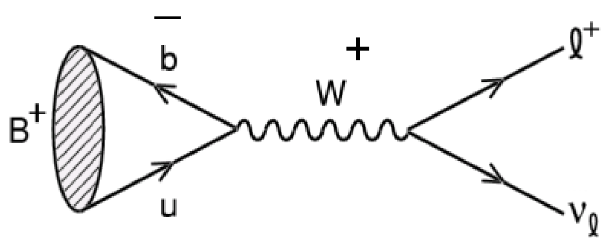
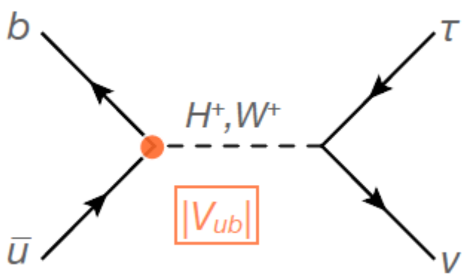
$B \to \tau\nu$ report for B2TiP

Mario Merola, Elisa Manoni, Guglielmo De Nardo

Belle II Italia

Trieste, 4/5/17





Outline

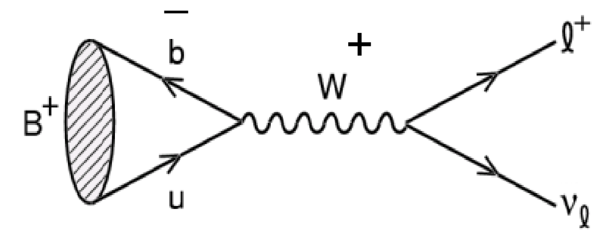
new w.r.t. last meeting

- Theory introduction and recent results overview
- B-tag reconstruction and signal selection
- *BR measurement projections for different lumi scenarios*
- *Comparisons (Belle/BelleII and BGx0/BGx1)*
- Conclusions and future plans

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

- Helicity suppressed

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

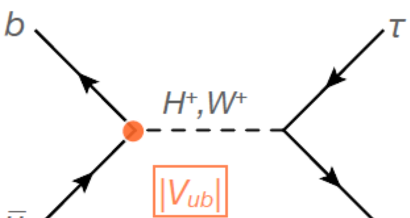


$\tau:\mu:e \to 1 : 10^{-3} : 10^{-7}$

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

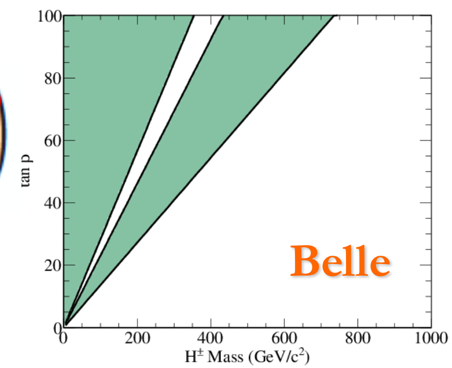
http://ckmfitter.in2p3.fr/

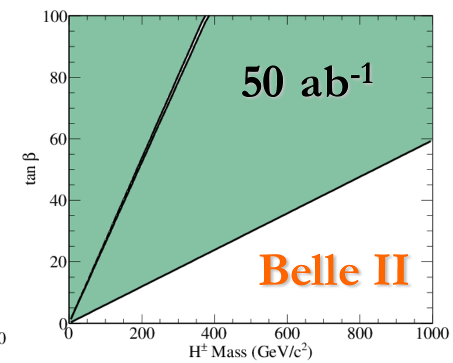
Higgs doublet models predict interference with SM decay with a modification of the branching ratio [PhysRevD.86.054014]



