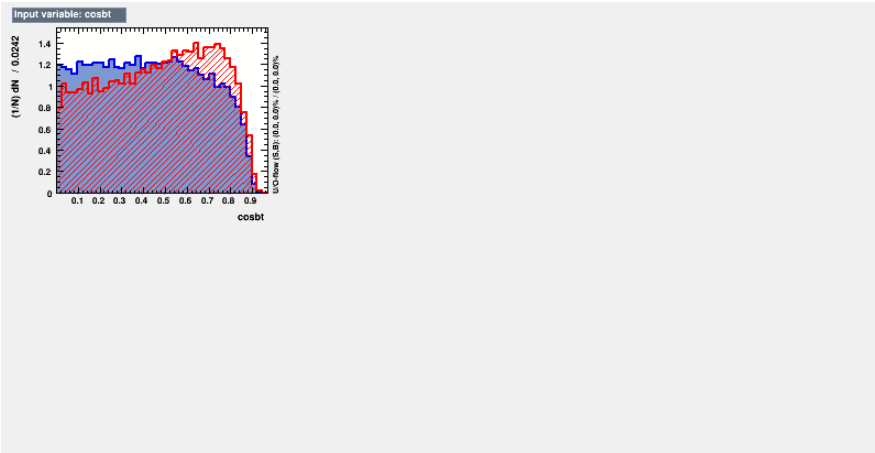
To train TMVA discriminators we use similar numbers of generic background and signal MC events

