Axion-Like Particle Search in $B \to K^{(*)} + a$, $a \to \gamma\gamma$

Michael De Nuccio
Bertrand Echenard
Brian Shuve

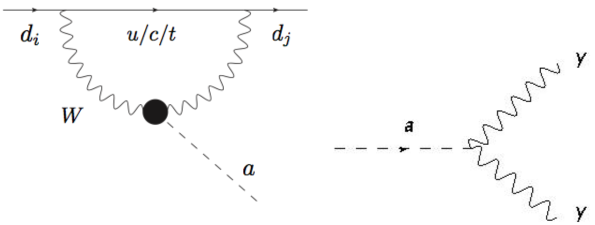
University of Pisa - INFN - BaBar

13 September 2017 - BaBar Jamboree

ALPs (Axion-Like Particles: pseudoscalar neutral massive particles) could couple to gauge bosons and induce FCNC.

$$\text{Coupling: } \mathcal{L} \supset -\frac{g_{aV}}{4} a V_{\mu\nu} \widetilde{V}^{\mu\nu}$$

Where:
$\widetilde{V}^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} V_{\rho\sigma}/2$, a and V are ALP and gauge boson fields.

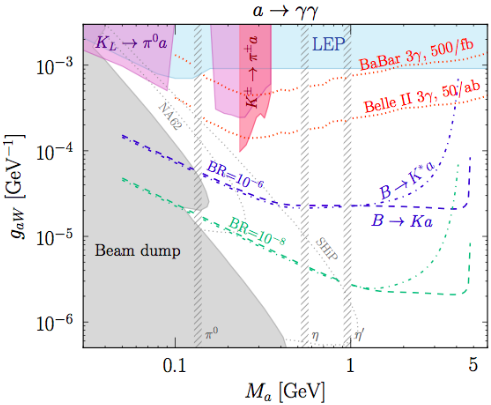


See Brian's paper: https://arxiv.org/pdf/1611.09355.pdf

Searched channel (diphoton resonance):

$$B \to K^{(*)} a, a \to \gamma\gamma$$

Exploits (often-overlooked) coupling of ALPs to $W^{\pm}$ $\Rightarrow$ FCNC production of ALPs (direct quark coupling unnecessary).



Overview

What was done till last time (June 2017):

- General overview of the ntuples;
- Characterize signal properties and resolution;
- Compare signal to bkgs (continuum and B^+B^-);
- Explore first discriminant variables;
- First steps into MVA.

What's **new**:

- Finalization of exploration of discriminant variables;
- Optimization of MVA method and strategy;
- Check stability of MVA over ALP masses and cuts;
- Validation MC bkgs with Run3 data;
- Fit and characterization of signal.

Backgrounds

Sources of bkg: $B\bar{B}$, Bhabha, continuum, $\tau^+\tau^-$, $\mu^+\mu^-$, ...
Main sources of bkg are continuum and $B\bar{B}$:

- Continuum ($e^+e^- \to q\bar{q}$; $q = uds + c$) generally really different behavior than signal but a random K and 2 γ can be reconstructed as signal $\to$ combinatorial bkg.
  Shape variables are powerful to discriminate: signal is uniform (Bs from Υ are slow), bkg is highly directional (jets).
- $B\bar{B}$ can look like signal, but $B \to K + X$, $X = \pi^0, \eta, \eta\prime \to \gamma\gamma$ have low BR and we veto over them; remains the combinatorial.

MVA - Strategies

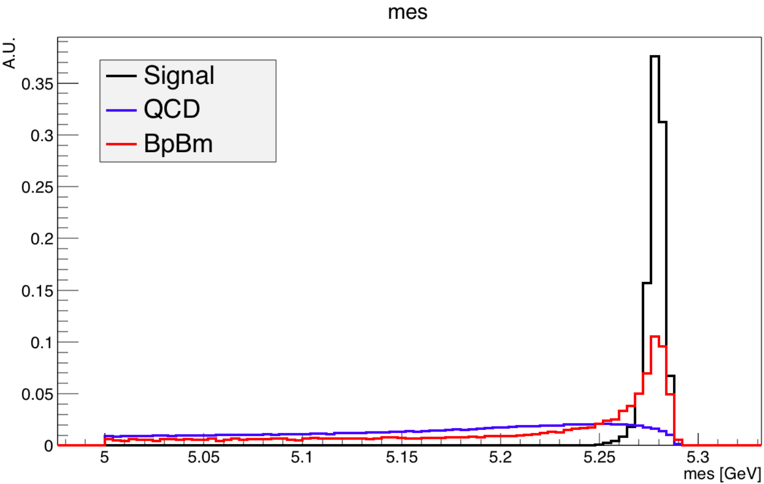
- Lot of variables $\Rightarrow$ **MVA**;
- Training MVA for the 2 main bkgs, check a posteriori that is good for the secondary bkgs too;
- Need to choose a training strategy:
  - Strategy 1step: train MVA for signal vs (B^+B^- and continuum) together, weighted, including m_{ES} and ΔE;
  - Strategy 2steps: train MVA for signal vs continuum, then separately for B^+B^-, including m_{ES} and ΔE;
  - Strategy 3steps: train MVA for signal vs continuum without m_{ES} nor ΔE, cut over m_{ES} and ΔE, train for signal vs B^+B^-.
- Also need to choose a MVA method (neural network, boosted decision tree, ...).

Discriminant variables

Variables used for selection:

- Δ_E;
- m_{ES};
- Legendre0;
- Legendre2;
- $\theta_{sphericity}$;
- θ_{thrust};
- γ energy;
- γ helicity angle;
- K helicity angle;
- K PID selector;
- Pi0Veto.

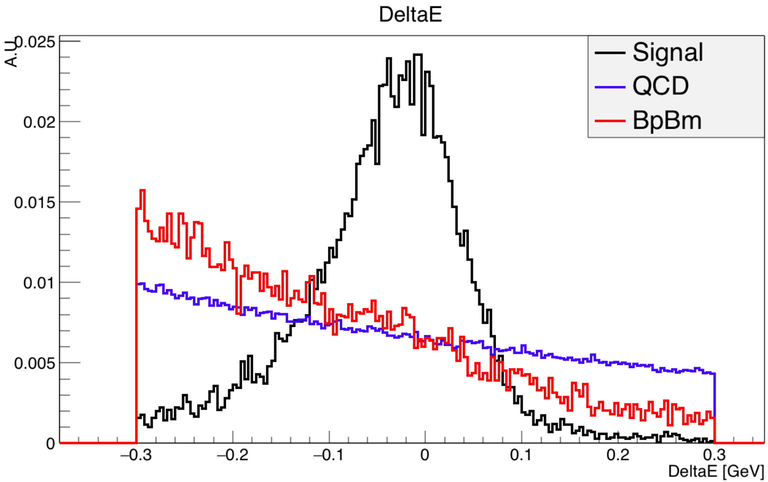
Discriminant variables
m_{ES}



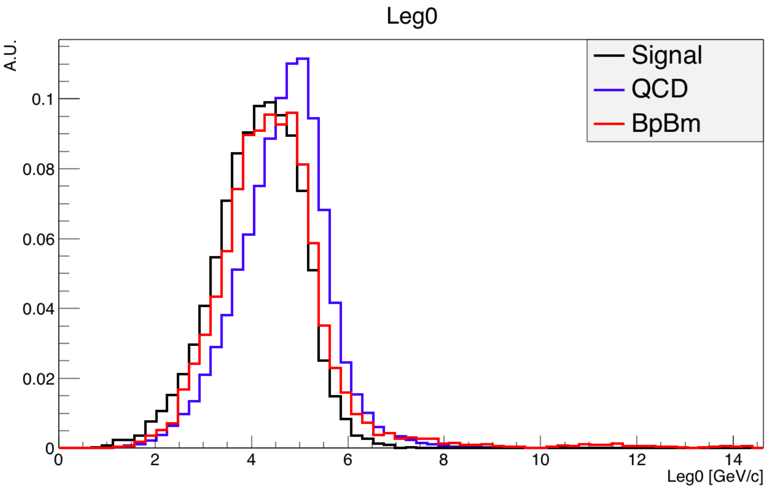
Discriminant variables

Δ_E

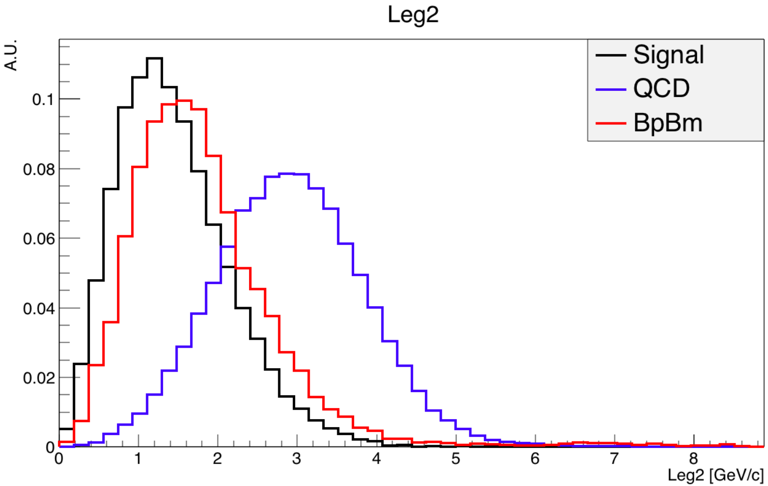
Pre-preselections applied at the time of ntuple generation.



