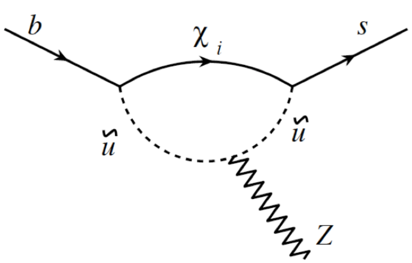
$B \to \tau\nu$ and $B \to K^*\nu\bar{\nu}$ status

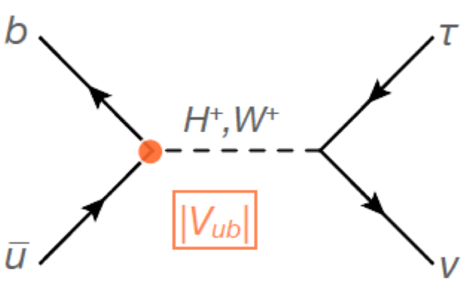
Mario Merola, Elisa Manoni, Claudia Cecchi, Guglielmo De Nardo

Belle II Italia, 23-05-18

Outline

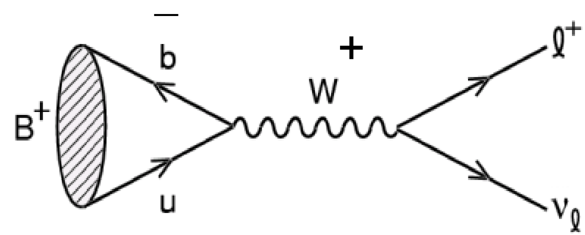
- $B \to \tau\nu$ analysis summary
- Intermezzo: MC9 / MC10 comparison and background impact study
- $B \to K^*\nu\bar{\nu}$ analysis summary





B leptonic decays ($B \to l\nu$)

- Helicity suppressed



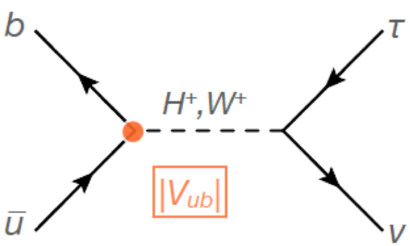
$$BR_{SM}(B \to \ell\, \nu) = \frac{G_F^2 m_B \tau_B}{8\pi} f_B^2 \,|V_{ub}|^2 m_\ell^2 \left[1 - \frac{m_\ell^2}{m_B^2}\right]^2$$

$\tau:\mu:e \to 1 : 10^{-3} : 10^{-7}$

- The SM predicts a branching ratio of $\mathcal{B}(B^+ \to \tau^+ \nu_\tau) = 0.817^{+0.054}_{-0.031} \times 10^{-4}$

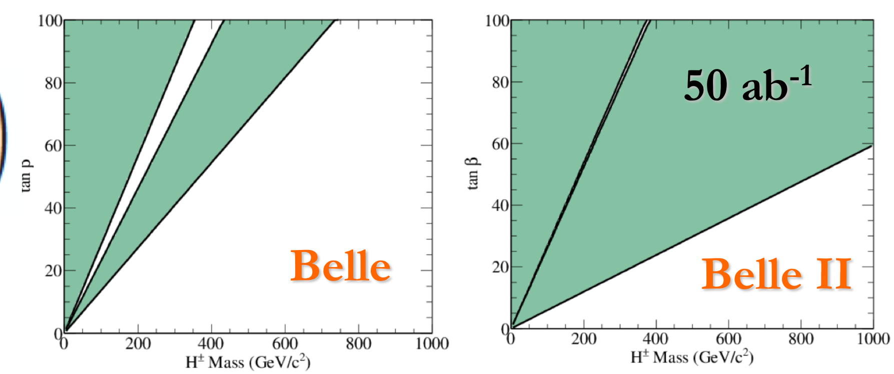
http://ckmfitter.in2p3.fr/

Higgs doublet models predict interference with SM decay with a modification of the branching ratio [PhysRevD.86.054014]



$$B = B_{SM} \times \left(1 - m_B^2 \frac{\tan^2\beta}{m_{H^\pm}^2}\right)^2$$

ratio of the two Higgs vacuum expectation values

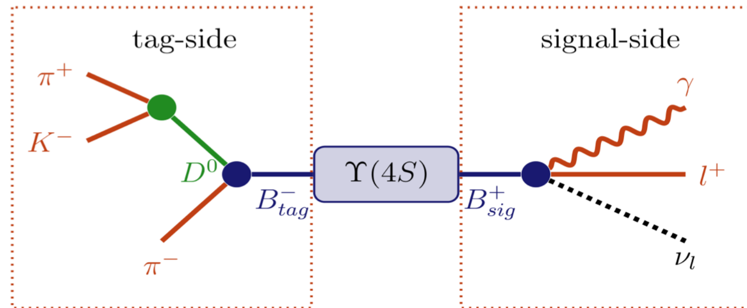


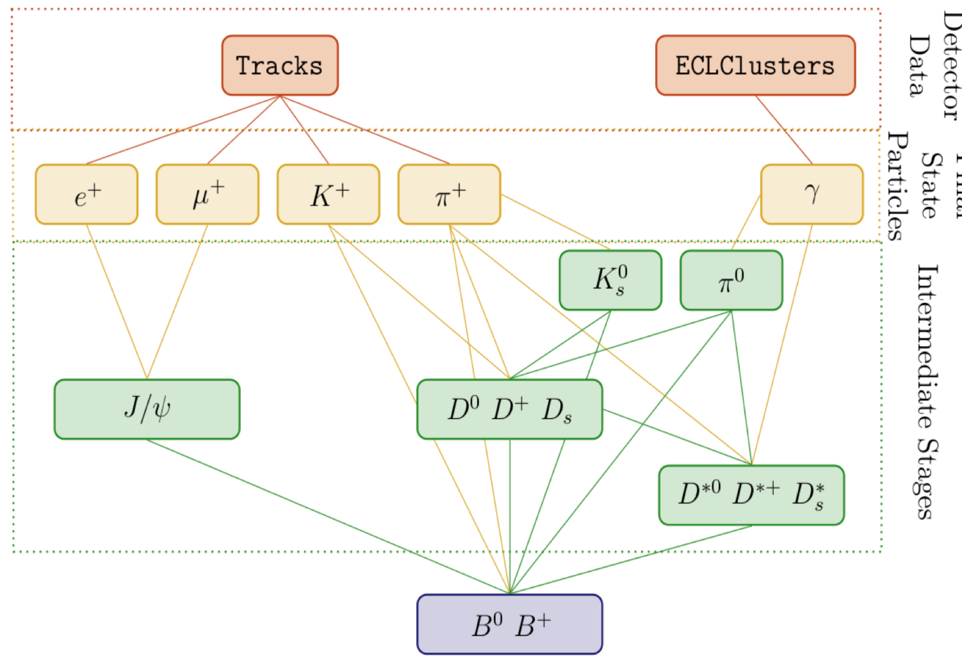
Tag side reconstruction with Full Event Interpretation (FEI)

- Developed by Thomas Keck*, it's an extension of the Full Reconstruction used in Belle, and uses a multivariate technique to reconstruct the B-tag side through lots of decay modes in a Υ(4S) decay.

- Hierarchical approach: first train multivariate classifiers (MVC) on FSP, then reconstruct intermediate particles and build new dedicated MVC. For each candidate a signal probability ("sigprob") is defined, which represents the "goodness" of its reconstruction.

- Training used here is performed on 100*10^6 $B^+B^-/B^0\bar{B}^0$ events with beam background, MC7 campaign





*https://ekp-invenio.physik.uni-karlsruhe.de/record/48602/files/EKP-2015-00001.pdf

Selection (1)

B tag side

Hadronic tag using FEI

1) Pre-selection on B-tag kinematics*
2) Loose cut on FEI output discriminant
3) Pick the highest sigprob B candidate

* Beam-constrained mass: $M_{bc} = \sqrt{E_{beam}^{*2} - p_B^{*2}}$

* Energy difference: $\Delta E = E_B^* - E_{beam}^*$

B sig side

$B \to \tau\nu$

- 4 tau modes: $\mu\nu\nu$, $e\nu\nu$, $\pi\nu$, $\pi\,\pi^0\nu$
- PID, ECL cluster cleaning (see next slides)
- $110 < M(\pi^0) < 160$ MeV
- $625 < M(\rho) < 925$ MeV

Require full reconstruction of tag side and only one additional track in the event

Run on MC9 production:
- $80*10^6$ events of $B \to \tau\nu \to$ generic with beam background
- 0.8 ab^{-1} of $B^+B^-/B^0\overline{B}^0$, continuum and $\tau\tau$, with beam background

https://confluence.desy.de/display/BI/Data+Production+MC9

Selection (2)

PID selection

- Use the release-09 working points (99% - 95% efficiencies)

  https://confluence.desy.de/display/BI/Physics+StandardParticles

$e\nu\nu$	- eid>0.750
$\mu\nu\nu$	- muid>0.625 and eid<0.750
$\pi\nu$	- piid>0.429 and muid<0.625 and eid<0.750
$\pi\,\pi^0\nu$	- same as pion + m(ρ) window

Extra clusters and pi0 selection

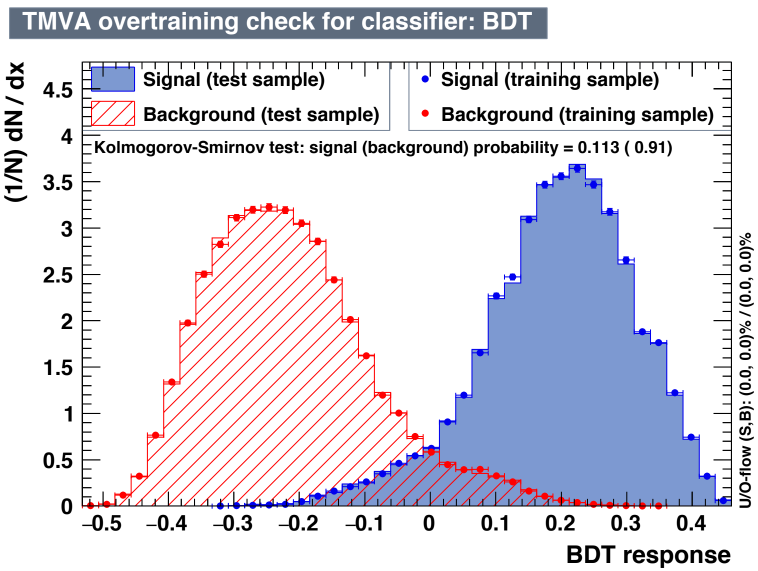
- Two MVA classifiers trained separately for the extra clusters and pi0s