Input variables used in TMVA



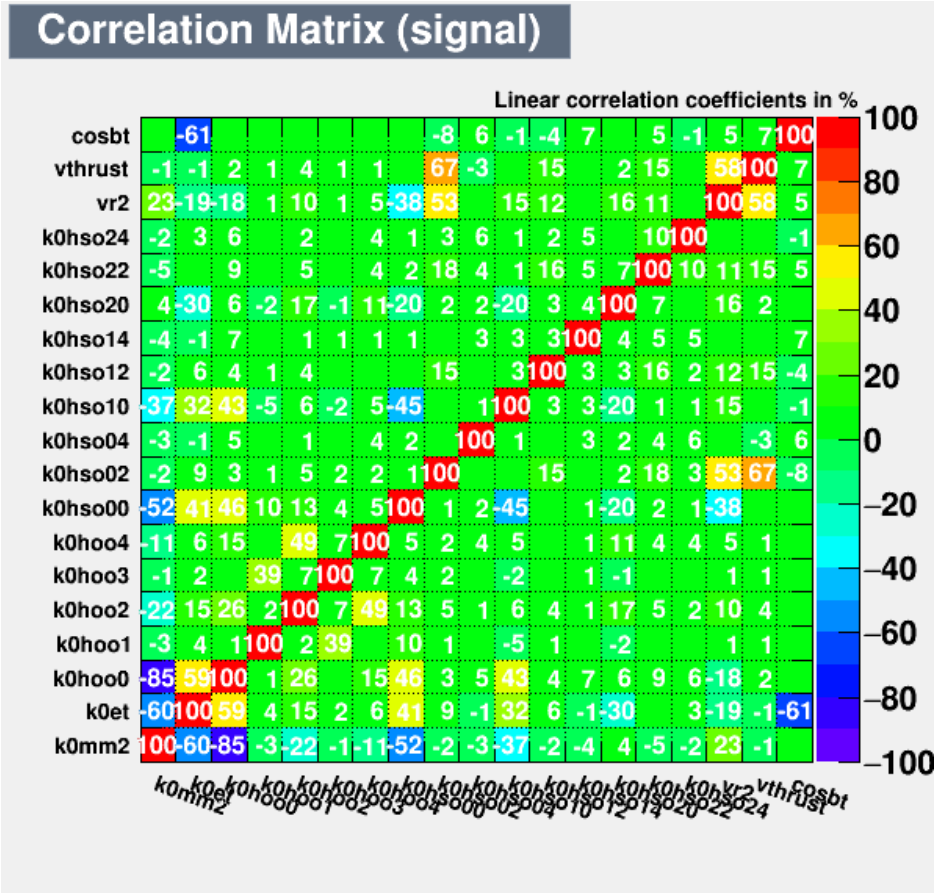




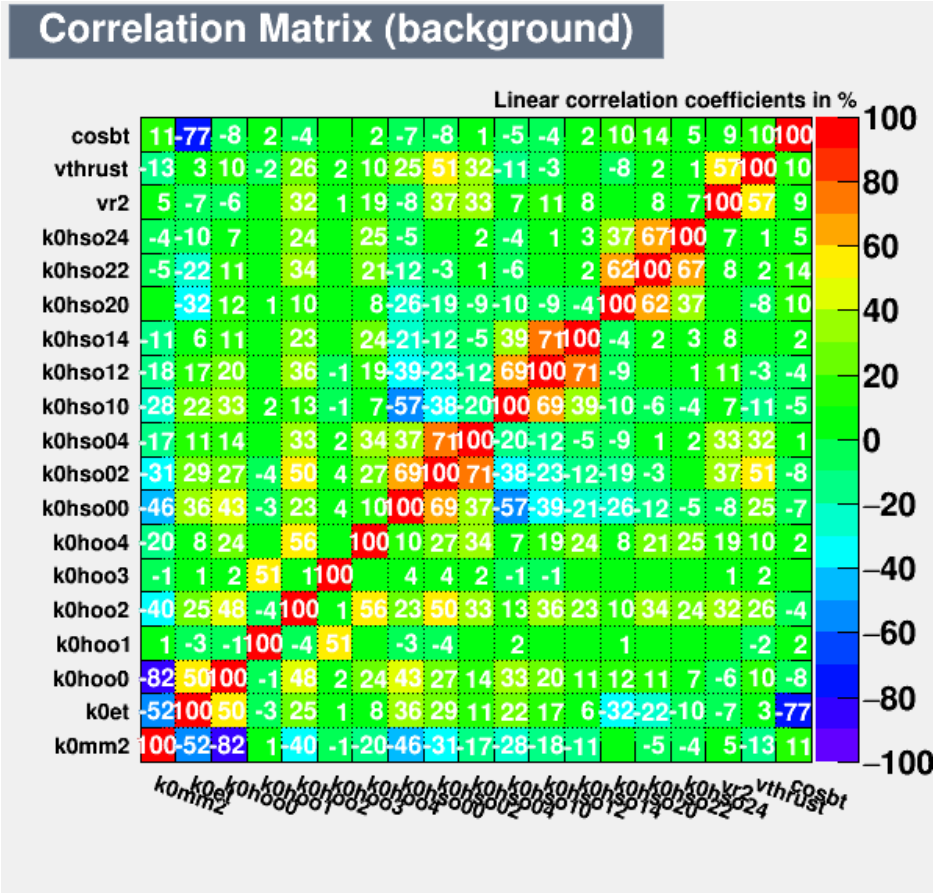


Correlation matrix for all input variables

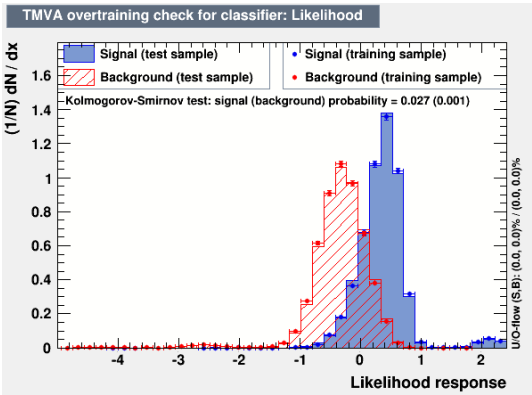
Signal

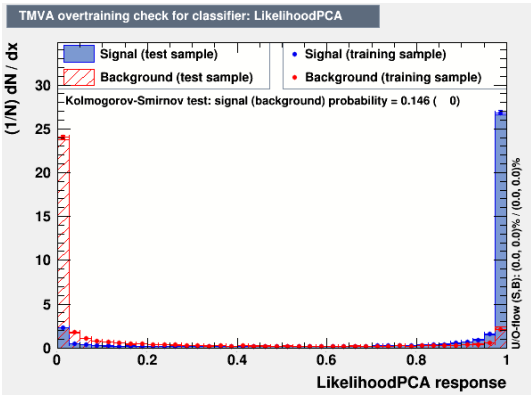


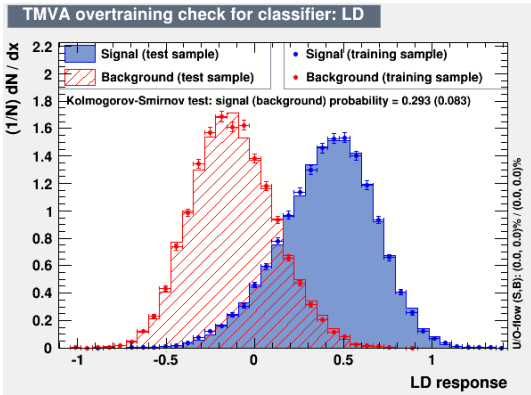