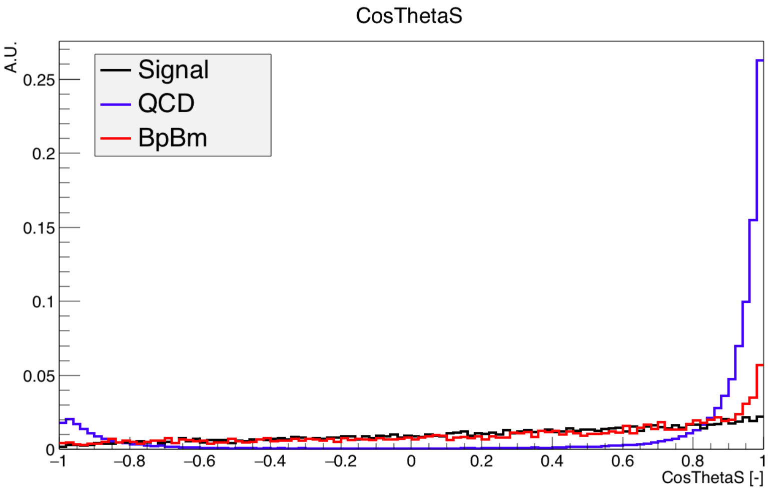
Legendre moment 0.



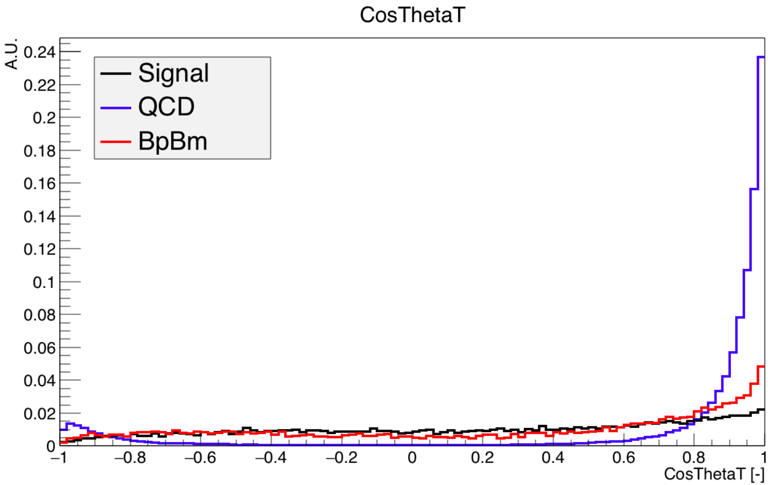
Legendre moment 2.



θ_S = sphericity angle (between sphericity axis of B and ROE).



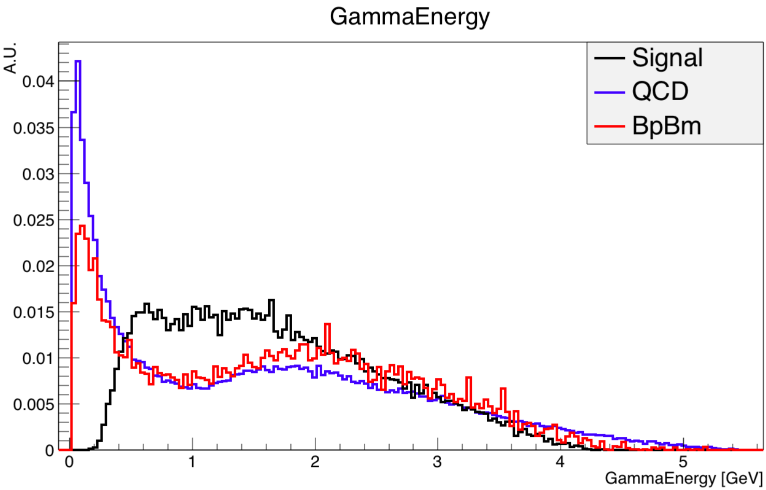
θ_T = thrust angle (between thrust axis of B and ROE).



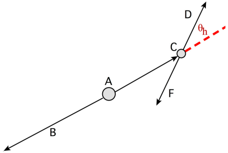
Discriminant variables

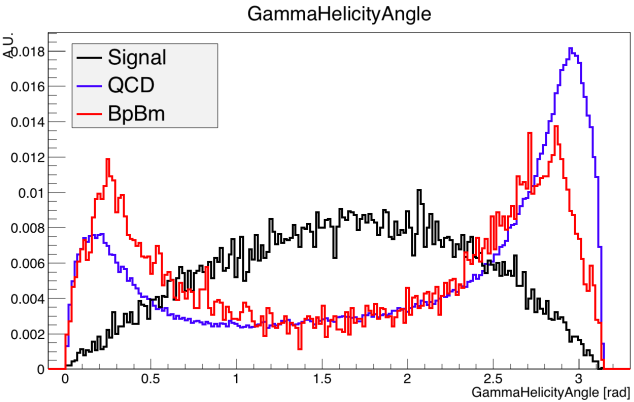
γ energy

Energy of γ from ALP; will lead to a preselection.

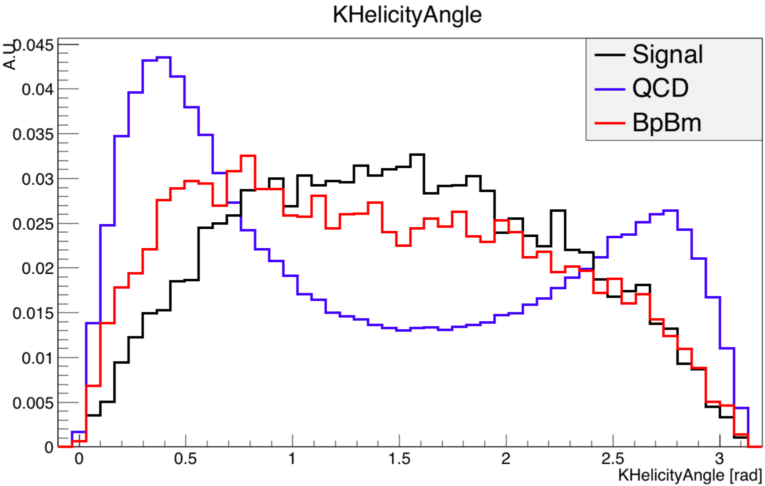


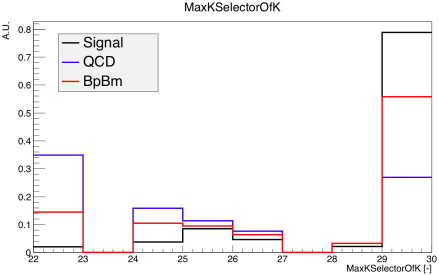
Helicity angle of γs from ALP.





Discriminant variables
K helicity angle



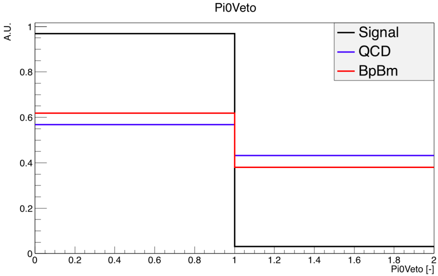


"K Selector" stands for "PID Selector".

A PID selector is passed if some quality requirements on the track are satisfied.

Highest the number (24-29), highest the probability that the K candidate is really a K.

- 29: SuperTightKMKaonMicroSelection
- 28: VeryTightKMKaonMicroSelection
- 27: TightKMKaonMicroSelection
- 26: LooseKMKaonMicroSelection
- 25: VeryLooseKMKaonMicroSelection
- 24: SuperLooseKMKaonMicroSelection
- 22: None of the above selectors has been passed



Pi0Veto is to eliminate contamination from true $X = \pi^0, \eta, \eta\prime$ (i.e. ALP candidates with mass close $\sim$ to X's).

True $X \to \gamma_A \gamma_B$: if in an event an ALP candidate is made with either γ_A or γ_B the Pi0Veto is 1.

Powerful to recognize bkg (Pi0Veto $= 1 \rightsquigarrow$ event is bkg).

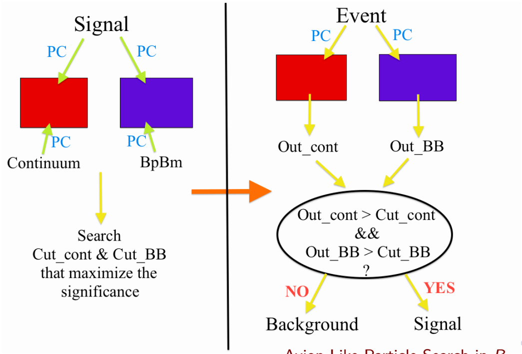
Strategy	Method	FoM
1step	MLP	139
	BDT	132
2steps	MLP	130
	BDT	169
3steps	MLP	152
	BDT	133

Table: MVA: comparison strategies and methods. 2 GeV signal used.