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

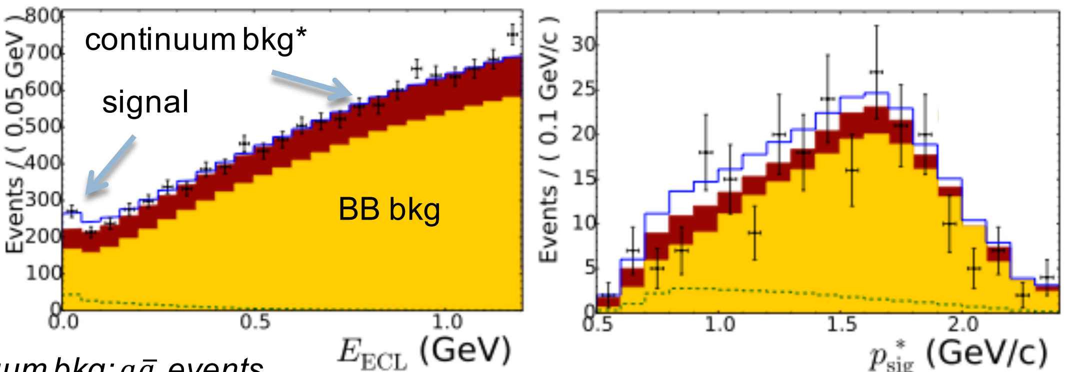
ratio of the two Higgs vacuum expectation values





Recent results on $B \to \tau\nu$

- First evidence at Belle (2006) and Babar (2012)
- Most recent measurement (Belle – 2015, using semileptonic tag):
  - use of multivariate techniques (neural network) to reconstruct the tag side
  - the signal side is reconstructed in four modes: $\tau \to \mu\nu\nu, e\nu\nu, \pi\nu, \rho\nu$
  - the signal is extracted through a two-dimensional maximum likelihood fit to the E_{ECL} and p^*_{sig} distributions



*continuum bkg: $q\bar{q}$ events

- E_{ECL} (later on called E_{extra}) is the sum of the energies of clusters in the ECL not associated to reconstructed B mesons
- p^*_{sig} is the momentum of the signal side particle in the CM

$$\mathcal{B} = [0.91 \pm 0.19(\text{stat.}) \pm 0.11(\text{syst.})] \times 10^{-4}$$ (evidence at ~4.6 σ level)

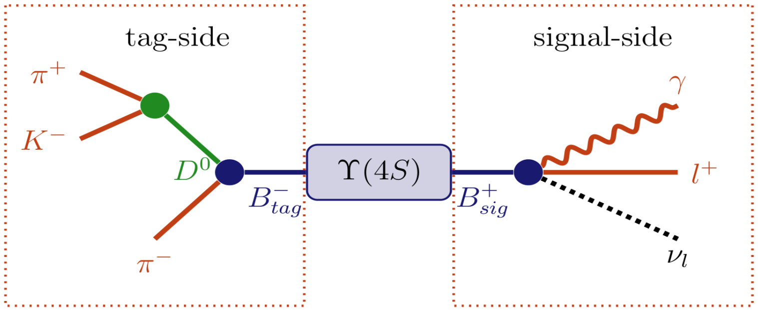
http://arxiv.org/abs/1503.05613v2

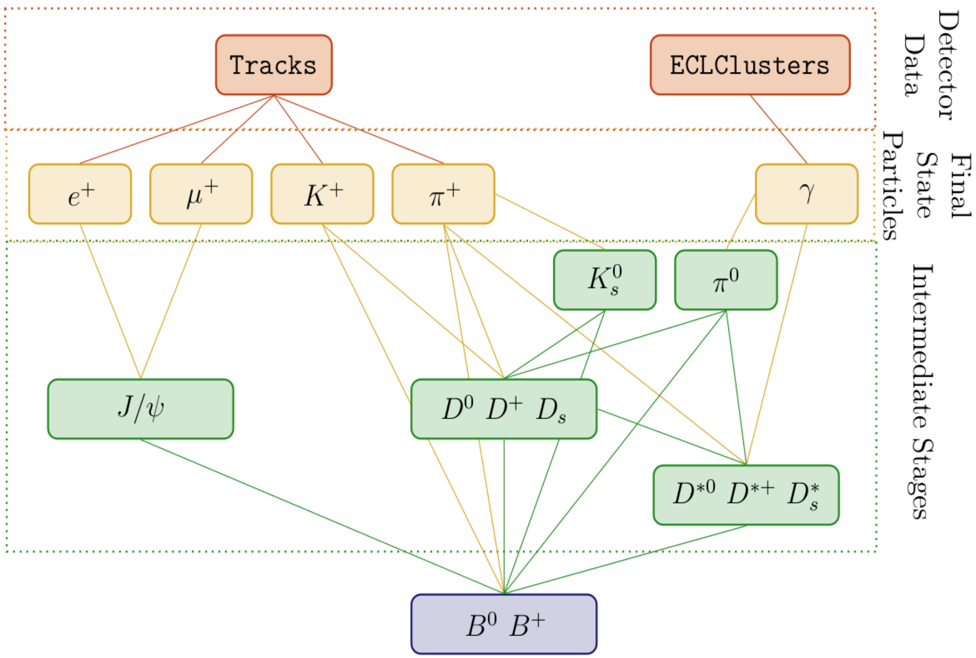
Tag side reconstruction: Full Event Interpretation (FEI)