Background

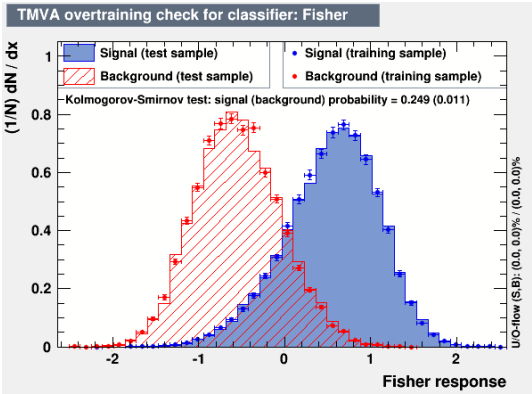


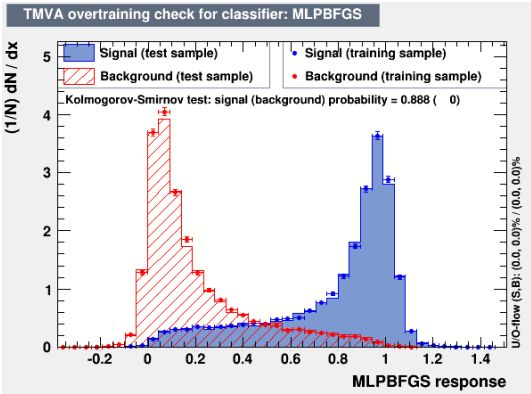
TMVA output for signal and background

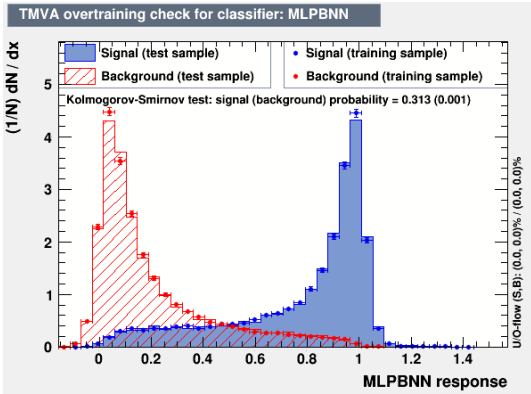


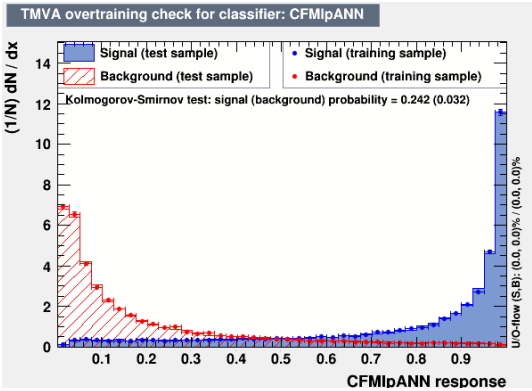


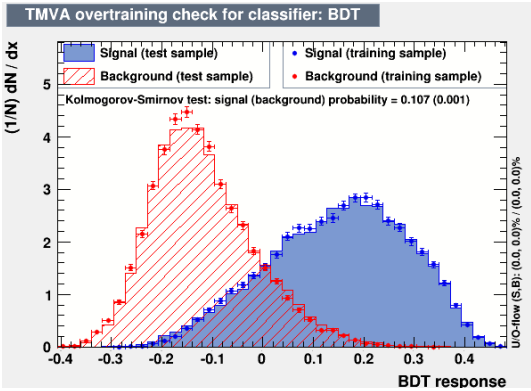