The FoM (Figure of Merit) is directly proportional to the actual significance, computed with continuum and B^+B^-:

$$\text{FoM} = \frac{N_{signal}^{passed}}{\sqrt{\sigma_{B^+B^-} \cdot \frac{N_{B^+B^-}^{passed}}{N_{B^+B^-}^{generated}} + \sigma_{continuum} \cdot \frac{N_{continuum}^{passed}}{N_{continuum}^{generated}}}} \quad (1)$$

1. Preselections: $\sim$ 30 MeV away from $X = \pi^0, \eta, \eta\prime$ masses; threshold for E_γ of 250 MeV;
2. $MVA_{continuum}$: train with Leg0, Leg2, $\cos(\theta_T)$, $\cos(\theta_S)$, MaxKSelectorOfK, m_{ES}, ΔE, γ helicity angle;
3. $MVA_{B^+B^-}$: train with $\cos(\theta_T)$, $\cos(\theta_S)$, Pi0Veto, m_{ES}, ΔE, K helicity angle, γ helicity angle.



Example of the distribution of the MVAs outs, for 1.5 GeV ALP mass training.

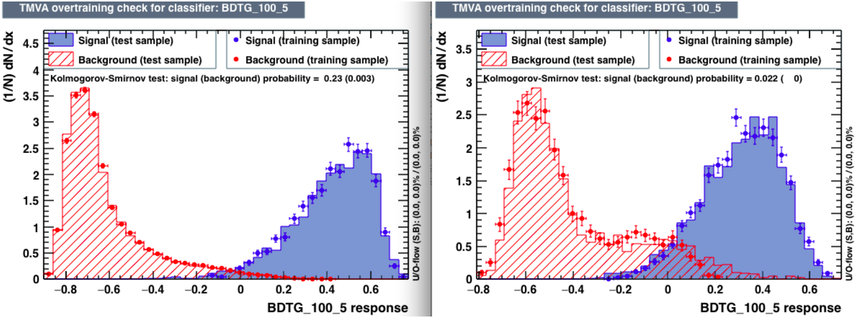
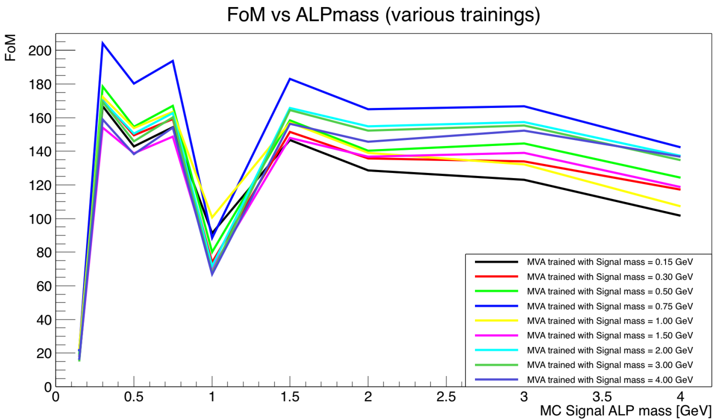
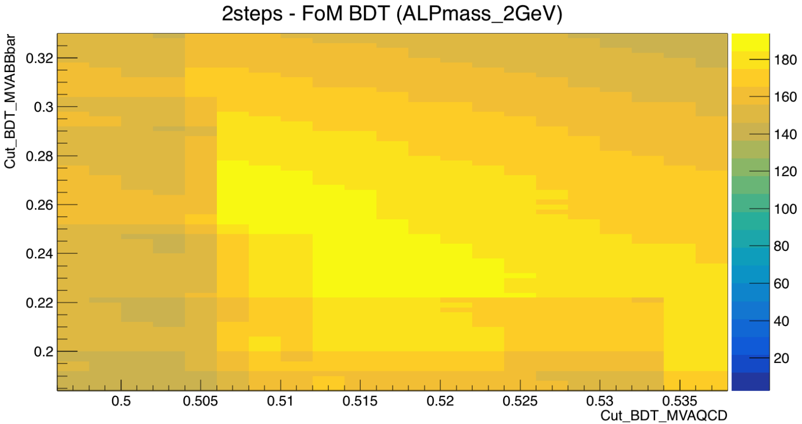


Figure: Left: continuum MVA out. Right: B^+B^- MVA out.

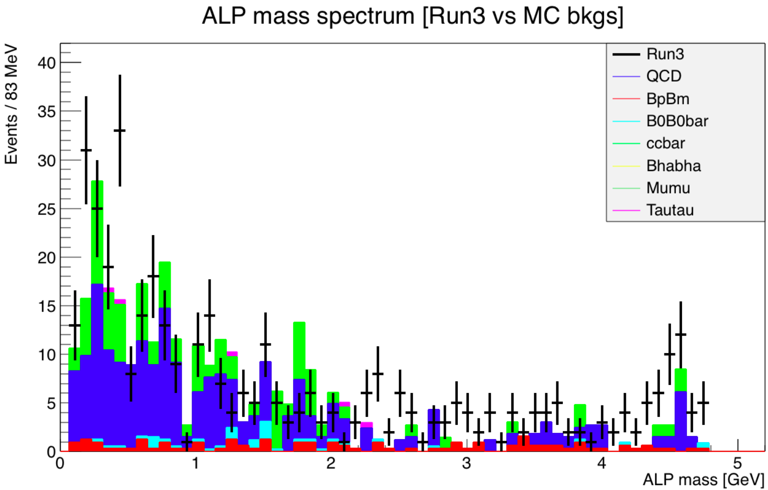
Intermediate value of cuts, interpolated between different masses (see next slide).
Significance slightly decreasing as ALP mass increases.
Drops near π^0 and $\eta\prime$ masses.



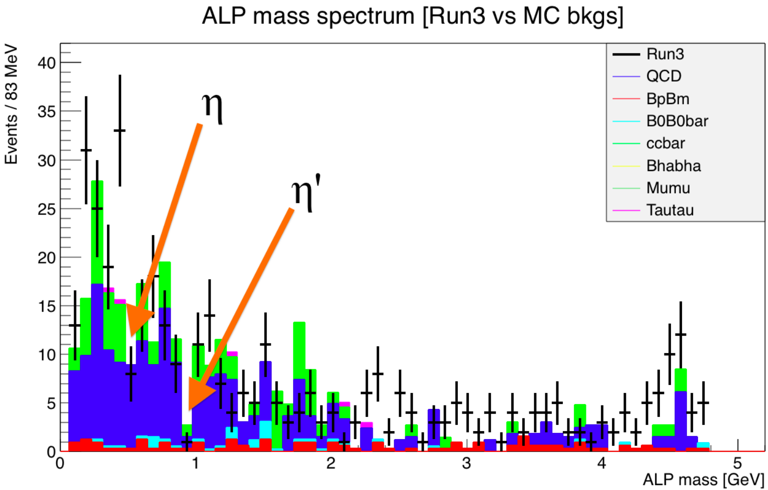
Searched not only point with highest significance but also stable in cuts variations.
Intersection between different MC Signal ALP masses' plots $\Rightarrow$ average cuts ($cut_{continuum} = 0.45$, $cut_{B^+B^-} = 0.16$).



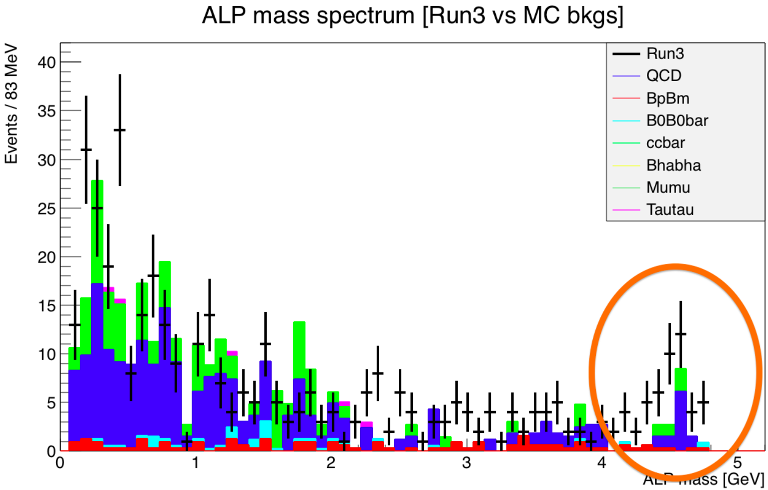
Good agreement between Run3 and MC bkgs.



We can appreciate the cuts around the vetoed masses.