- 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 performed on 100*10⁶ B^+B^- / $B^0\bar{B}^0$ events of MC5 with beam background (for training on MC8 see my presentation on the grid usage)





*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) 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

$B \rightarrow \tau\nu$

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

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

Run on MC5 production:
- $100*10^6$ events of $B \rightarrow \tau\nu \rightarrow$ generic with beam background
- 1 ab^{-1} of $B^+B^-/B^0\overline{B}^0$ and continuum with beam background

https://confluence.desy.de/display/BI/Computing+MC5Release4Physics

Selection (2)

PID selection

- Likelihood function based on E/p and dE/dx
- Cut on the LR = L(particle) / (L(e) + L(mu) + L(pi))

Photon selection

- Cluster cleaning (to reject photons from beam background) with cuts on photon energy, cluster timing and E9/E25 (separately in forward, barrel and backward detector regions).

Continuum rejection

- MVA with boosted decision trees to separate back-to-back topology from events with spherical symmetry (BB). See backup for details

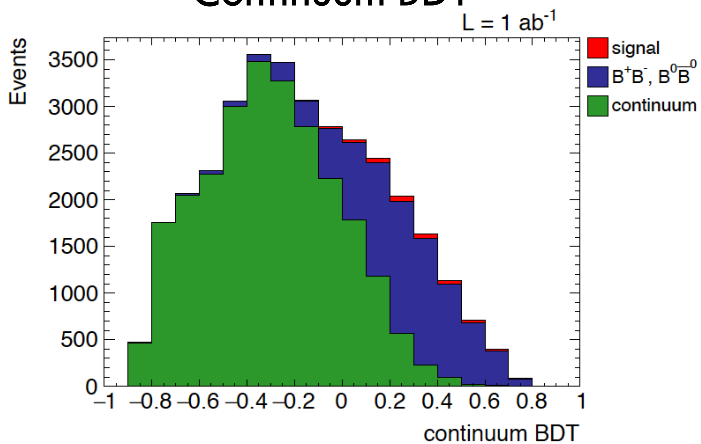
Detailed talk in WG1 meeting:
https://kds.kek.jp/indico/event/21392/contribution/0/material/slides/0.pdf

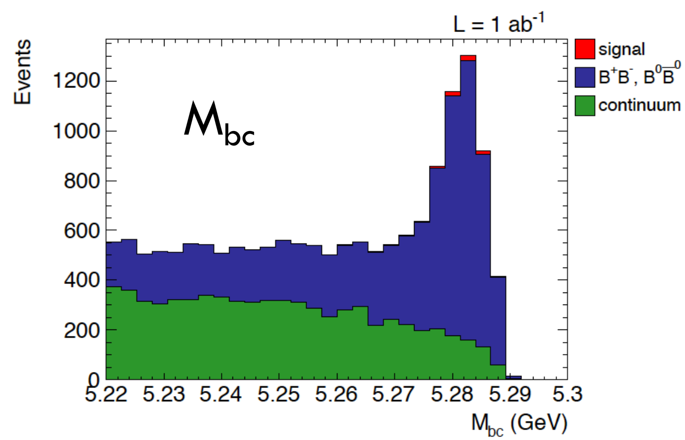
More details at the Twiki page:
https://confluence.desy.de/display/BI/Physics+Pi0Reco

To be optimized with new photons and pi0 lists available in release 08 and with the new PID recommendations:
https://confluence.desy.de/display/BI/Physics+StandardParticles
(see my talk on photons and pi0s ID this afternoon)