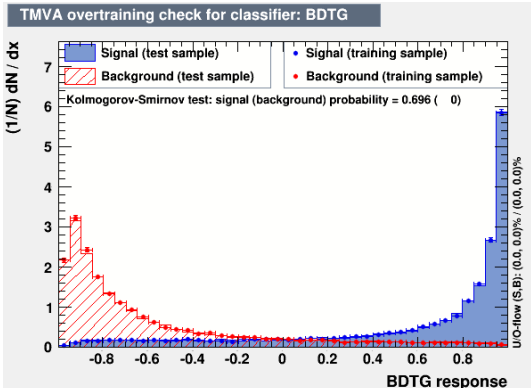




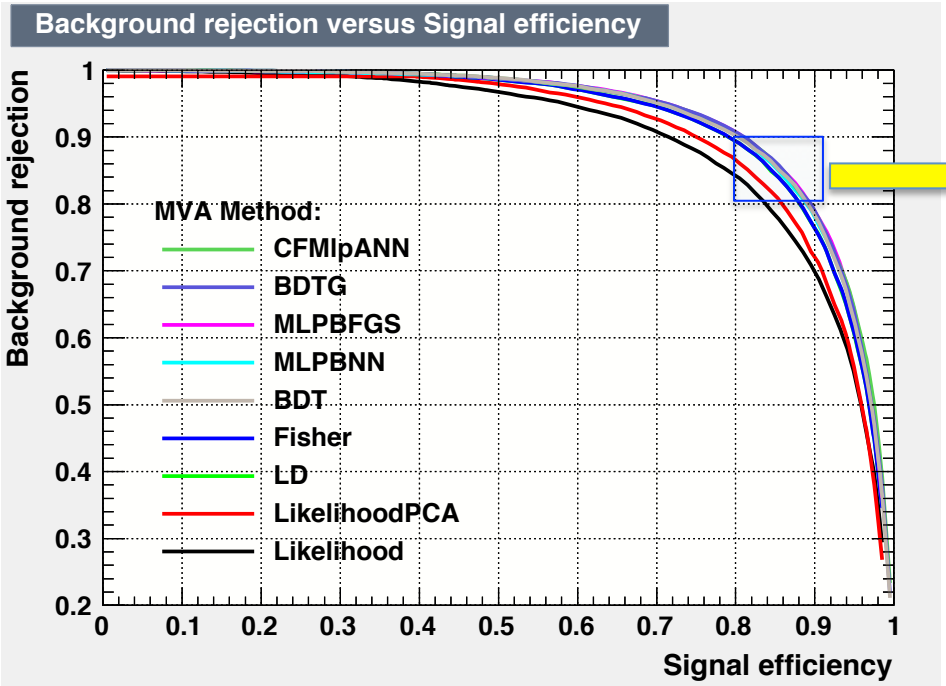


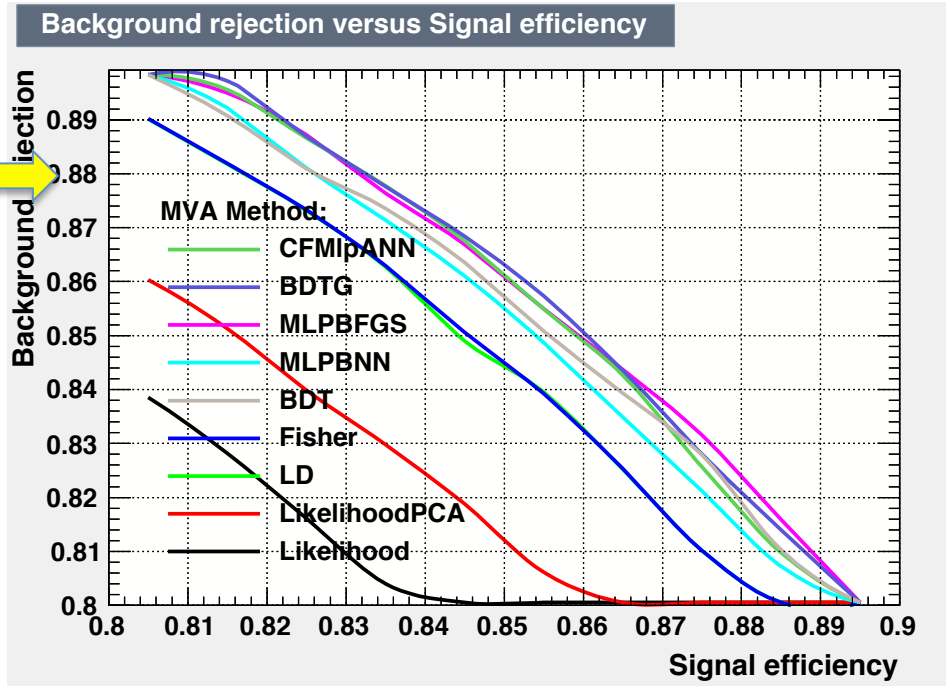






Background rejection vs signal efficiency





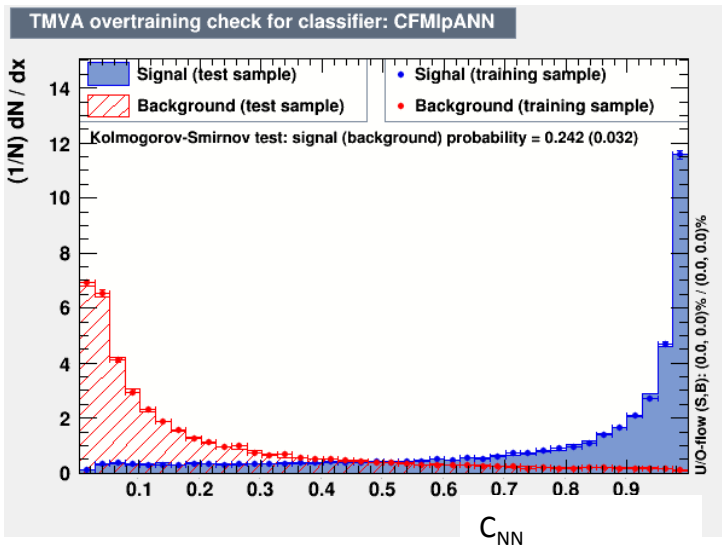
- BDTG and CFMlpANN exhibit the best performance and better shapes