Continuum rejection

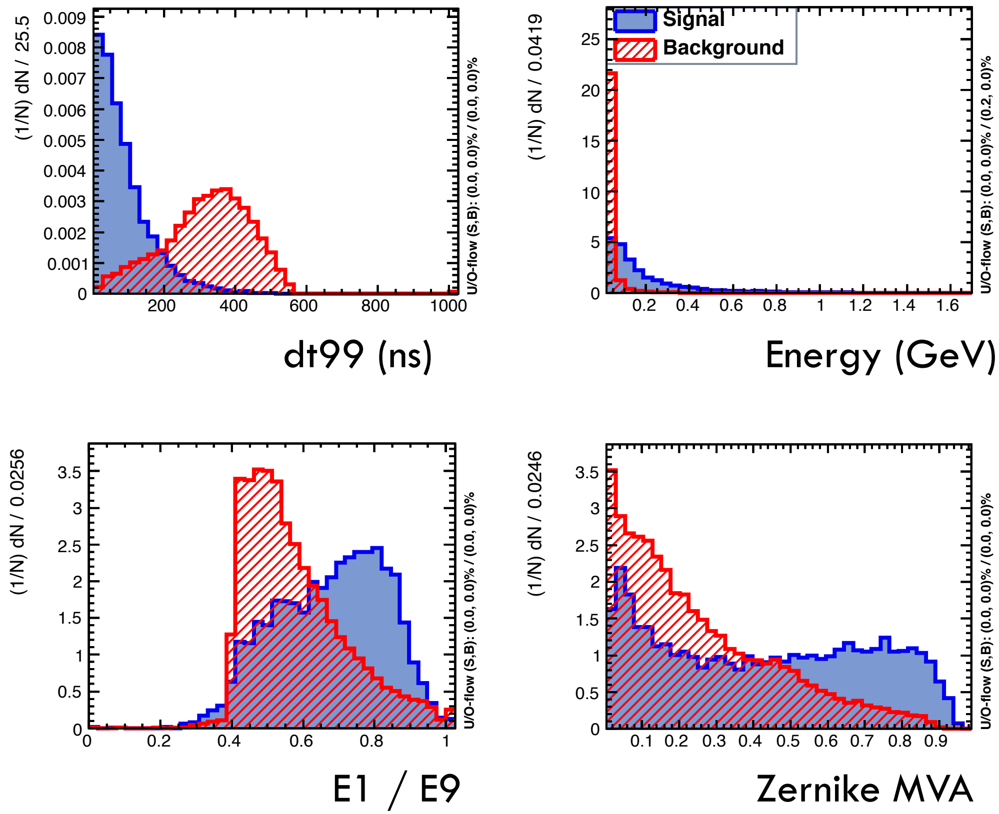
- MVA to separate back-to-back topology from events with spherical symmetry (BB).

Extra clusters MVA

BDT output classifier for signal (physics photons) and background (photons from beam)



Most important variables

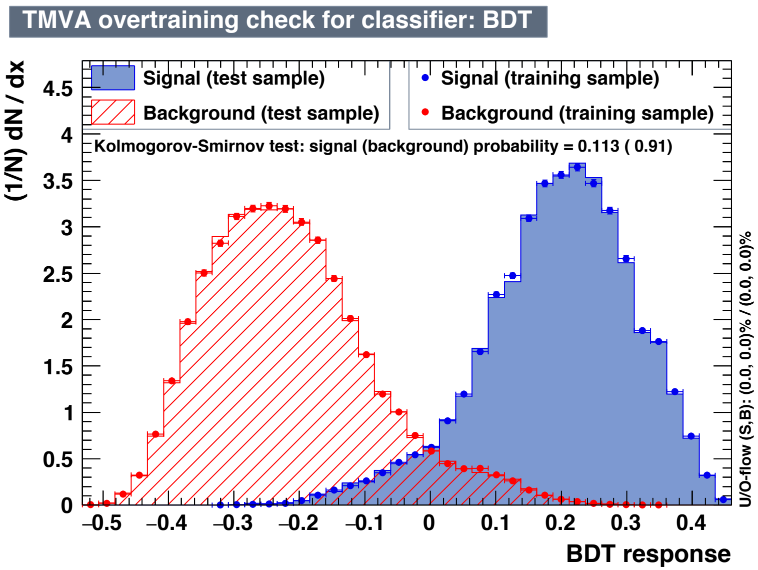


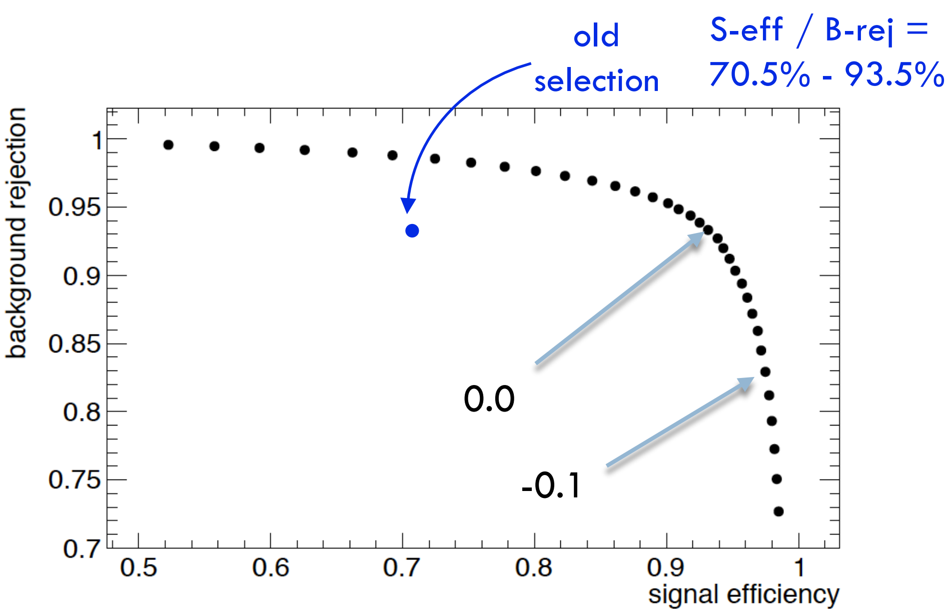
Variables correlation:

- Shower shape variables slightly correlated (E1/E9, Zernike and LAT)
- Some level of correlation between dt99 and the cluster energy

full set of variables and correlations in the backup slides

Extra clusters MVA: ROC curve





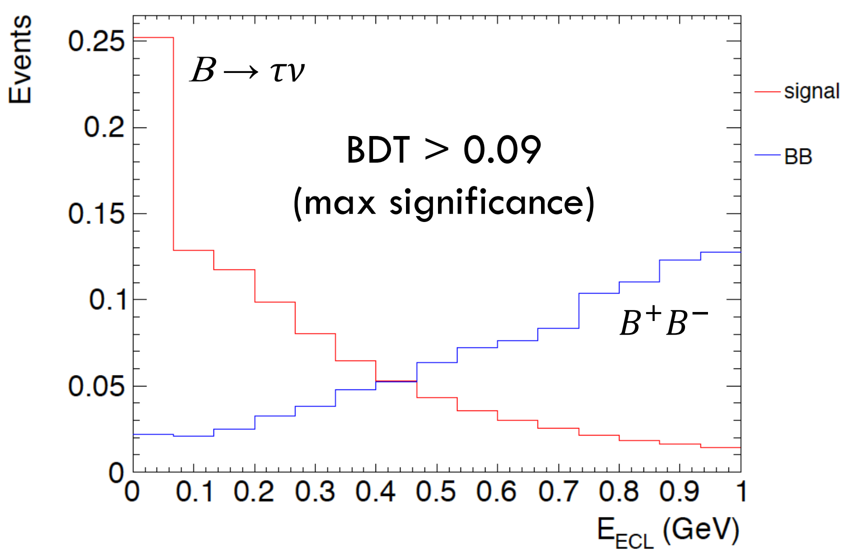
Scan of the BDT from -0.15 to 0.2 with step of 0.01 and plot the signal efficiency vs background rejection

At same signal efficiency level, we have ε_{bkg}=1.5% with respect to 6.5% of old selection –> ~80% more bkg rejected

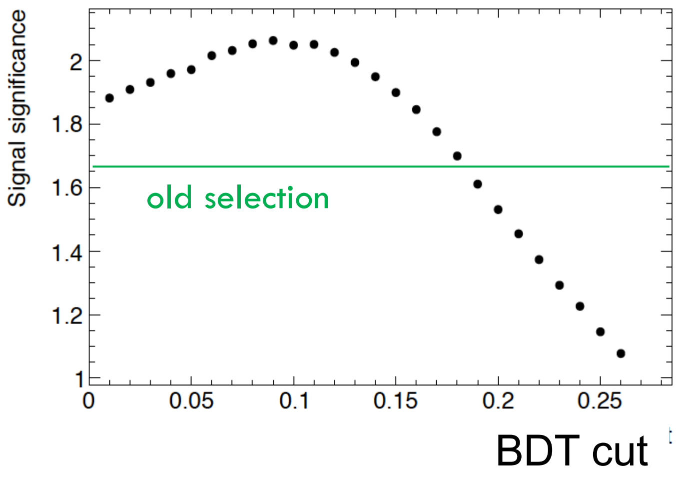
Signal efficiency: N physics photons after BDT cut / Ntot physics photons
Background efficiency: N bkg photons after BDT cut / Ntot bkg photons
Background rejection = 1 - Background efficiency

Extra clusters MVA: performance on E_{extra} distribution

- E_{extra} distribution for tau nu signal and B^+B^- background with $M_{bc} > 5.27$ GeV



Scan of the cut on the BDT



Significance estimated as S/sqrt(S+B) in $E_{extra} < 0.2$ GeV and $M_{bc} > 5.27$ GeV, where **S** is $B \to \tau\nu$ and **B** is BB+continuum+$\tau\tau$ bkg, normalized to 1 ab^{-1}

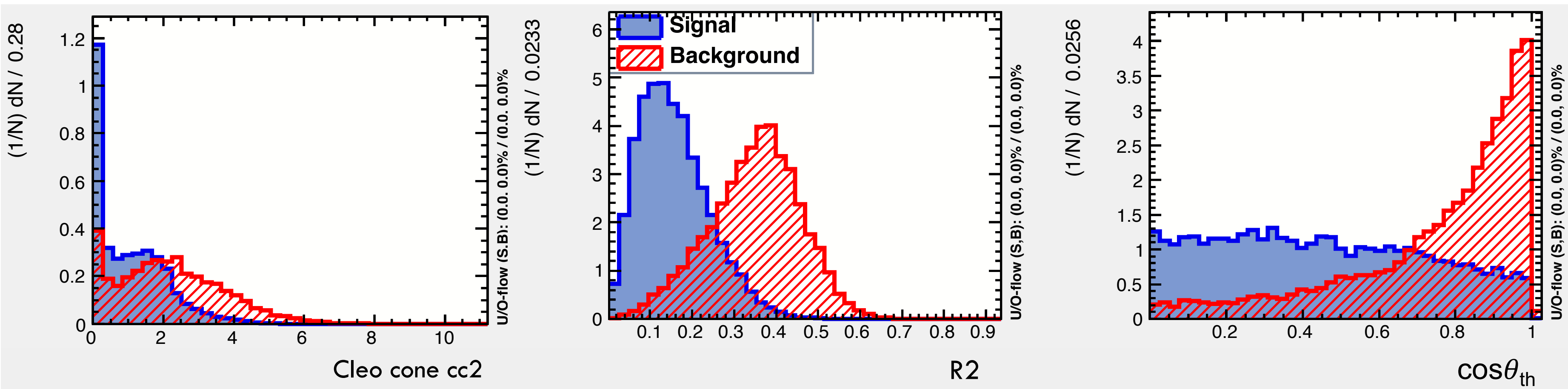
Continuum rejection applied (see next slides)