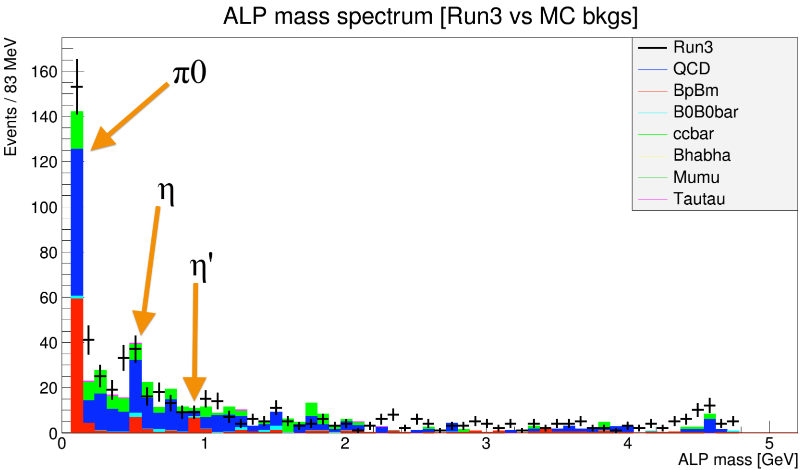


Bump at end of spectrum. Partially modeled by MC but worth investigation.



Run3 vs MC bkgs
Withous vetoing the critical masses

Same plot without vetoing the $\pi^0, \eta, \eta\prime$ masses.



Bump at $\sim$ 4.5 GeV could be caused by a combination of:

- geometrical acceptance;
- m_{ES} requirement.

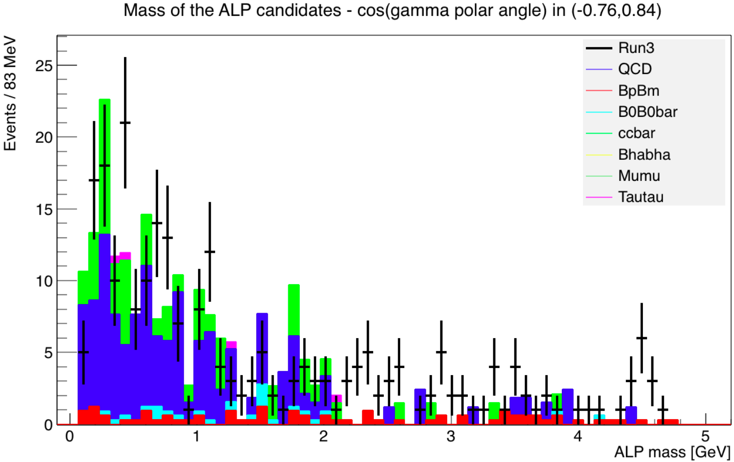
or by some not-well modeled bkgs. So we tried the following:

- cut on polar angle of γ from ALP;
- modifying m_{ES} window.

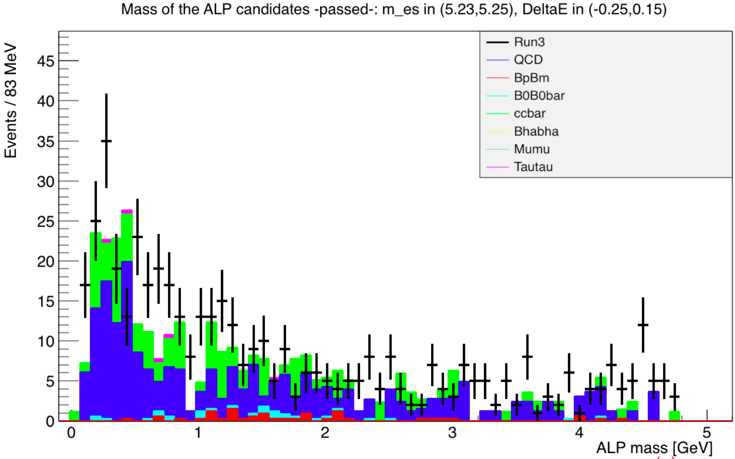
Run3 Bump

γ polar angle cut

Putting a cut on γ polar angle was not decisive: heavy loss in statistic and bump still present.

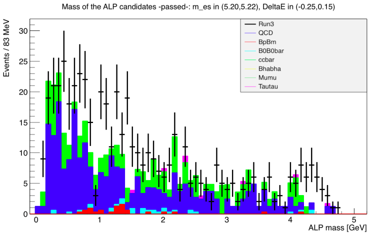


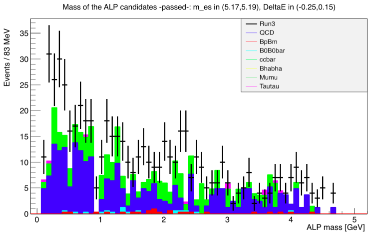
Here we trained the MVA with the same variables except m_{ES} and ΔE and cut over them at the end.
Bump is reduced.

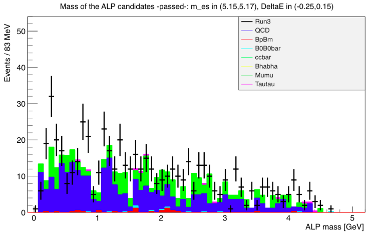


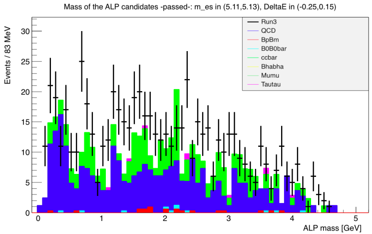
Run3 Bump

m_{ES} window modification - overview





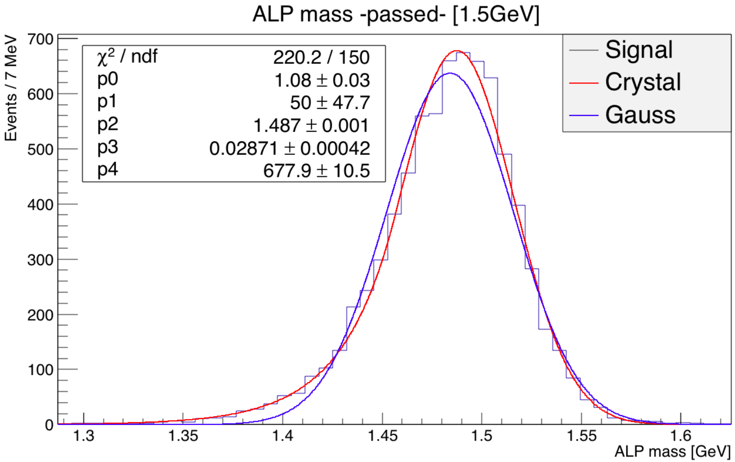




Fit signal with Gaussian and with **Crystal Ball** (CB), the second being significantly better (see χ^2 and plots).
Problem with parameter n of the CB, went too high and fitter crashed (n is an exponent); limited at max 50, has to be fixed.

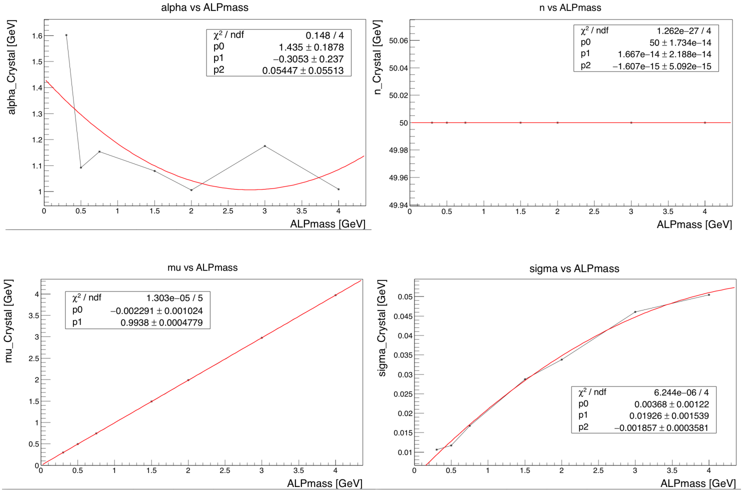
ALPmass	Gauss χ^2/DoF	CB χ^2/DoF (n max $= 50$)
0.3	$296/132 = 2.24$	$242/130 = 1.86$
0.5	$421/125 = 2.27$	$231/123 = 1.87$
0.75	$432/140 = 3.08$	$255/138 = 1.85$
1.5	$437/152 = 2.88$	$220/150 = 1.47$
2	$369/110 = 3.60$	$151/108 = 1.39$
3	$228/64 = 3.58$	$83/62 = 1.34$
4	$286/66 = 4.34$	$95/64 = 1.49$

Table: Signal fit: Gaus vs CB χ^2_{red}.



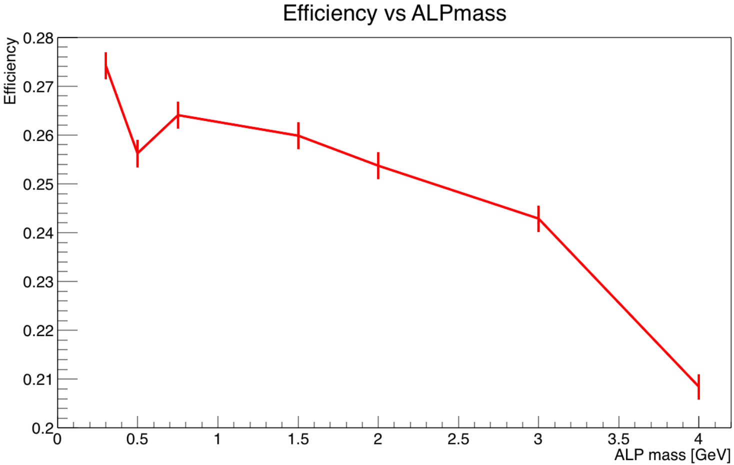
$p0 = \alpha$, $p1 = n$, $p2 = \mu$, $p3 = \sigma$, $p4 = N$.