- When the background rejection rate is about 85%, the signal efficiency is about 86%

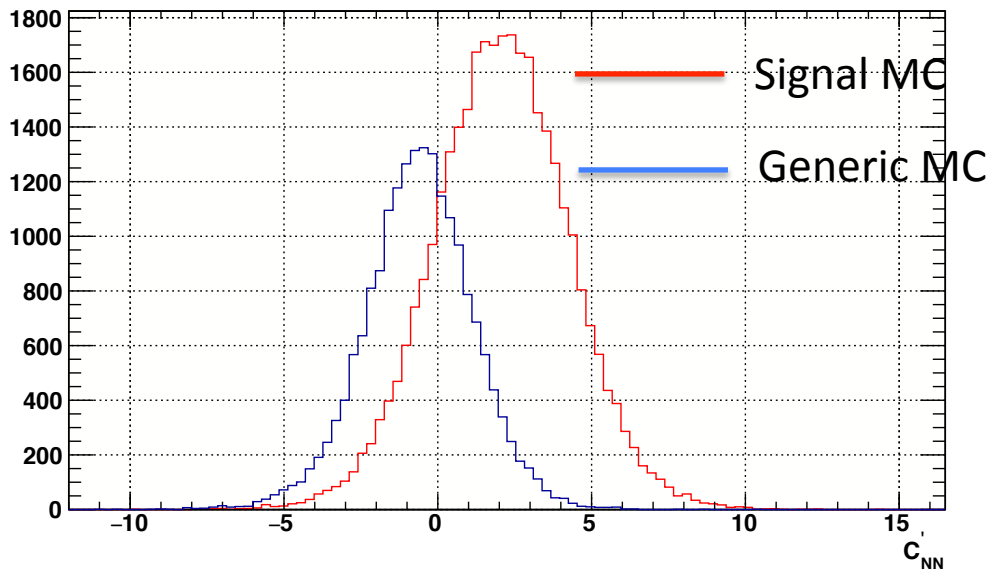
TMVA output (C_{NN}) for signal and generic MC (using CFMlpANN)

- Convert C_{NN} to C_{NN}'

$$C'_{NN} = \ln\left(\frac{C_{NN} - C_{NN}^{\min}}{C_{NN}^{\max} - C_{NN}}\right)$$