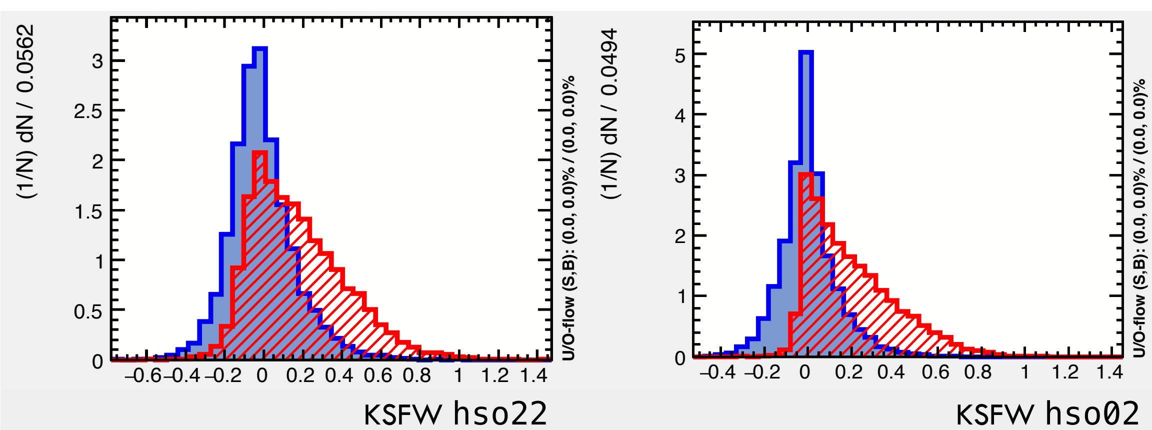
Continuum rejection - BDT

- Input Variables: R2, Cosθ_{th} , Cleo Cones and Kakuno Super Fox-Wolfram (KSFW) moments: 30 variables
  - R2: $R_2 = H_2/H_0$ where $H_l = \sum_{i,j} \frac{|\vec{p}_i||\vec{p}_j|}{W^2} P_l(\cos\vartheta_{ij})$ are the Fox-Wolfram moments
  - Cosθ_{th} : $\left|\cos(\vartheta_{thrust})\right| = \frac{|\vec{p}_B \cdot \hat{T}|}{|\vec{p}_B|}$ where *T* is the thrust axis of the rest of the event
  - Cleo Cones: momentum flow around the B thrust axis in 9 angular bins
  - KSFW: Extension of Fox-Wolfram moments

Detailed explanation here:
https://kds.kek.jp/indico/event/26297/session/1/contribution/11/material/slides/0.pdf

Highest ranked variables

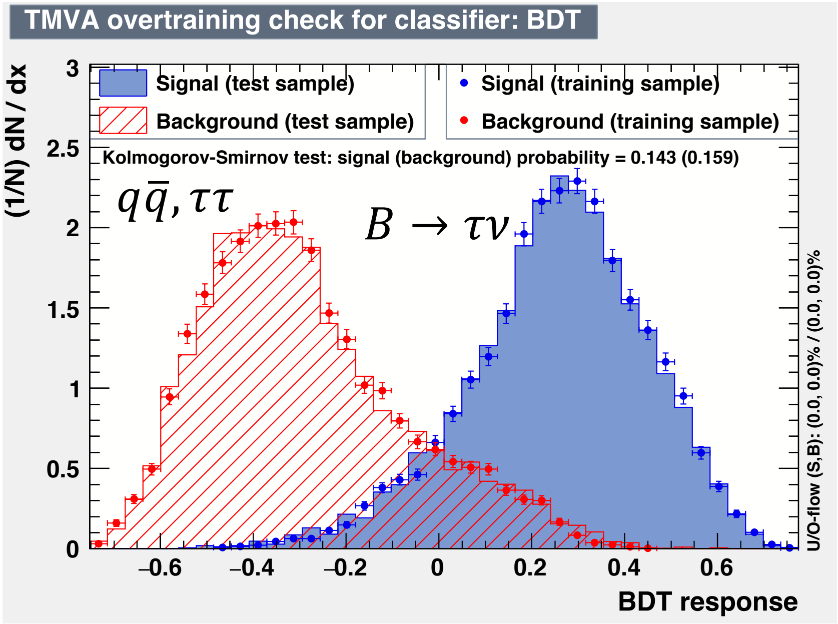


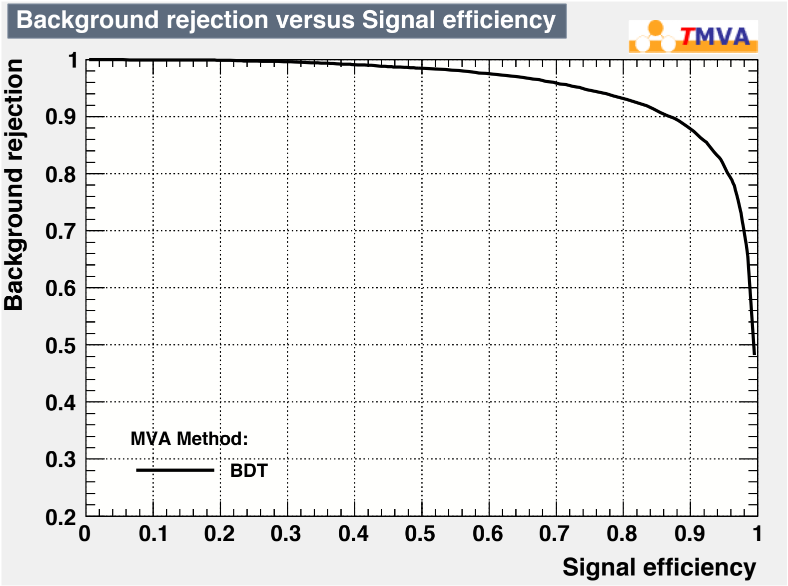


Signal: $B \to \tau\nu$

Background: $q\bar{q}, \tau\tau$

Training results





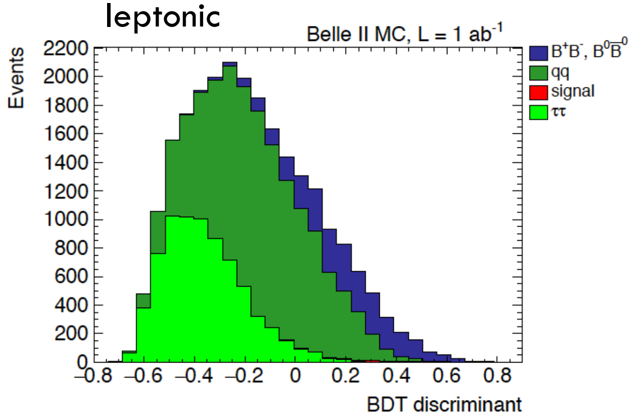
Overtraining under control
Limited statistics for the backgrounds

ROC curve

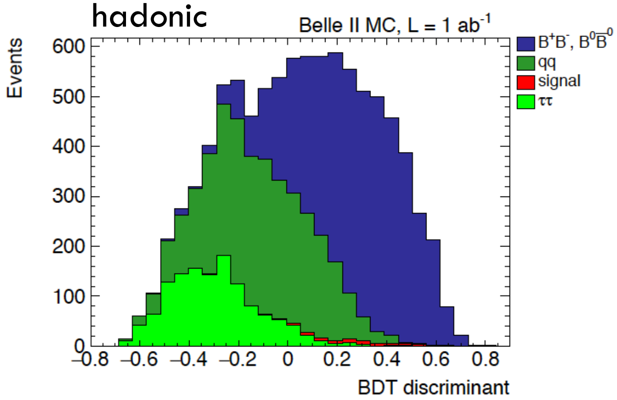
- Good separation power

Continuum rejection: BDT

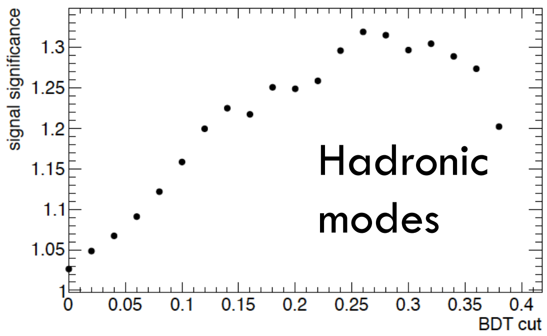
leptonic

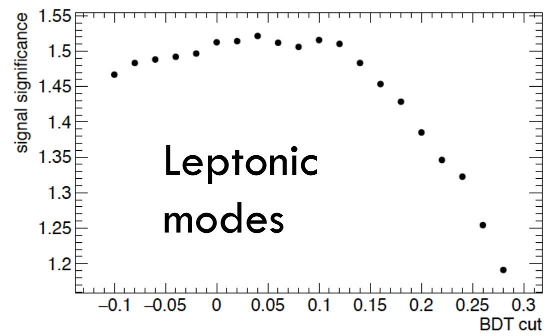


hadronic



The BDT cut is optimized in order to maximize the FOM* in the M_{bc} and E_{extra} signal windows (respectively 5.275-5.29 GeV/c^2 and 0-0.2 GeV)





The continuum background mostly affects the hadronic modes → apply a tighter cut

In all the plots shown here and in the next slides the signal and bkg are normalized to 1 ab^{-1}

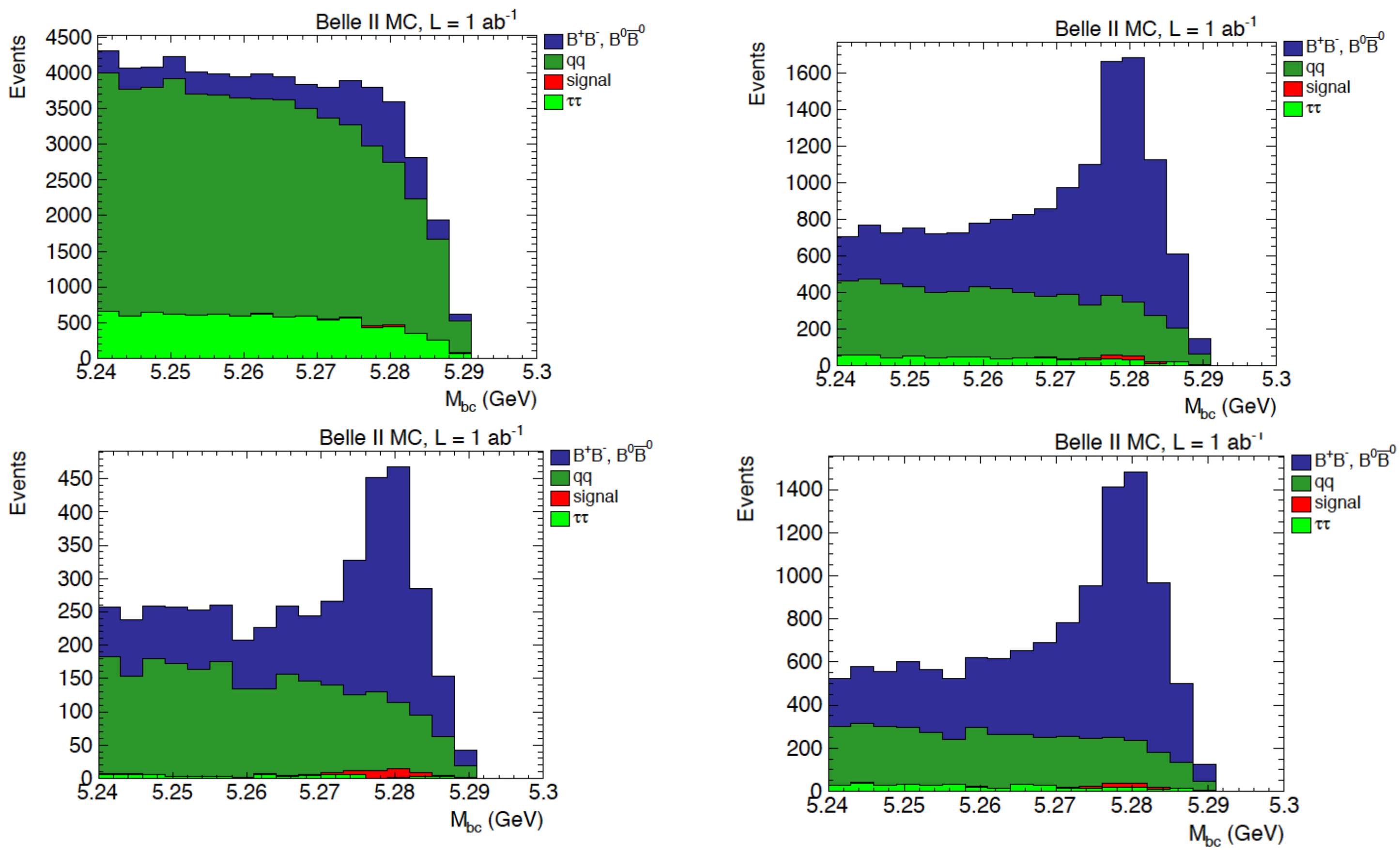
*estimated as S/sqrt(S+B) where **S** is $B \to \tau\nu$ and **B** is BB+$q\bar{q}, \tau\tau$ bkg, normalized to 1 ab^{-1}