Signal fit
CB parameters fits



Signal efficiency

Using the average MVA cuts we obtain high efficiency:



What's next?

- Finalize the investigation about the bump;
- Fix the CB fitting procedure;
- Use of flattened mass signal to investigate mass-dependent variables (as γ energy or γ polar angle);
- Simulate signal together with bkg to compute yield extraction efficiency;
- Systematic errors.

Thank you!

Backup slides

$$m_{ES} = \sqrt{\left(\frac{\frac{s}{2} + \vec{p}_i \cdot \vec{p}_B}{E_i}\right)^2 - \vec{p}_B^2}$$

With E_i, $\vec{p}_i$ referred to initial four-momentum and $\vec{p}_B$ is the B candidate momentum, both evaluated in the laboratory frame.

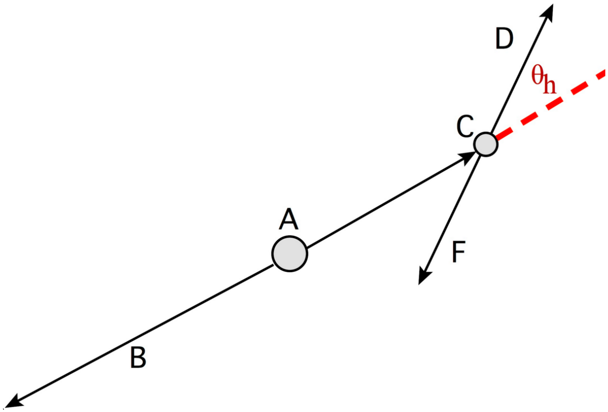
$$\Delta_E = E_B - \frac{\sqrt{s}}{2}$$

Where E_B is the energy of the B cand in the CM frame.

$$L_k = \sum_{j \in ROE} p_j cos^k(\theta_j)$$

where θ_j is the angle between B thrust axis and the particle in the sum (ROE = Rest Of the Event).

Helicity angle
Definition



The **thrust axis** $\vec{T}$ is the unit vector along which the projection of the total momentum of group of particle is maximized.

θ_T is the angle between the thrust axis of the B candidate and the one of the ROE.

The **sphericity tensor** is given by: $S^{\alpha,\beta} = \frac{\sum_{i=1}^{N} p_i^\alpha p_i^\beta}{\sum_{i=1}^{N} |\vec{p}|^2}$ and has 3 eigenvalues and eigenvectors.

Sphericity axis is the versor of the eigenvector with the highest eigenvalue.

θ_S is the angle between the sphericity axis of the B candidate and the one of the ROE.

$$f(x;\alpha,n,\bar{x},\sigma) = N \cdot \begin{cases} \exp(-\frac{(x-\bar{x})^2}{2\sigma^2}), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\bar{x}}{\sigma})^{-n}, & \text{for } \frac{x-\bar{x}}{\sigma} \leqslant -\alpha \end{cases}$$

where

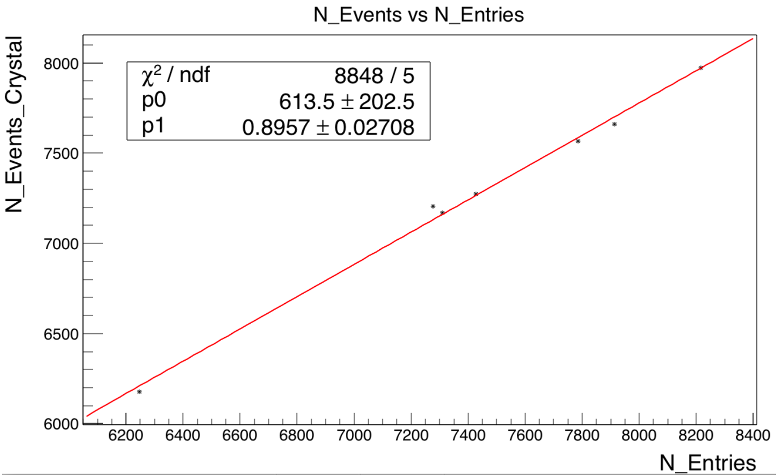
$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right),$$

$$B = \frac{n}{|\alpha|} - |\alpha|,$$

$$N = \frac{1}{\sigma(C+D)},$$

$$C = \frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right),$$

$$D = \sqrt{\frac{\pi}{2}}\left(1 + \operatorname{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right).$$



N_Events vs N_Entries
N_Events_Crystal
N_Entries
χ² / ndf 8848 / 5
p0 613.5 ± 202.5
p1 0.8957 ± 0.02708
8000
7500
7000
6500
6000
6200
6400
6600
6800
7000
7200
7400
7600
7800
8000
8200
8400

Pre-preselections

The following pre-preselections are applied at the time of ntuple generation:

- K candidate has to be at least a "KCombinedSuperLoose";
- γ candidate has to be at least "GoodPhotonsLoose";
- $\Delta E \in (-0.3; 0.3)$ GeV;
- $m_{ES} \in (5.0; 5.4)$ GeV;
- $m_B \in (4.5; 6.0)$ GeV.

Away from $X = \pi^0, \eta, \eta\prime$ masses:

$$|m_{ALP} - m_{\pi^0}| < 25 \text{ MeV} \;\wedge\; |m_{ALP} - m_{\eta}| < 30 \text{ MeV} \;\wedge \quad (2)$$

$$\wedge\; |m_{ALP} - m_{\eta\prime}| < 35 \text{ MeV} \quad (3)$$

Threshold for E_γ:

$$E_\gamma > 250 \text{ MeV} \quad (4)$$

MLP stands for Multilayer perceptron, a kind of neural network.
BDT stands for Boosted Decision Tree.

$MVA_{continuum}$:

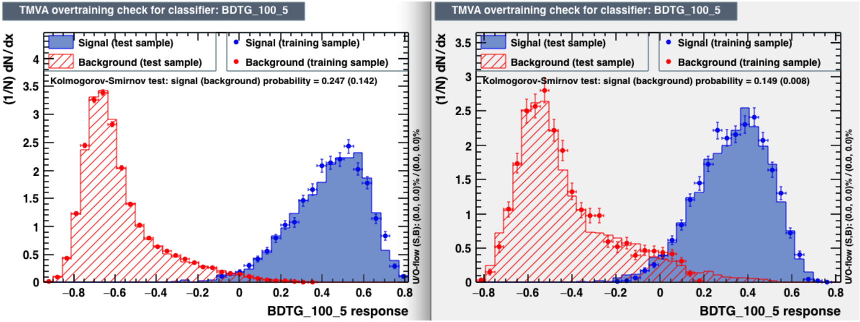
$$MVA_{out} > 0.45 \tag{5}$$

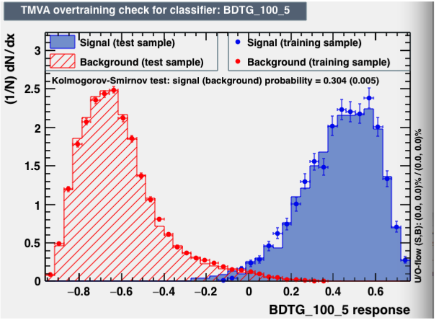
$MVA_{B^+B^-}$:

$$MVA_{out} > 0.16 \tag{6}$$

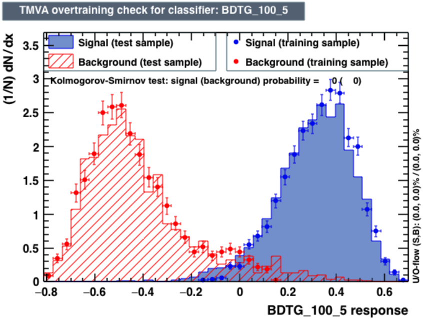
MVA out

0.3 GeV



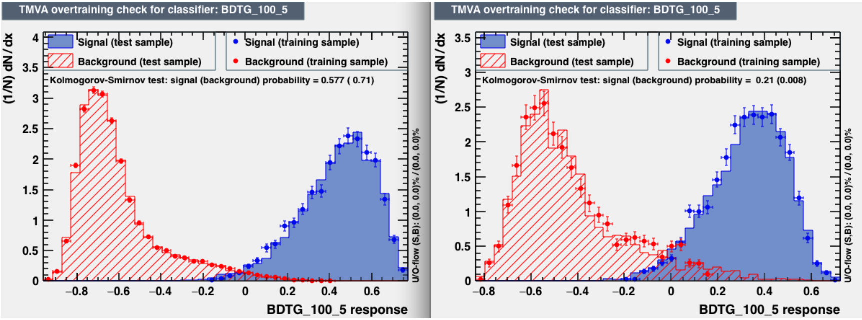


TMVA overtraining check for classifier: BDTG_100_5
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.304 (0.005)
(1/N) dN/dx
BDTG_100_5 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