Background rejection

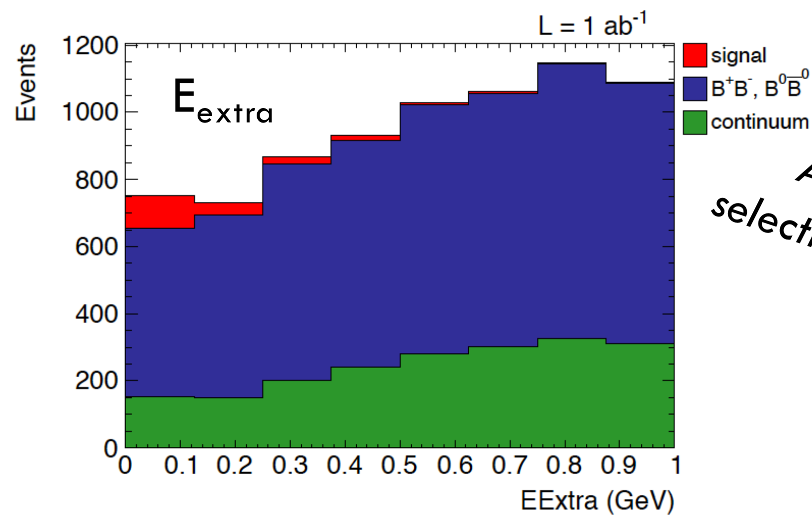
Continuum BDT





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

Cuts on the BDT, M_{bc} ΔE, missing mass, and signal side track momentum are optimized maximizing the FOM* in the M_{bc} and E_{extra} signal windows (respectively 5.275-5.29 GeV/c^2 and 0-0.2 GeV). Optimization for hadronic and leptonic tau decay modes separately



After selection cuts

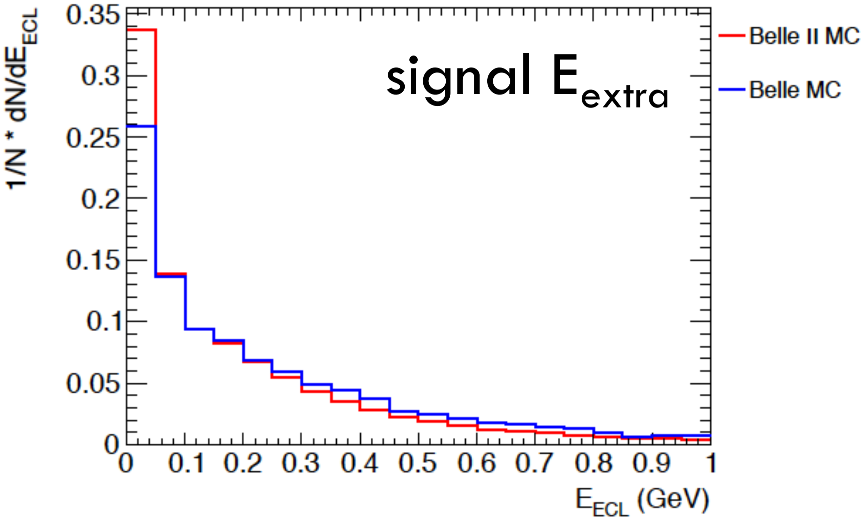
* S/sqrt(S+B) where S is $B \to \tau\nu$ and B is BB+continuum bkg, normalized to 1 ab^{-1}

Selection efficiency and comparison with Belle MC

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

	E_{ECL}	< 1 GeV	< 0.25 GeV
Belle II	# background events	7420	1348
	# signal events	188	136
	signal efficiency (‰)	2.2	1.6
Belle*	# background events	2160	365
	# signal events	97	60
	signal efficiency (‰)	1.2	0.7

*PRL 110, 131801 (2013)



- In Belle II we have higher bkg contamination and higher signal efficiency
- Statistical improvement (S/sqrt(S+B)), but we need to evaluate the systematics impact (e.g. uncertainty on the peaking background)

- E_{extra} has a narrow peak at 0 for Belle II MC → better extra clusters reconstruction despite higher beam background