M_{bc} distribution

hadronic modes | leptonic modes

Before continuum rejection

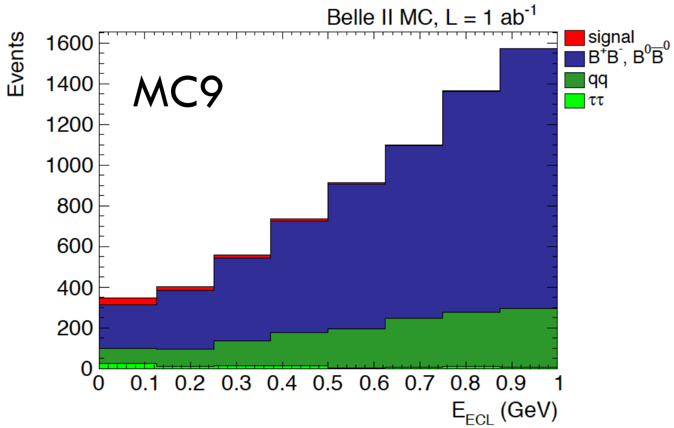
After continuum rejection



$5.27 < M_{bc} < 5.29$ GeV/c^2

E_{extra} and selection efficiency

E_{extra} distribution after selection



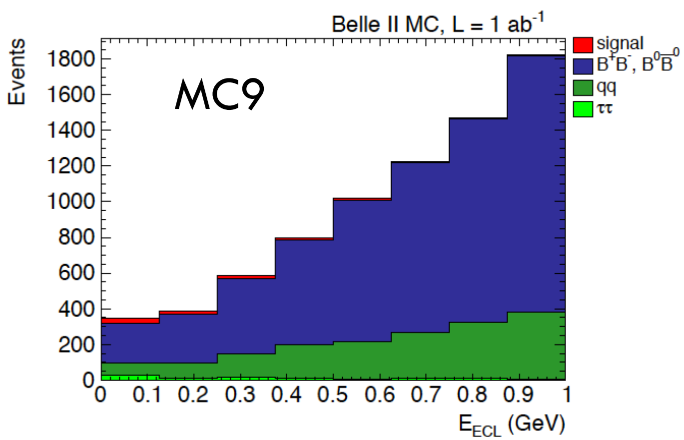
Signal and background event yields in 1 ab^{-1}

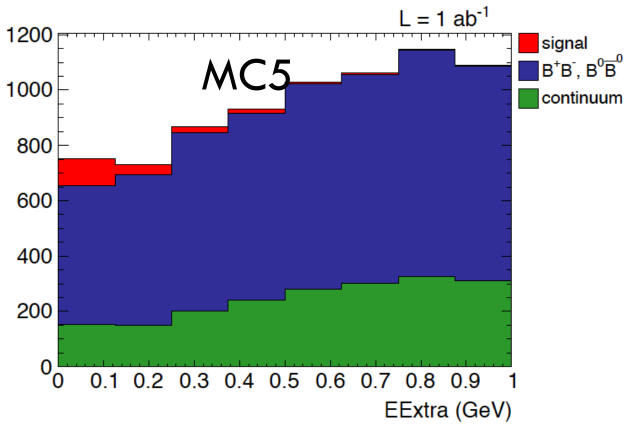
Eextra < 1 GeV
- sig: 97 events
- bkg: 6900 events

Eextra < 0.2 GeV
- sig: 49 events
- bkg: 512 events

E_{extra} < 1 GeV	Babar PRD 88, 031102 (2013)	Belle PRL 110, 131801 (2013)	Belle II (this analysis)
Signal Efficiency (‰)	**0.72**	**1.1**	**1.1**

old Eextra definition



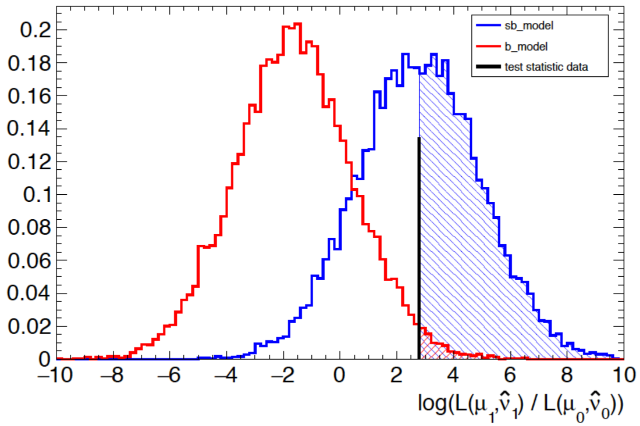


Signal efficiency in MC9 ~ half the MC5 efficiency

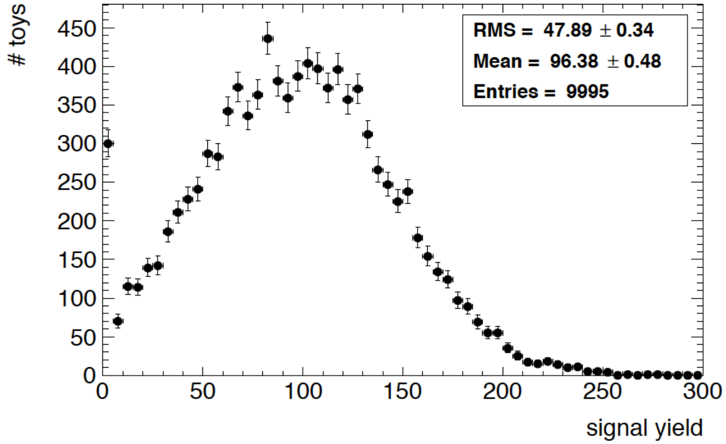
Toy MC study and expected sensitivity

- Perform a 1D fit to the E_{extra} distribution
  - Generate a pseudo-dataset according to the signal + background MC expectations
  - Perform a template maximum likelihood fit to E_{extra} with two components: signal and background pdfs built from the expected MC distributions

- Toy MC with 10000 pseudo-datasets:



Log Likelihood Ratio test statistics



Mean yield: 96 events
Mean uncertainty: 48 events

Significance: 2.10 +/- 0.02 (in MC5 it was ~3.4σ)

SL & L physics group analysis validation

https://confluence.desy.de/display/BI/Data+Production+Validation

- Four modes for analysis validation: $B \to \tau\nu$, $B \to \pi l\nu$, $B \to D^* l\nu$, $B \to D^* \tau\nu$
- The basic idea is set up a simple analysis for each mode and check the effect of the changes from a software release to an other (comparison between different MC campaigns)

$B \to \tau\nu$ MC9 vs MC10

Signal reconstruction efficiency

efficiency	MC9	MC10
TOT	3.5‰	7.1‰
$\mu\nu\nu$ mode	0.7‰	2.3‰
$e\nu\nu$ mode	0.8‰	2.3‰
$\pi\nu$ mode	1.4‰	1.8‰
$\pi\,\pi^0\nu$ mode	0.6‰	0.7‰

efficiency	MC9	MC10
TOT	0.66‰	1.3‰
$\mu\nu\nu$ mode	0.21‰	0.48‰
$e\nu\nu$ mode	0.21‰	0.46‰
$\pi\nu$ mode	0.15‰	0.26‰
$\pi\,\pi^0\nu$ mode	0.09‰	0.12‰

requiring MC matching

Selection steps:

- B-tag reconstruction with FEI
- PID
- Neutral quality criteria

Efficiency increases by factor ~2

B → πlν MC9 vs MC10

Selection steps

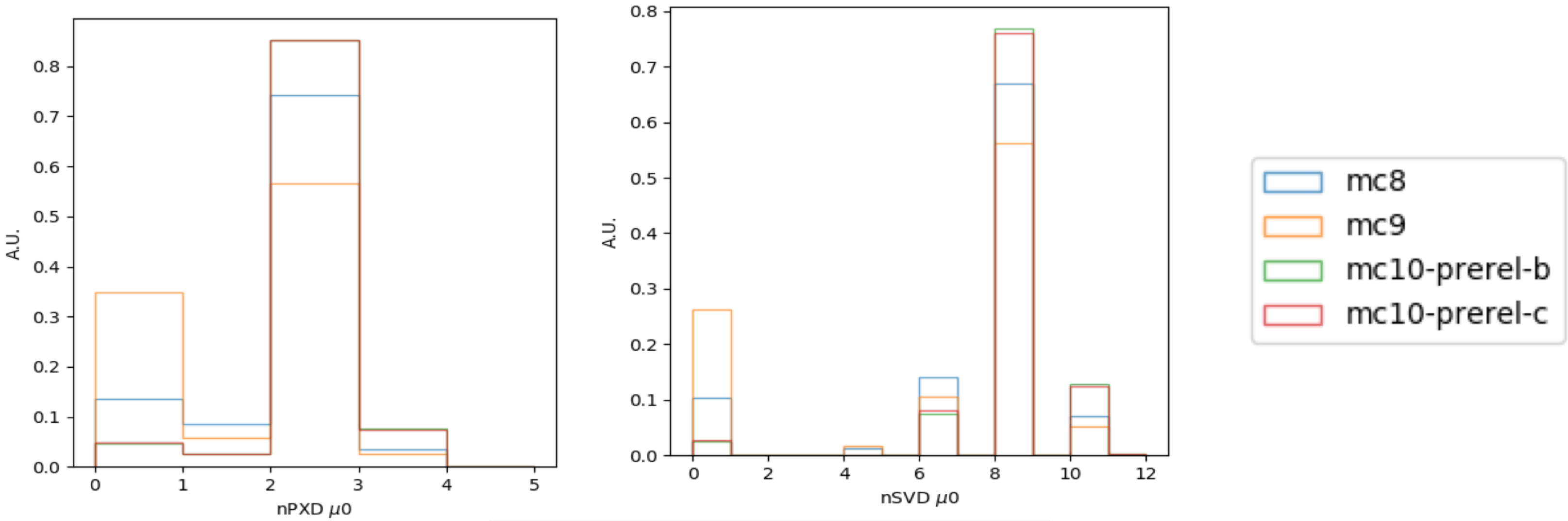
- Reconstruct the signal B ⟶ π l ν, requiring PID > 0.8 and $p^* > 1\ \mathrm{GeV}$
- Define ROE(signal) and remove tracks originating far from the IP and low energetic / out-of-time clusters
- Untagged