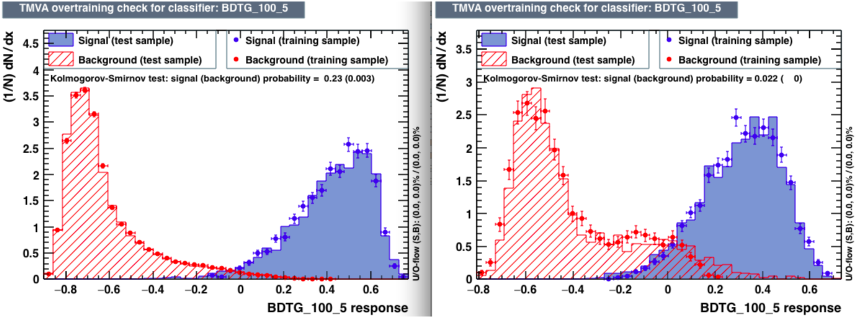


TMVA overtraining check for classifier: BDTG_100_5
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability =
(1/N) dN/dx
BDTG_100_5 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

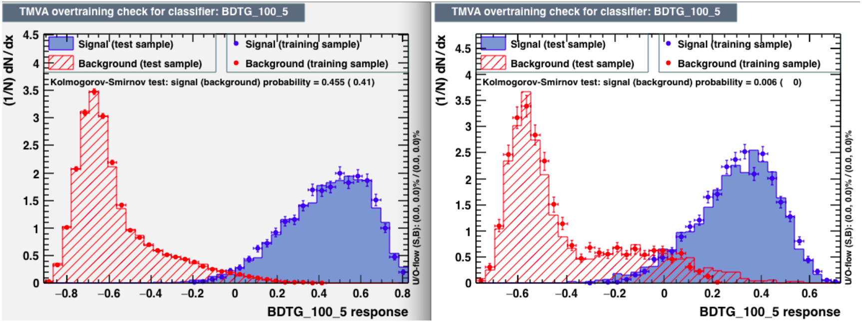
MVA out
0.75 GeV



MVA out
1.5 GeV

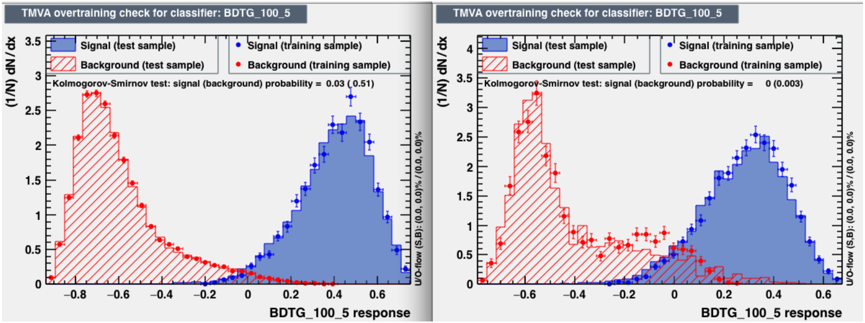


MVA out
2 GeV

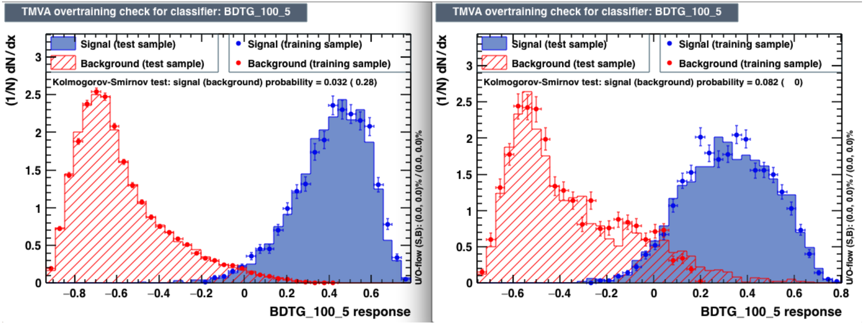


TMVA overtraining check for classifier: BDTG_100_5
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.455 (0.41)
(1/N) dN / dx
BDTG_100_5 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%
TMVA overtraining check for classifier: BDTG_100_5
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.006 (0)
(1/N) dN / dx
BDTG_100_5 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

MVA out
3 GeV


TMVA overtraining check for classifier: BDTG_100_5
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.03 (0.51)
(1/N) dN / dx
BDTG_100_5 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%
TMVA overtraining check for classifier: BDTG_100_5
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0 (0.003)
(1/N) dN / dx
BDTG_100_5 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

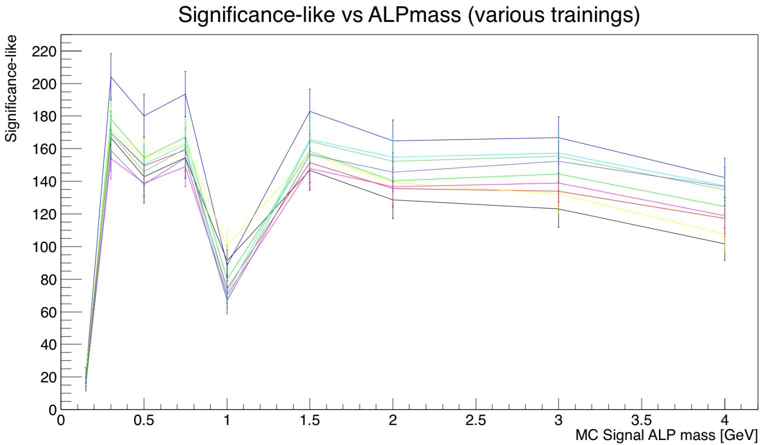
MVA out
4 GeV



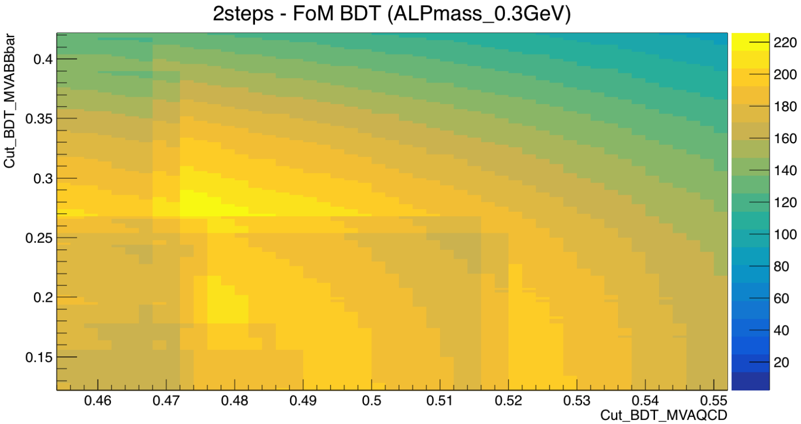
TMVA overtraining check for classifier: BDTG_100_5
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.032 (0.28)
(1/N) dN / dx
BDTG_100_5 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%
TMVA overtraining check for classifier: BDTG_100_5
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.082 (0)
(1/N) dN / dx
BDTG_100_5 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