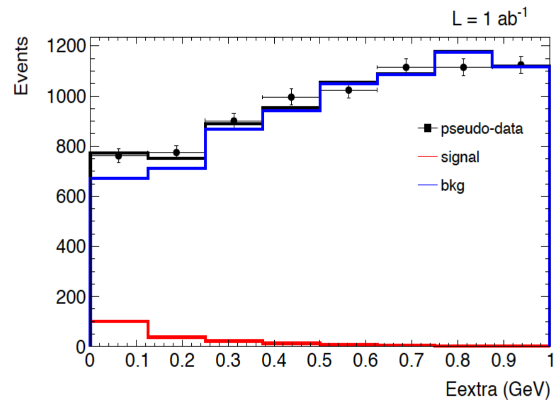
Toy MC and branching ratio measurement

- 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 PDFs taken from MC

- Toy MC with 20000 pseudo-datasets:

$$BR(B \to \tau\nu) = 0.82 \pm 0.24 \times 10^{-4}$$

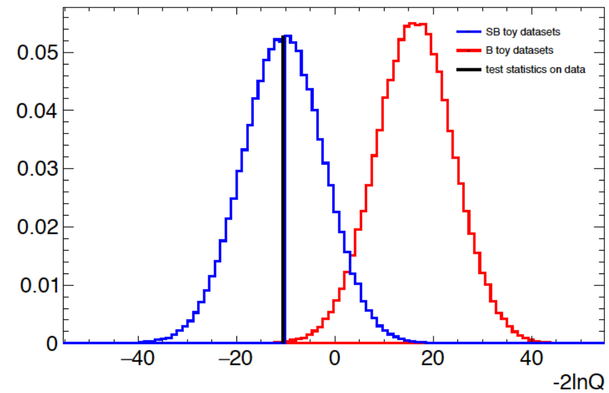
~30% precision



Toy example

- p-value determination: define the test statistics $Q = L(s+b)/L(b)$ and generate pseudo-datasets sampled from S+B and B only E_{extra} distributions. From the left tail of the B only distribution:

p-val = 0.000385 → significance: 3.4 σ



Test statistics distribution

Projections at 1, 5 and 50 ab^{-1} with systematic uncertainties

- The main systematic uncertainties are: signal and background E_{Extra} PDFs, branching fractions of the peaking backgrounds, tagging efficiency, and K^0_L veto efficiency (followed by the signal efficiency and others)
- Extrapolation based on Belle measurement and dedicated Belle II note*: systematic uncertainties splitted into a part behaving as statistical and a part which is expected to be limited (as the branching ratios of the peaking backgrounds).

Integrated Luminosity (ab^{-1})	1	5	50
statistical uncertainty (%)	29.2	13.0	4.1
systematic uncertainty (%)	12.6	6.8	4.6
total uncertainty (%)	31.6	14.7	6.2

luminosity needed for B → tau nu
5σ observation is 2.6 ab^{-1}

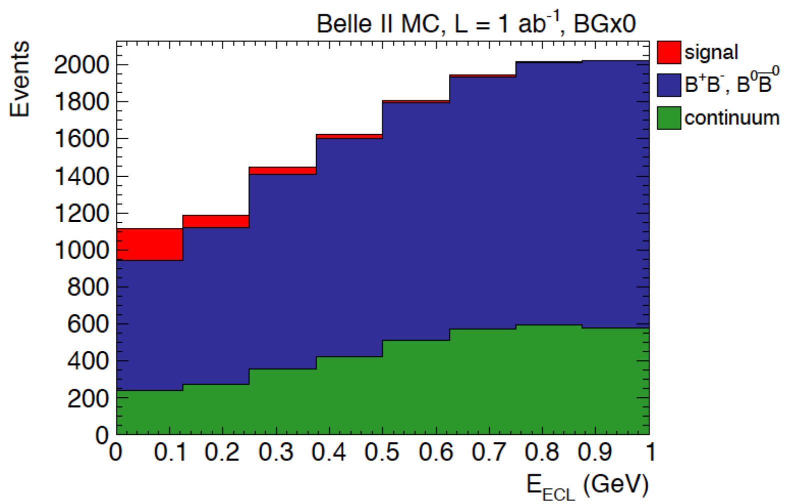
* Belle2-note-0021
https://confluence.desy.de/download/attachments/35838603/belle2-note-0021.pdf?version=1&modificationDate=1468937879217&api=v2

Impact of beam background

- Impact of beam background evaluated re-performing the study described above on MC5 samples without beam background
- Continuum background rejection and signal selection optimized for the new configuration
- Caveat: use the old FEI trained on BGx1 samples and use the extra cluster cleaning optimized on BGx1 samples