Efficiency	MC9	MC10
reco (iCand==0)*	61.9%	65.0%
truth	40.2%	44.2%

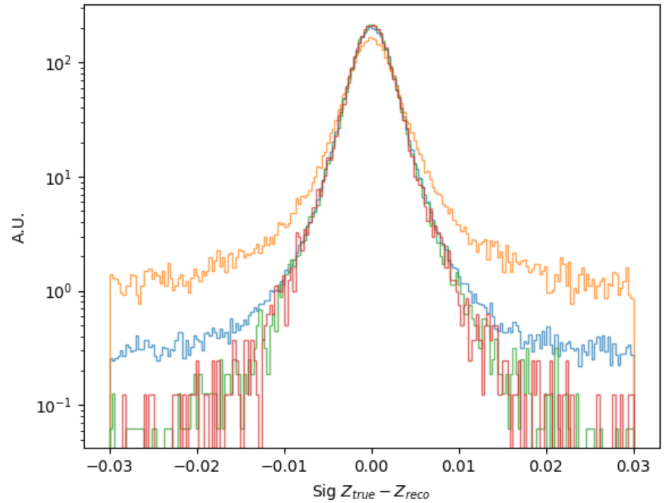
* number of events with reconstructed candidates

MC9 vs MC10

Number of PXD and SVD hits associated with mu track from $J/\psi \rightarrow \mu\mu$ ($B \rightarrow J/\psi\ K_s$)



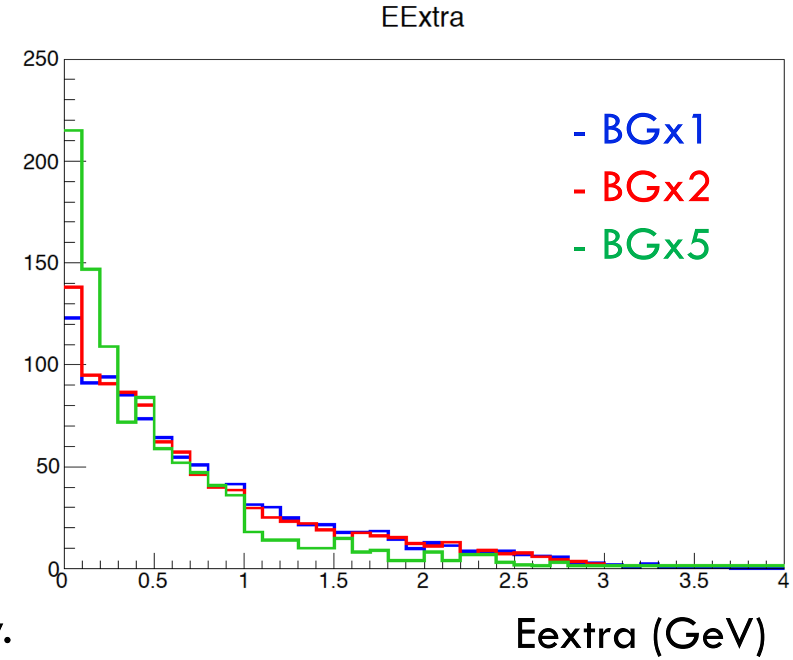
Resolution on decay vertexes



Background impact

$B \rightarrow \tau\nu$

efficiency	BGx1	BGx2	BGx5
TOT	1.3‰	1.0‰	0.06‰
$\mu\nu\nu$ mode	0.48‰	0.34‰	0.02‰
$e\nu\nu$ mode	0.46‰	0.40‰	0.03‰
$\pi\nu$ mode	0.26‰	0.19‰	0.01‰
$\pi\,\pi^0\nu$ mode	0.12‰	0.06‰	0.002‰



From bgx1 to bgx2: expected drop of efficiency.
Bgx5 has a catastrophic impact on the measurement

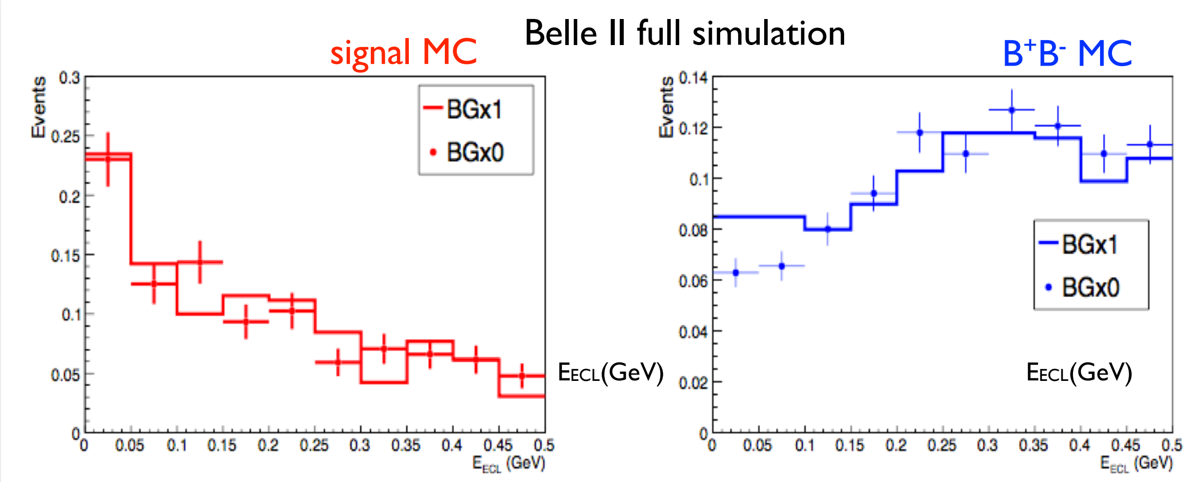
$B \to K^* \nu \bar{\nu}$ status

Robustness against machine background with MC5

- Nominal machine bkg (BGx1) and machine bkg-free (BGx0) simulated samples analysed
- Negligible impact of machine background both in terms of variables shape and signal significance

1 ab^{-1} equivalent statistics

	"BGx0"	"BGx1"
N_{bkg}	6415 ± 80	3678 ± 61
$\varepsilon\ (10^{-4})$	10.3 ± 0.3	5.38 ± 0.23
$N_{sig}/\sqrt{N_{bkg}}$	0.16	0.15
UL (10^{-4})	2.6	3.8



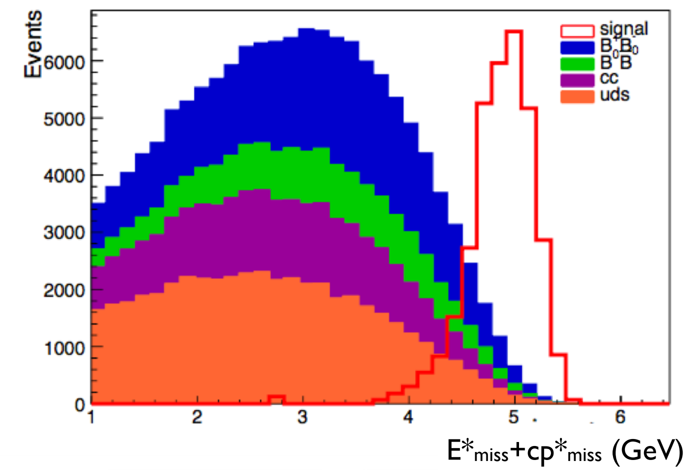
- Detector performances and reconstruction proves to be robust against machine background… considering MC5 machine background. In new PHASE III simulation, machine background increased of a factor of 3, studies will be repeated with latest MC production, including new extra neutral and neutral pion selection

Perspectives with full Belle II statistics

- Extrapolation on full Belle II statistics on Belle HAD and SL analyses, assuming two times better B_{tag} reconstruction efficiency:
  - observation with about 18 ab^{-1}
  - precision on the branching fraction at 50 ab^{-1}:

	stat only	total
$B^+ \to K^+ \nu\nu$	9,5%	10,7%
$B^+ \to K^{*+} \nu\nu$	7,9%	9,3%
$B^+ \to K^{*0} \nu\nu$	8,2%	9,6%

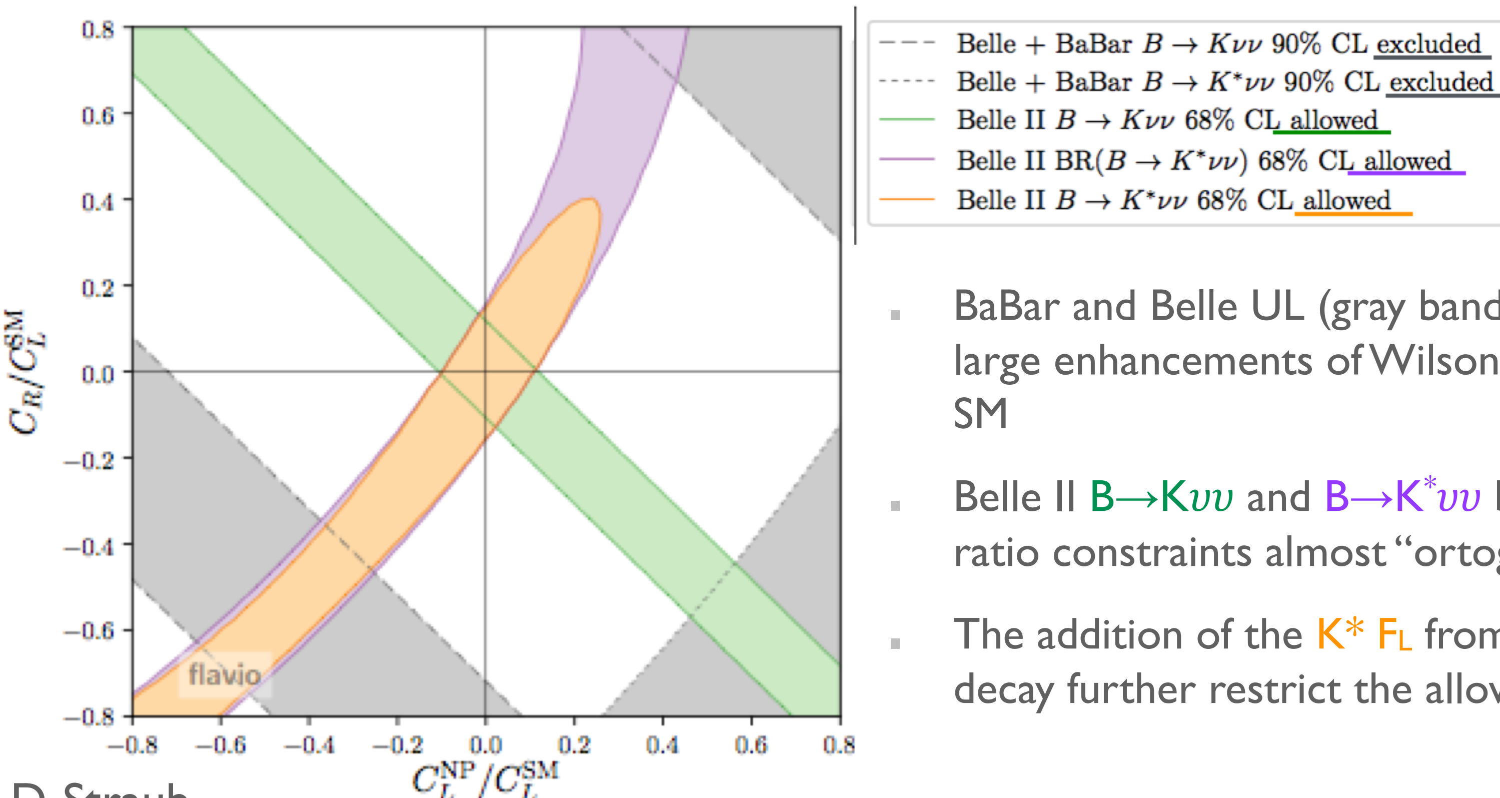
Belle II full simulation with machine background



- Fraction of longitudinally polarized K* may be measured, ~20% precision with full statistics
- Next steps: update the analysis to most recent MC campaign using HAD FEI skimmed sample, evaluate machine background impact, improve analysis strategy (continuum suppression, fit for yield extraction)