$C_{NN}^{\min} = 0.4$ and $C_{NN}^{\max} = 0.998$





Signal fitting

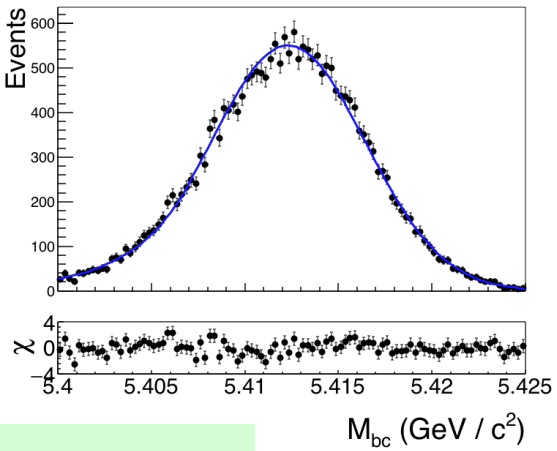
- Data set
  - Signal data set
    - 100K $\Upsilon(5S) \rightarrow B_s^* \bar{B}_s^*$ exclusive MC
  - Generic MC data set
    - 1 stream of $\Upsilon(5S)$ generic MC
- Signal region:
  - $\Delta E > -0.2$ and $\Delta E < 0.04$
  - $M_{bc} > 5.4$ and $M_{bc} < 5.425$
  - $C_{NN}(CFMlp) > 0.4$

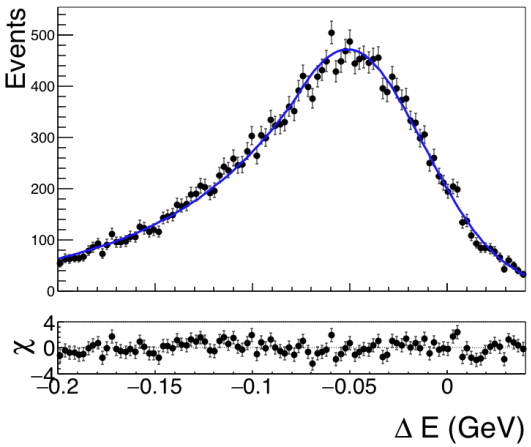
Modeling signal MC and generic MC

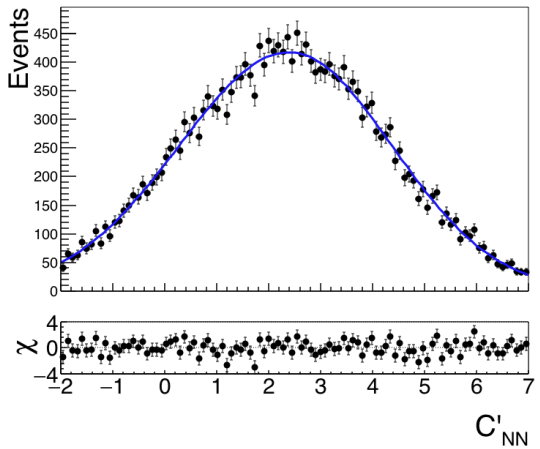
Components	M_{bc}	ΔE	C'_{NN}
Signal MC	Crystal Ball	Crystal Ball	Bifurcated Gaussian
Generic MC	Argus	2 order RooPolynomial	Bifurcated Gaussian