MVA stability for different ALP masses

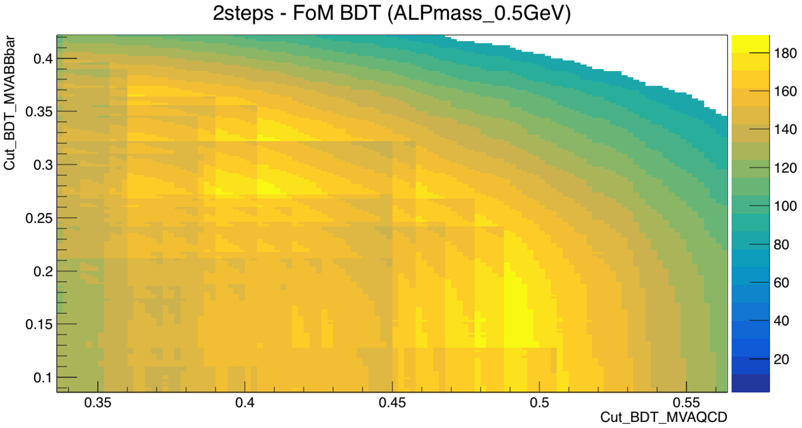
Uncertainties



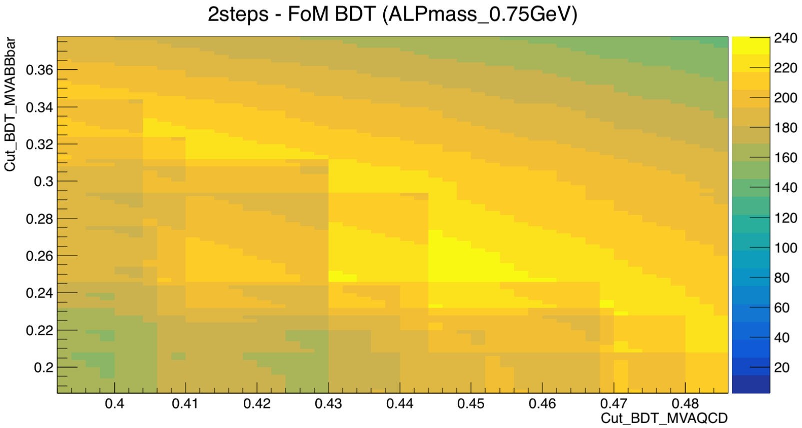
MVA stability in cut space
0.3 GeV



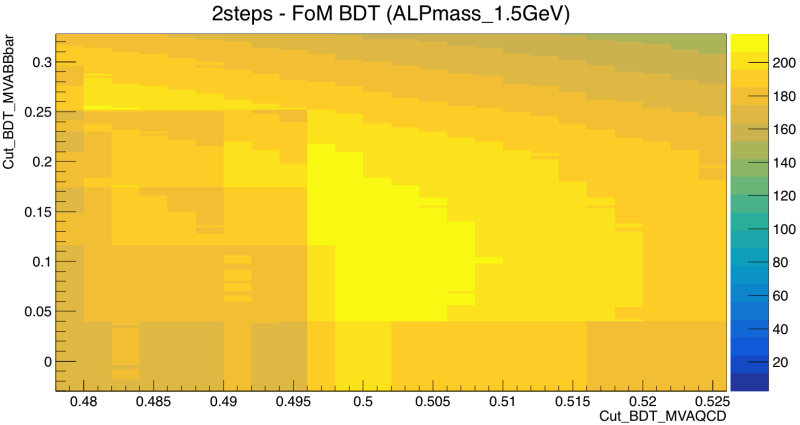
MVA out
0.5 GeV



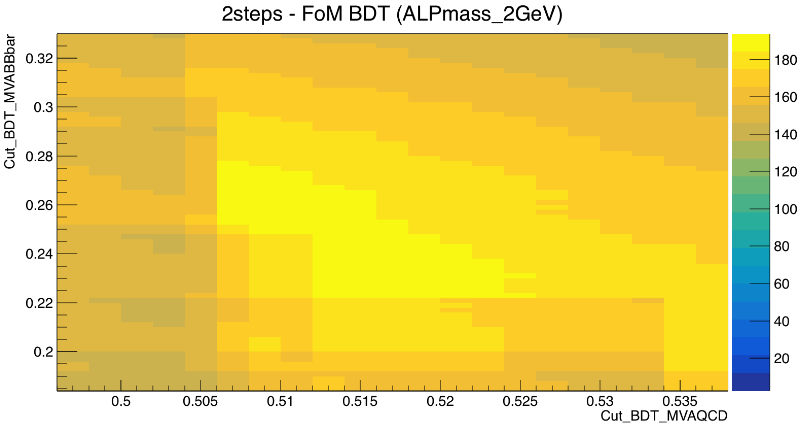
MVA out
0.75 GeV



MVA out
1.5 GeV

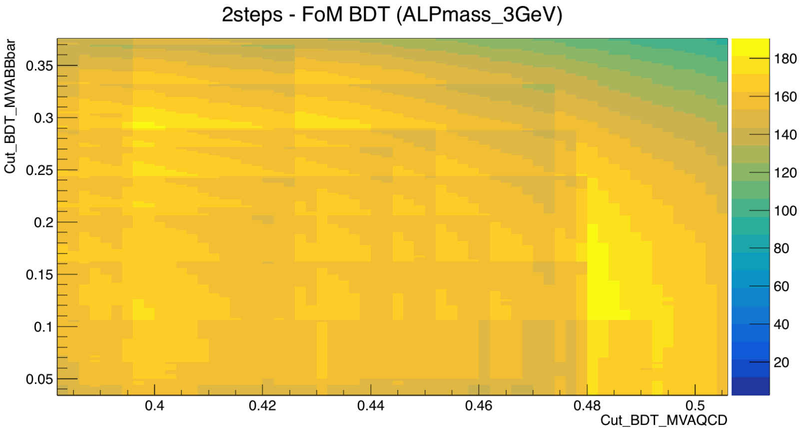


MVA out
2 GeV



2steps - FoM BDT (ALPmass_2GeV)
Cut_BDT_MVABBbar
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.5
0.505
0.51
0.515
0.52
0.525
0.53
0.535
Cut_BDT_MVAQCD
180
160
140
120
100
80
60
40
20

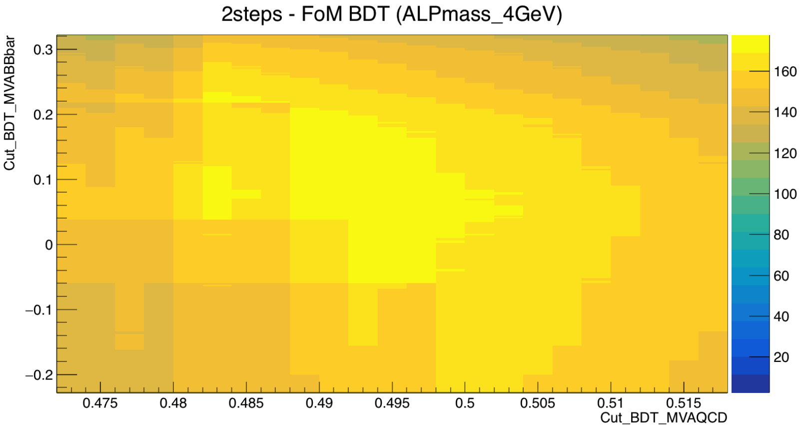
MVA out
3 GeV



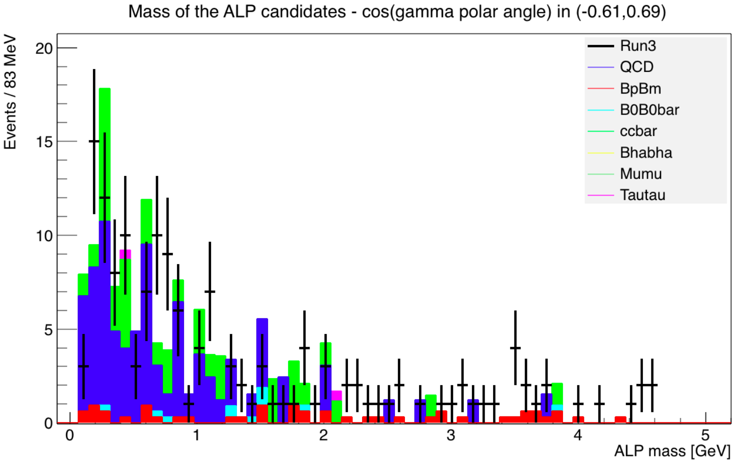
2steps - FoM BDT (ALPmass_3GeV)
Cut_BDT_MVABBbar
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0.4
0.42
0.44
0.46
0.48
0.5
Cut_BDT_MVAQCD
180
160
140
120
100
80
60
40
20

MVA out

4 GeV

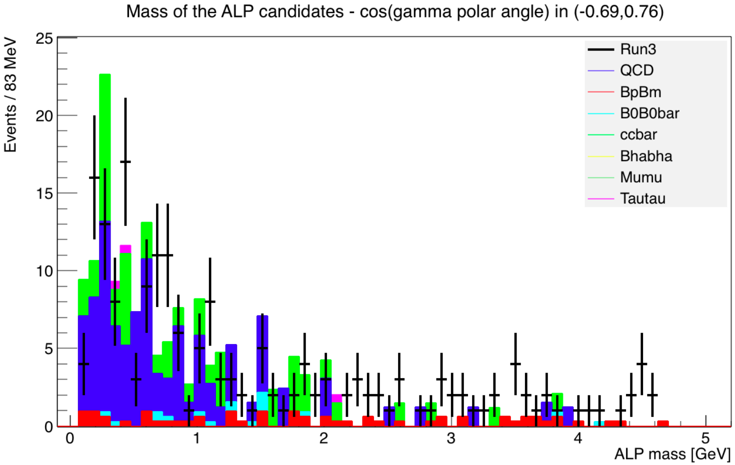

2steps - FoM BDT (ALPmass_4GeV)
Cut_BDT_MVABBbar
0.3
0.2
0.1
0
−0.1
−0.2
0.475
0.48
0.485
0.49
0.495
0.5
0.505
0.51
0.515
Cut_BDT_MVAQCD
160
140
120
100
80
60
40
20

Polar angle of γs from ALP candidate.