Constraints on NP models

- Constraints on (real and neutrino-flavour-independent) Wilson coefficients C_L^{NP} and C_R^{NP} normalised to SM C_L, assuming SM central values and sensitivities from previous page



D. Straub

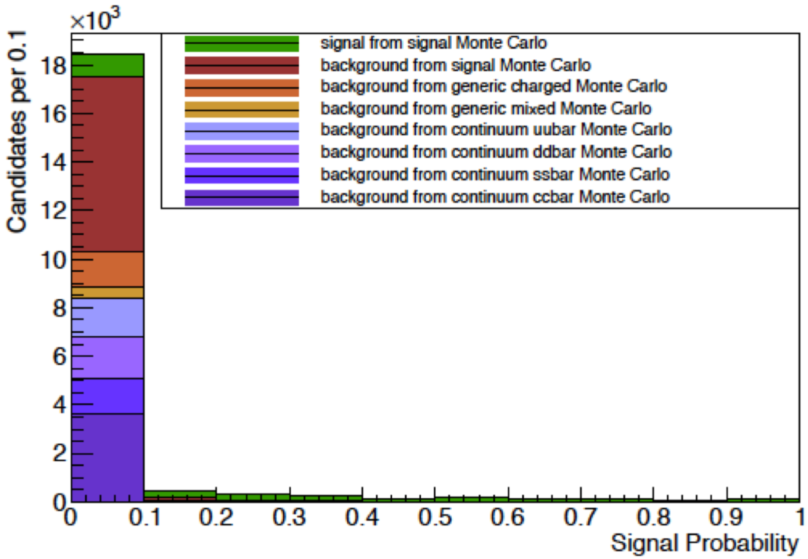
- BaBar and Belle UL (gray bands) ruled out large enhancements of Wilson coeffs. wrt SM
- Belle II $B \to K\upsilon\upsilon$ and $B \to K^*\upsilon\upsilon$ branching ratio constraints almost "ortoghonal"
- The addition of the K^* F_L from $B \to K^*\upsilon\upsilon$ decay further restrict the allowed region

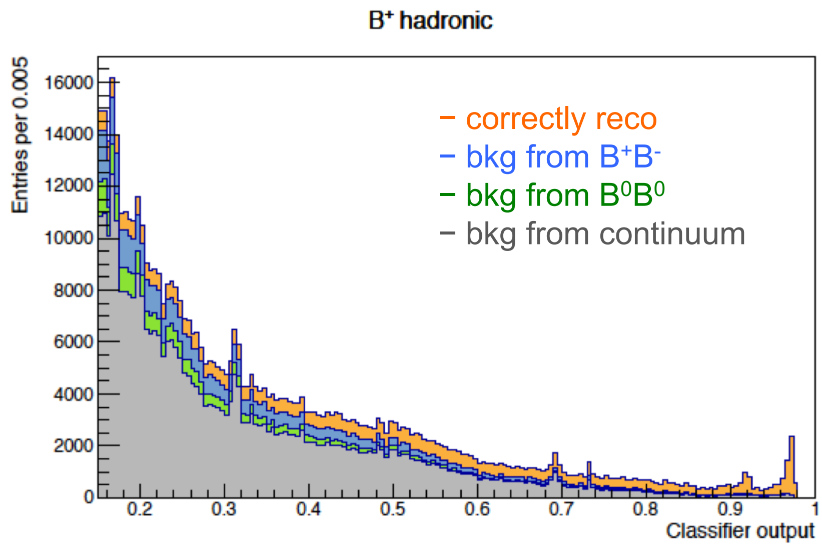
Conclusions

- Studies performed on MC9 ($B \to \tau\nu$) and MC5 ($B \to K^{*}\nu\bar{\nu}$)
- b-tag reconstructed hadronically with FEI, photons cleaning (BDT for extra clusters and pi0), continuum suppression (BDT)
- Extrapolation of precision on the BR with increasing integrated luminosity
- $B \to \tau\nu$: efficiency in signal region (Eextra<0.2GeV) is ~half w.r.t. previous version of the analysis on MC5
- Analyses validation studies on MC10 show that the efficiency is ~ doubled w.r.t. MC9 –> it may be due to the worst vertexing in MC9 –> low FEI reconstruction efficiency.

Backup

Full Event Interpretation (FEI) performances





from Christian Pulvermacher PhD thesis

Total reconstruction efficiency compared with Belle I

Belle II

B^+ (hadronic)	0.78 %		B^+ (semileptonic)	1.05 %
B^0 (hadronic)	0.59 %		B^0 (semileptonic)	1.17 %

Belle I

B^+ (hadronic)	0.39 %		B^+ (semileptonic)	0.80 %
B^0 (hadronic)	0.28 %		B^0 (semileptonic)	0.86 %

Belle paper

Belle paper, hadronic tag,
PRL 110, 131801 (2013)

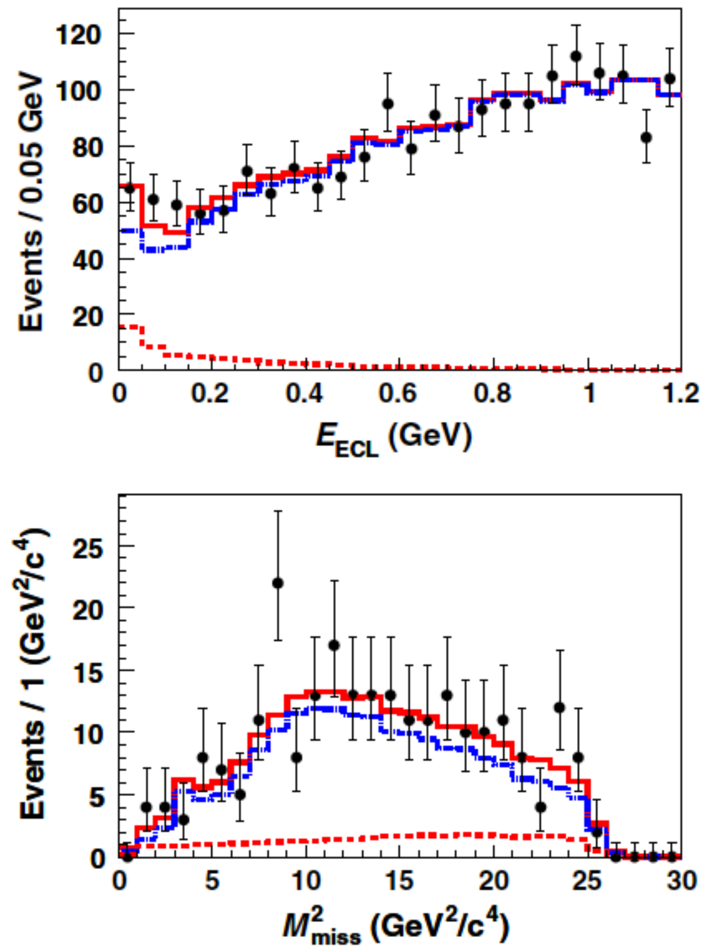
Entire Belle data sample ~ 700 fb^{-1}

TABLE I. Results of the fit for $B^- \to \tau^- \bar{\nu}_\tau$ yields ($N_{\rm sig}$), detection efficiencies (ϵ), and branching fractions ($\mathcal{B}$). The efficiencies include the branching fractions of the τ^- decay modes. The errors for $N_{\rm sig}$ and $\mathcal{B}$ are statistical only.

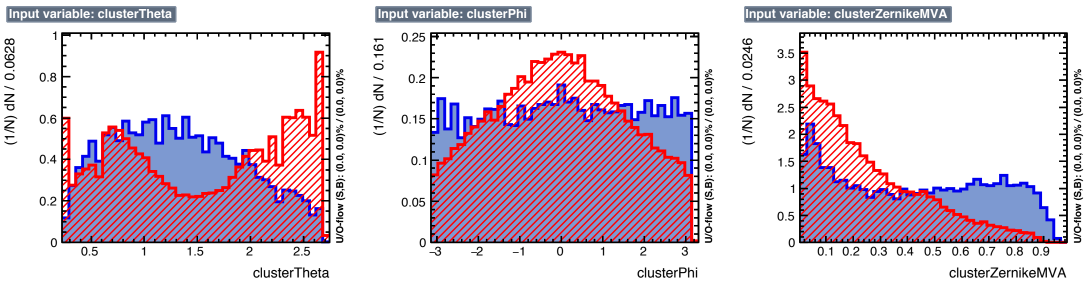
Submode	$N_{\rm sig}$	ϵ (10^{-4})	$\mathcal{B}$ (10^{-4})
$\tau^- \to e^- \bar{\nu}_e \nu_\tau$	16^{+11}_{-9}	3.0	$0.68^{+0.49}_{-0.41}$
$\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$	26^{+15}_{-14}	3.1	$1.06^{+0.63}_{-0.58}$
$\tau^- \to \pi^- \nu_\tau$	8^{+10}_{-8}	1.8	$0.57^{+0.70}_{-0.59}$
$\tau^- \to \pi^- \pi^0 \nu_\tau$	14^{+19}_{-16}	3.4	$0.52^{+0.72}_{-0.62}$
Combined	62^{+23}_{-22}	11.2	$0.72^{+0.27}_{-0.25}$

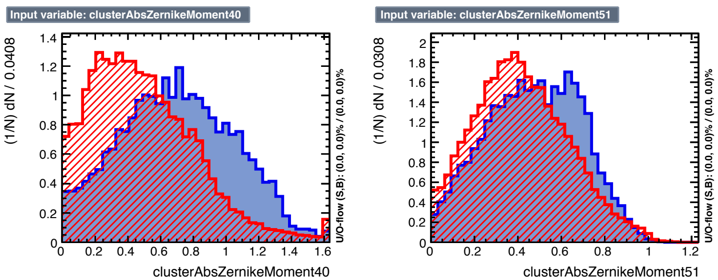
$$\mathcal{B}(B^- \to \tau^- \bar{\nu}_\tau) = [0.72^{+0.27}_{-0.25}(\text{stat}) \pm 0.11(\text{syst})] \times 10^{-4}$$