Signal fitting status

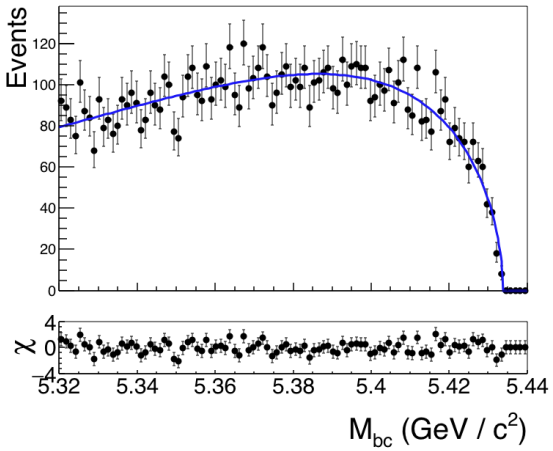
Signal MC

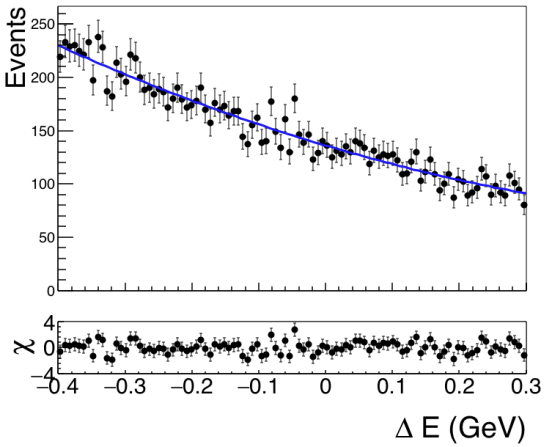


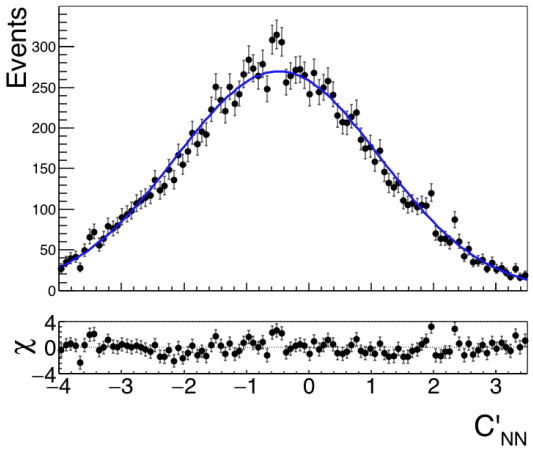




Generic MC







Summary

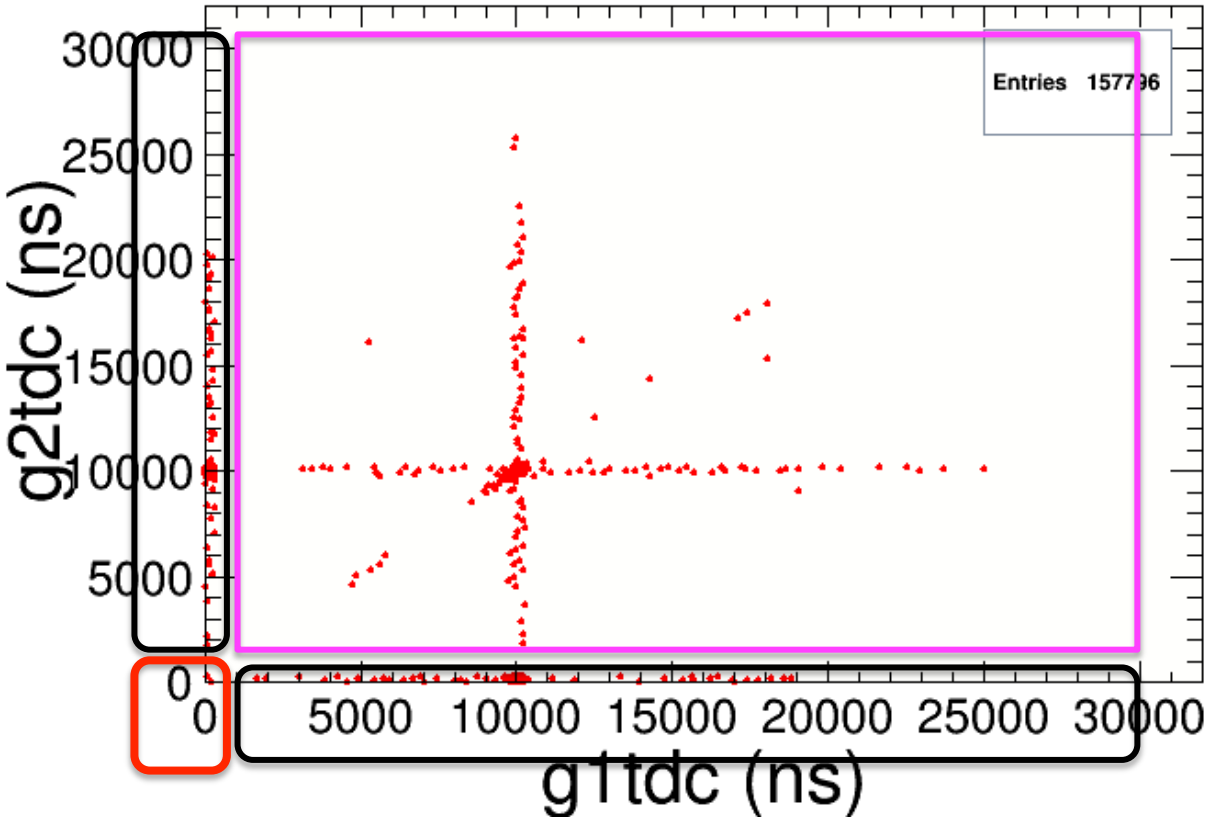
- Selection program has been written, we validated our program using signal MC; best candidate selection is now based on χ^2
- Generic MC is used to study background.
- TMVA is used to implement background suppression.
  - Identified BDTG and CFMlpANN as the best discriminators
  - The training and validation samples are in excellent agreement with each other. No indication of overtraining has been observed
  - For these two discriminators (BDTG and CFMlpANN), when the background rejection is about 85%, the signal efficiency is about 86%
  - Neural Network output (C_{NN}) and the converted variable C_{NN}' are now calculated in the analysis program
- Started to develop the fitting model