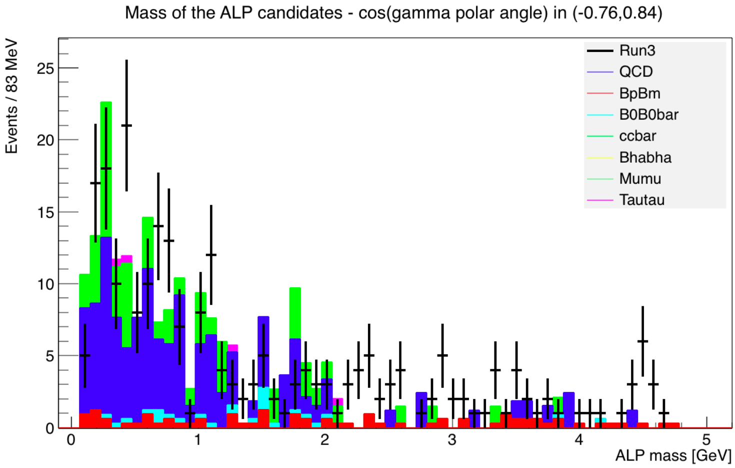


Polar angle investigation
cos(γ polar angle) in (-0.69;0.76)

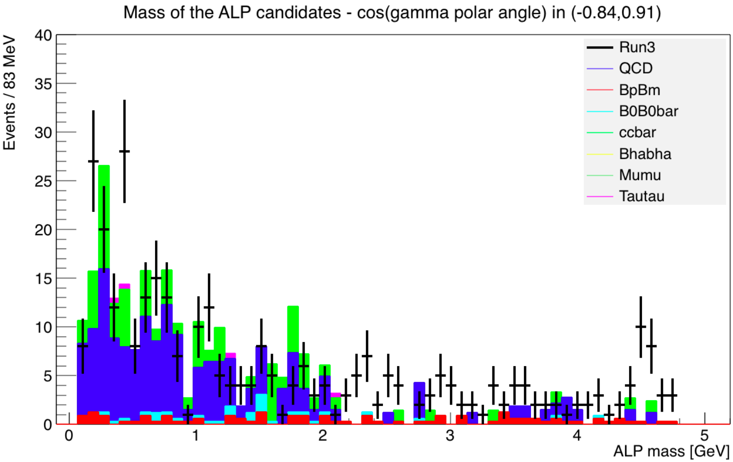
Polar angle of γs from ALP candidate.



Polar angle of γs from ALP candidate.

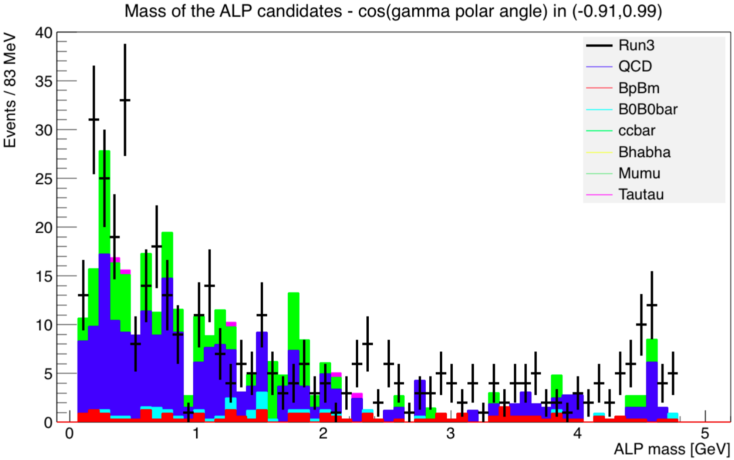


Polar angle of γs from ALP candidate.

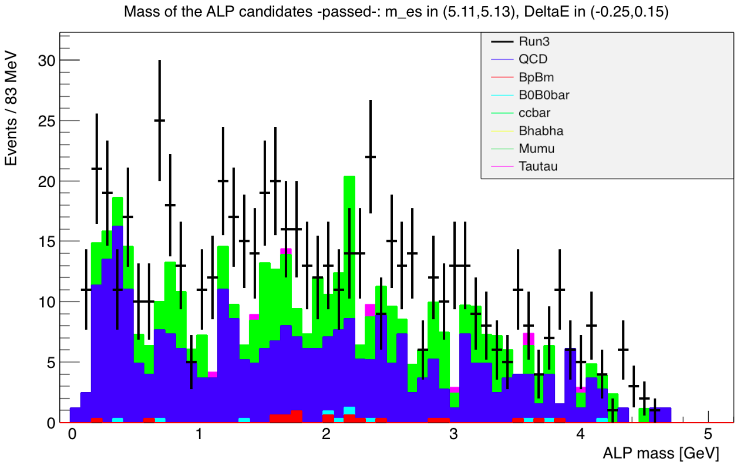


Polar angle investigation
cos(γ polar angle) in (-0.91;0.99)

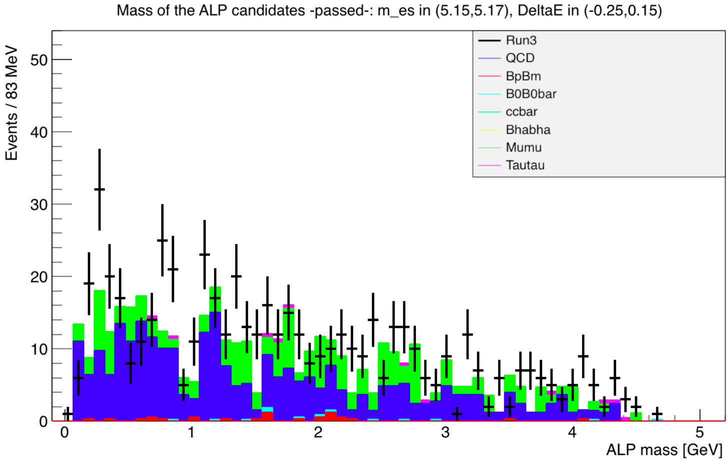
Polar angle of γs from ALP candidate.



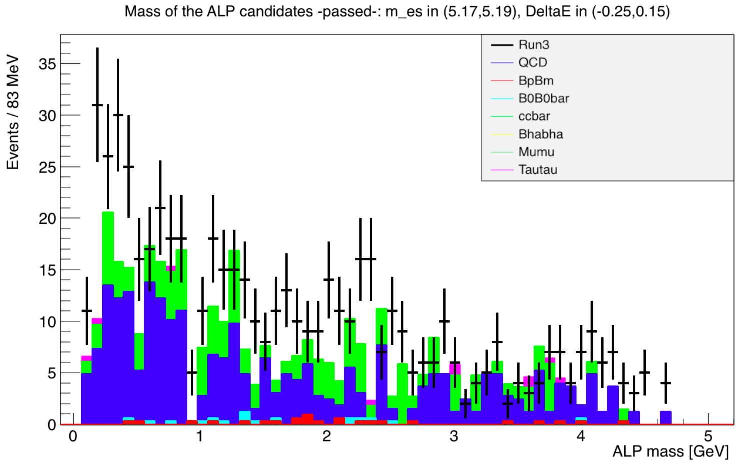
m_{ES} window
5.11 - 5.13

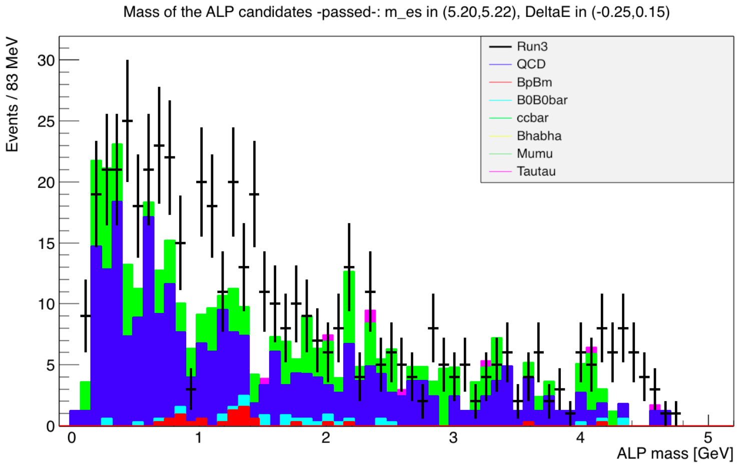


m_{ES} window
5.15 - 5.17



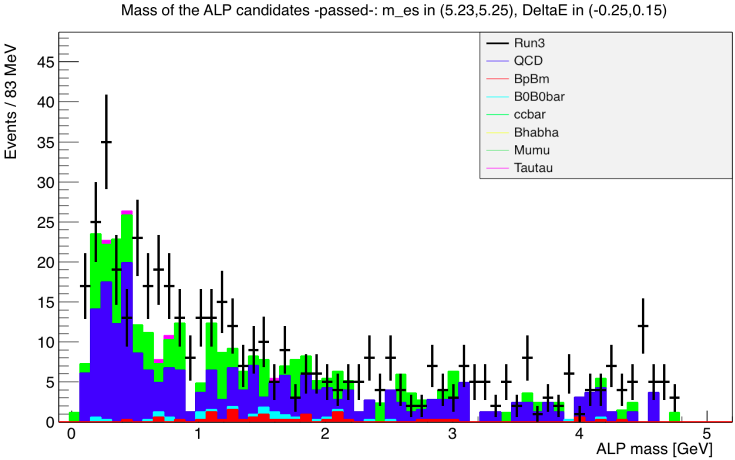
m_{ES} window
5.17 - 5.19





Mass of the ALP candidates -passed-: m_es in (5.20,5.22), DeltaE in (-0.25,0.15)
Events / 83 MeV
ALP mass [GeV]
Run3
QCD
BpBm
B0B0bar
ccbar
Bhabha
Mumu
Tautau

m_{ES} window
5.23 - 5.25



References

- `https://arxiv.org/pdf/1611.09355.pdf` paper from which the analysis is based on;
- `https://arxiv.org/pdf/0707.2798.pdf` comparison paper for $B \to K\pi^0$.