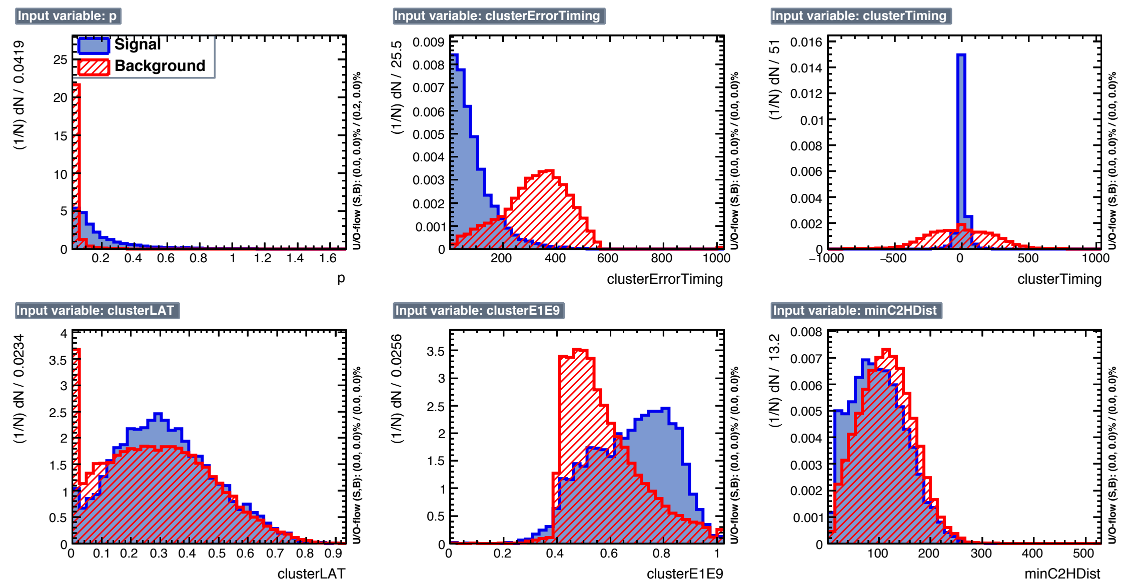
Significance: 3.0 σ



Extra clusters MVA: input variables





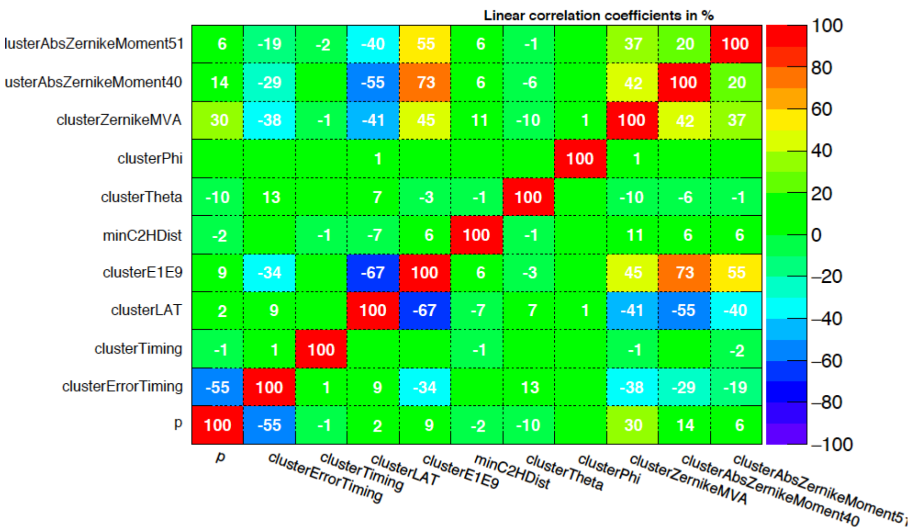


cluster LAT:

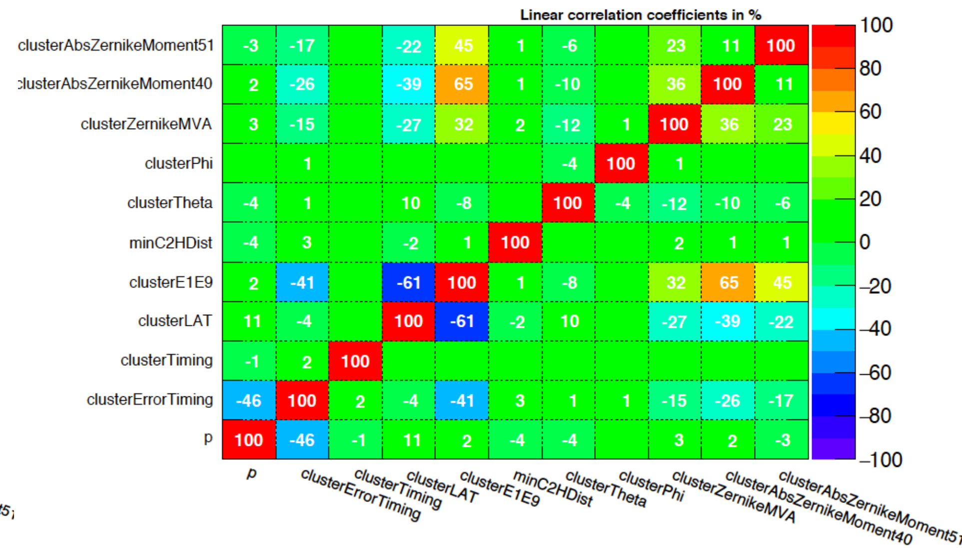
$$S = \frac{\sum_{i=3}^{n} w_i E_i r_i^2}{\sum_{i=3}^{n} w_i E_i + w_0 E_0 r_0^2 + w_1 E_1 r_0^2}$$

Extra clusters MVA: variables correlation

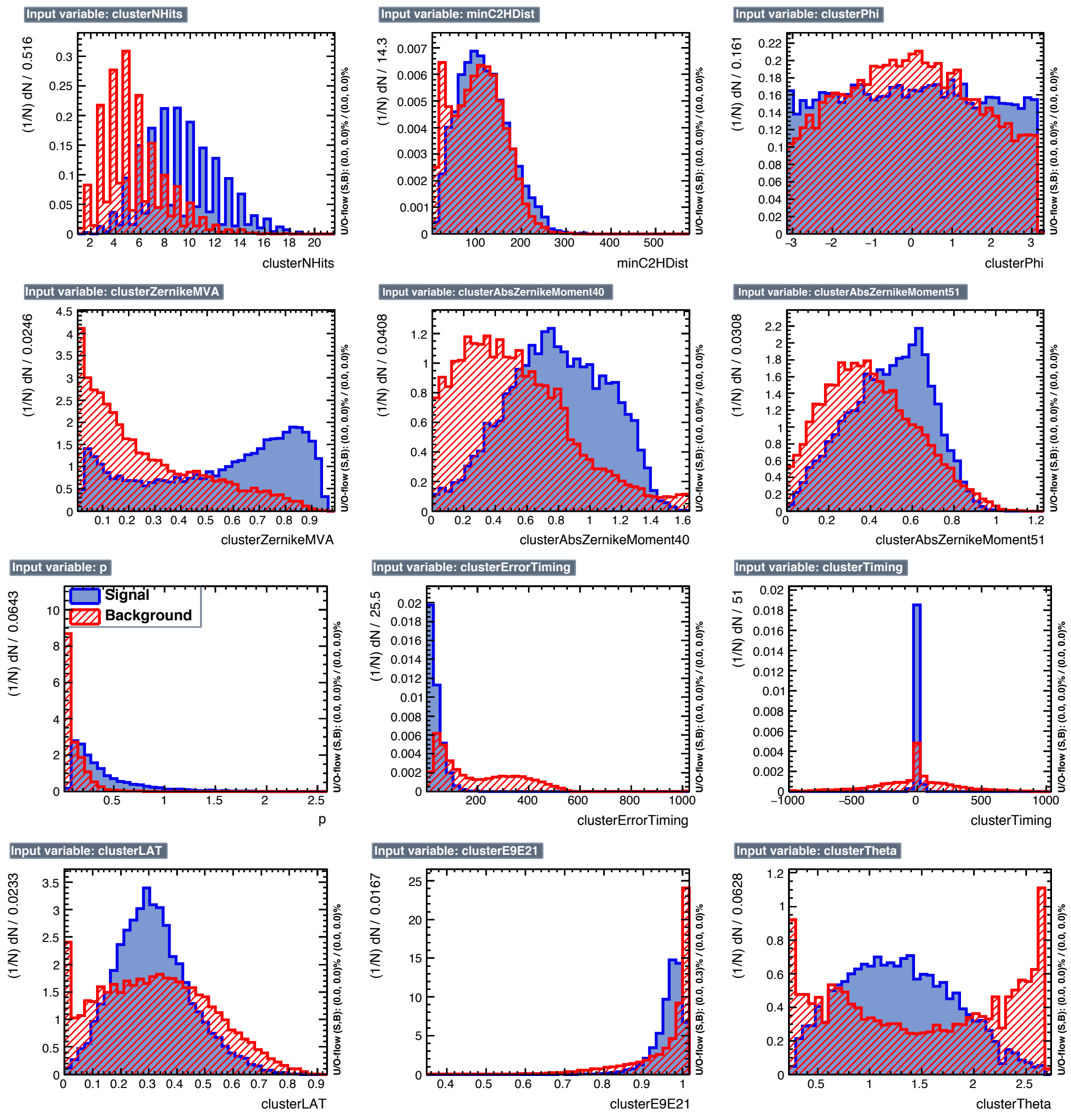
Correlation Matrix (signal)



Correlation Matrix (background)



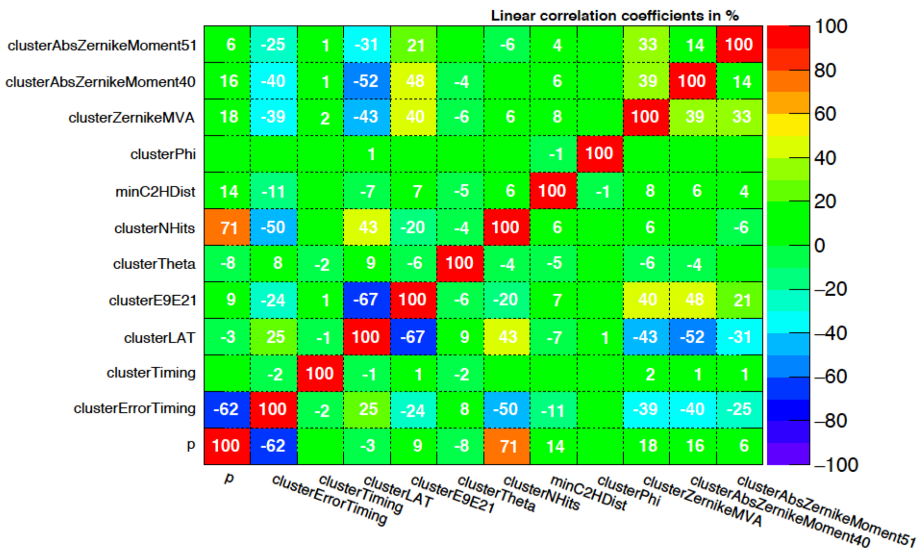
Pi0 MVA



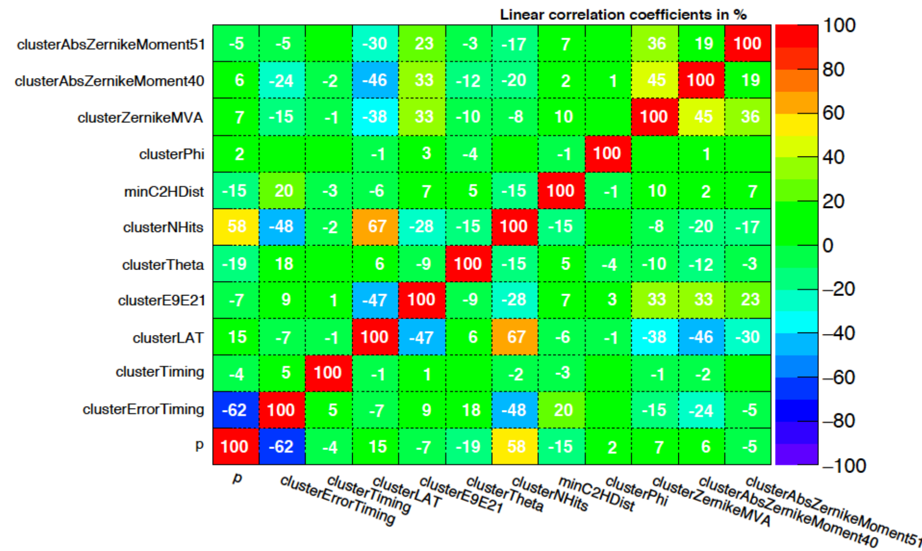
Input variable: clusterNHits
clusterNHits
Input variable: minC2HDist
minC2HDist
Input variable: clusterPhi
clusterPhi
Input variable: clusterZernikeMVA
clusterZernikeMVA
Input variable: clusterAbsZernikeMoment40
clusterAbsZernikeMoment40
Input variable: clusterAbsZernikeMoment51
clusterAbsZernikeMoment51
Input variable: p
Signal
Background
p
Input variable: clusterErrorTiming
clusterErrorTiming
Input variable: clusterTiming
clusterTiming
Input variable: clusterLAT
clusterLAT
Input variable: clusterE9E21
clusterE9E21
Input variable: clusterTheta
clusterTheta