End

Additional slides

Pi0 daughter photons trigger TDCs

TDC values for daughter photons (one vs the other)



Red: both TDCs = 0 (8 events)

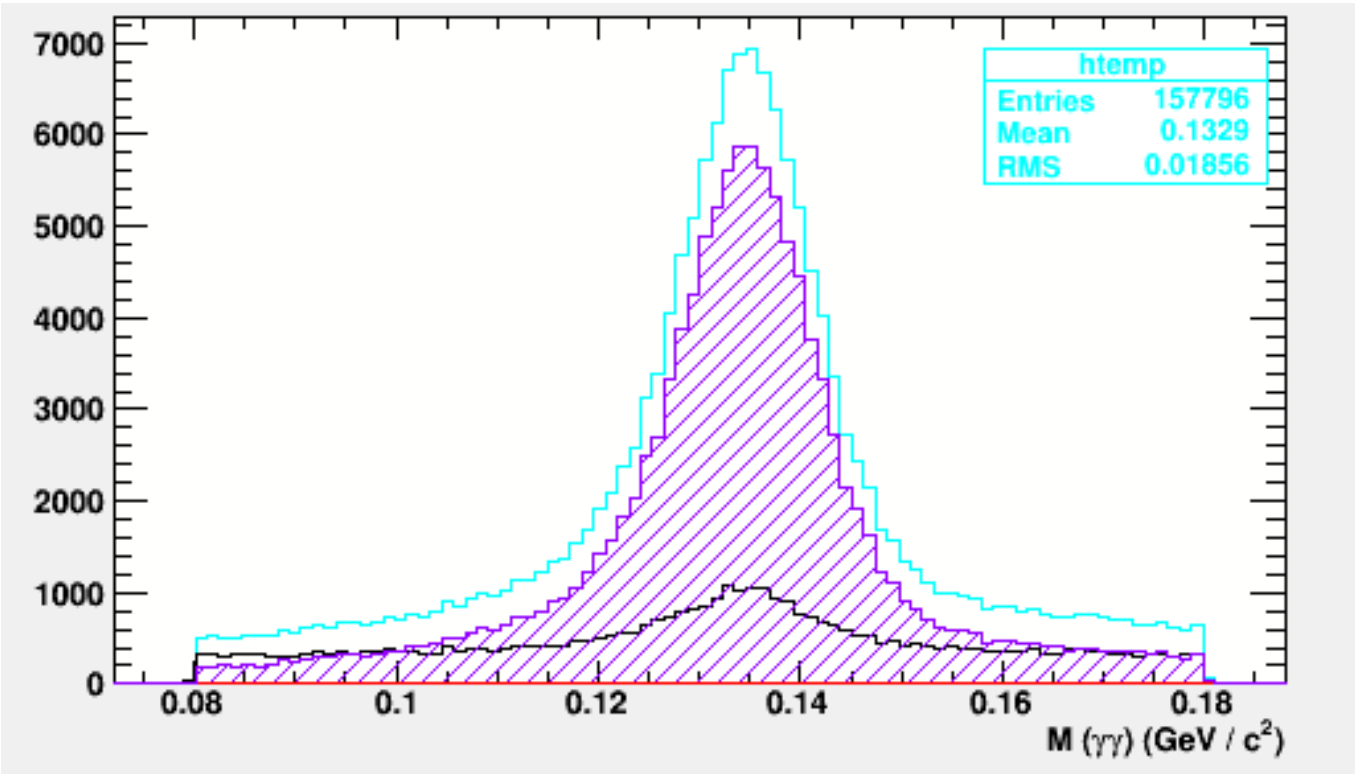
Black: only one TDC = 0 (40819 events)

Magenta: BOTH TDCs are != 0 (116977 events)

Thanks to Vikas Bansal for sharing with us his expertise concerning this matter

γγ invariant mass

All Y(5S) data after applying selection criteria for signal even candidates. Only signal pi0 candidates are shown.



Red: both TDCs = 0
(8 events)

Black: only one TDC = 0
(40819 events)

Magenta: BOTH TDCs are ! = 0
(116977 events)

Cyan: all events

No need to worry about out-of-time hits in Y(5S) data in this analysis