Pi0 MVA: variables correlation

Correlation Matrix (signal)



Correlation Matrix (background)



MC comparisons

B –> τν validation: loose cut on signal prob, no continuum suppression, loose cuts on other variables

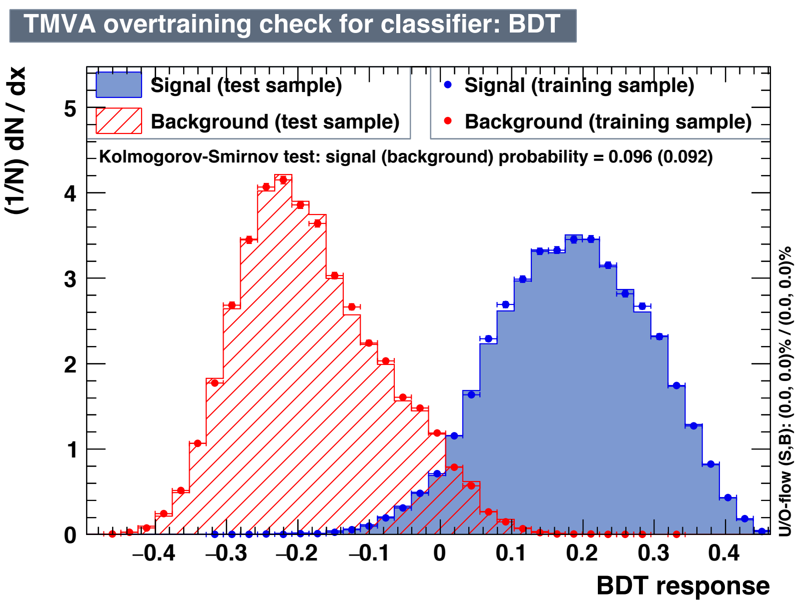
efficiency	MC9	MC10
TOT	3.5‰	7.1‰
$\mu\nu\nu$ mode	0.7‰	2.3‰
$e\nu\nu$ mode	0.8‰	2.3‰
$\pi\nu$ mode	1.4‰	1.8‰
$\pi\,\pi^0\nu$ mode	0.6‰	0.7‰

#events in (0 - 1.0) GeV	Lep channels	Had channels	Total
MC5	**126**	**62**	**188**
MC9	**57**	**40**	**97**

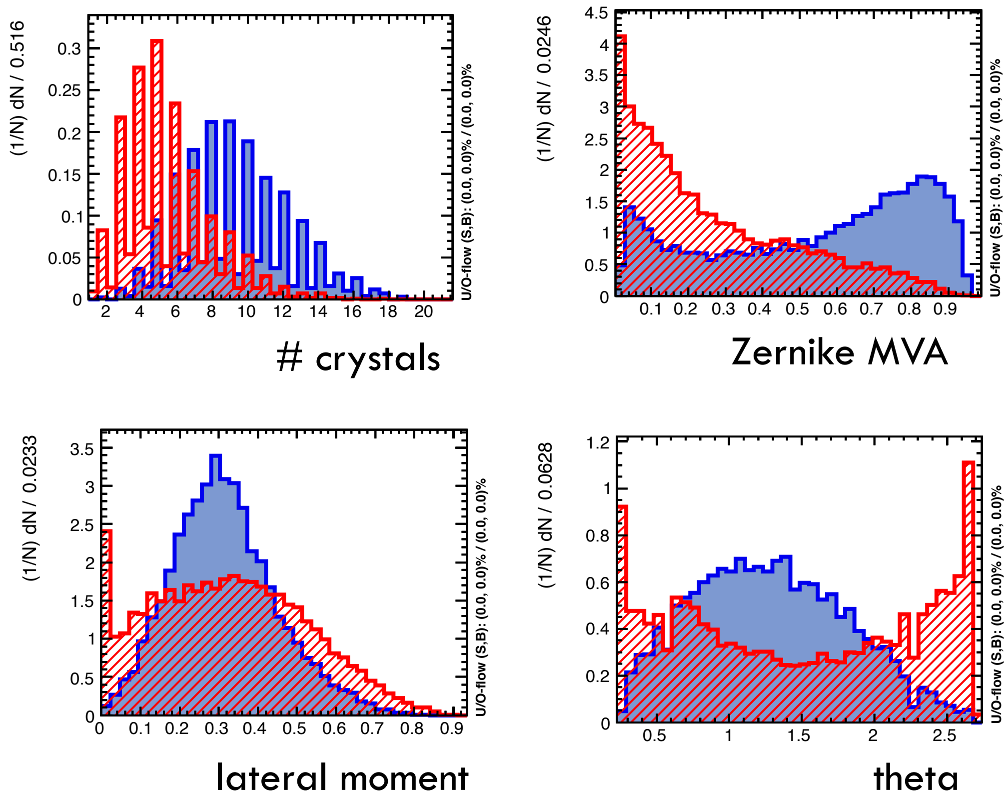
#events in (0 - 0.2) GeV	Lep channels	Had channels	Total
MC5	**88**	**35**	**123**
MC9	**31**	**17**	**48**

π^0 MVA

BDT output classifier for signal (physics photons) and background (photons from beam)



Most important variables

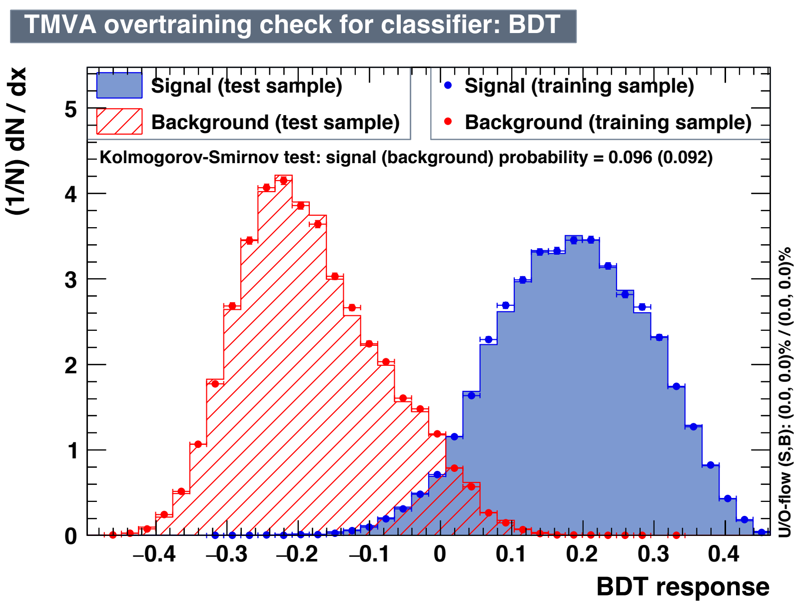


Variables correlation:

- Shower shape variables slightly correlated (E1/E9, Zernike and LAT)
- Number of crystals highly correlated with energy and timing (~60%)

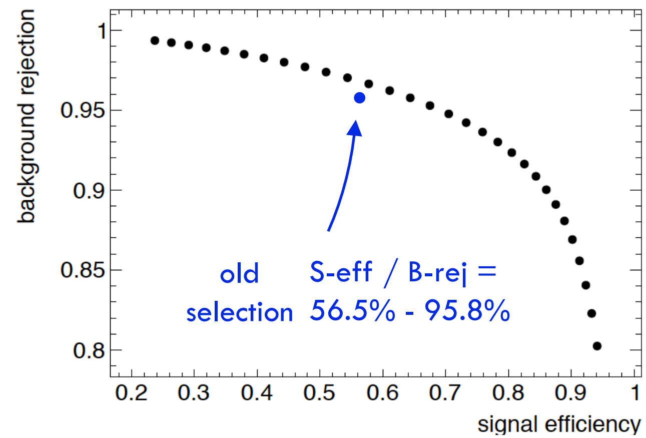
full set of variables and their correlation in the backup slides

π^0 MVA: performances

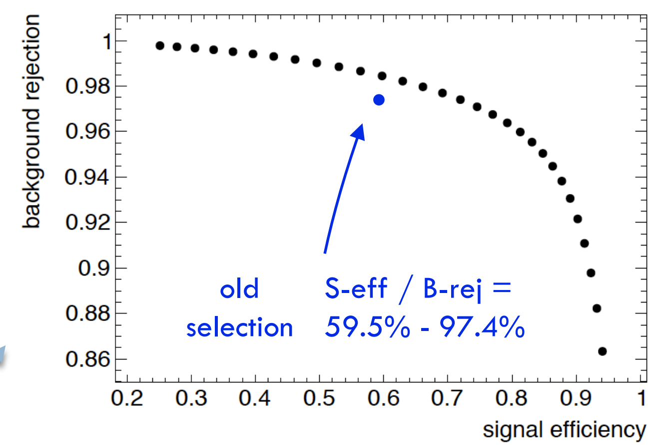


Scan of the BDT from -0.15 to 0.15 with step of 0.01, and plot the signal efficiency vs background rejection

B^+B^- events



$B \rightarrow \tau\nu$ events

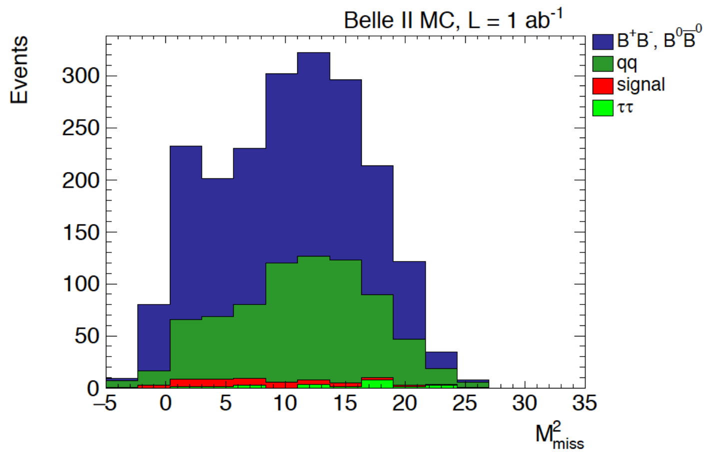


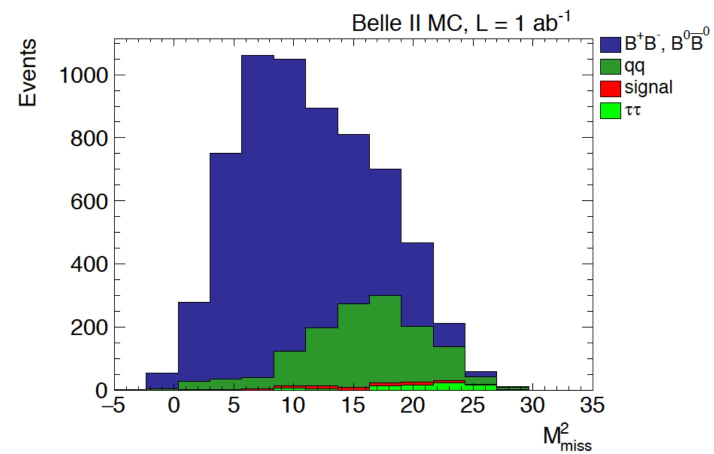
At same signal efficiency level, we have ε_{bkg}=1.5% with respect to 2.6% of old selection –> 40% more bkg rejected

M2miss and signal side momentum

hadronic modes / leptonic modes

M2miss





Signal side momentum