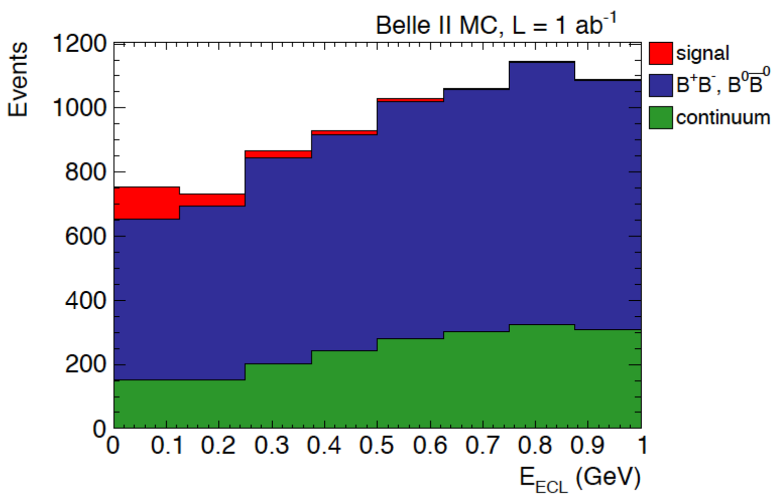
Impact of beam background: E_{extra} and efficiencies

BGx0



E_{ECL}	< 1 GeV	< 0.25 GeV
# background events	12835	2062
# signal events	332	238
signal efficiency (‰)	3.8	2.7

BGx1



E_{ECL}	< 1 GeV	< 0.25 GeV
# background events	7420	1348
# signal events	188	136
signal efficiency (‰)	2.2	1.6

- Toy MC: expected precision for BGx0 is 20% (30% for BGx1) with a significance $> 5\sigma$

Summary and plans

- $B \rightarrow$ tau nu sensitivity study performed on 1 ab^{-1} of MC5 production w and w/o beam background superimposed to physics events (12th bkg campaign)

- Precision of the branching ratio measurement at 1 ab^{-1} of collected luminosity is found to be 30% (3.4σ stat. only) with beam bkg, and 20% (>5σ stat. only) without beam bkg

- Extrapolation of sensitivity with higher luminosity scenarios and including systematic uncertainties has been included in the B2TiP report. With the current performances the $B \longrightarrow \tau\nu$ observation (5σ) is expected at 2.6 ab^{-1}

Summary and Plans

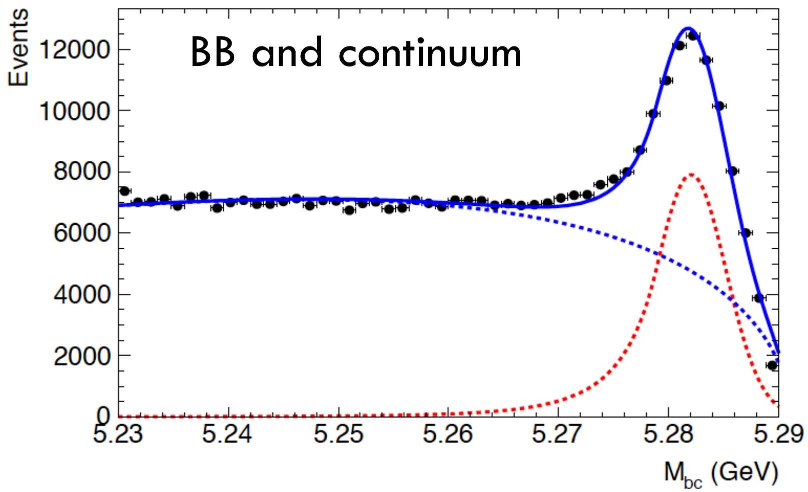
- Comparison with Belle MC (hadronic tag measurement) shows that the E_{extra} distribution is robust against increasing of beam background, and that our selection leads to higher efficiencies resulting in higher S/sqrt(B) in the signal window

- To do list:
  - Analyse MC8 (on the grid):
    - more realistic beam background estimation
    - profiting from new ecl reconstruction optimize extra clusters and pi0 selection (see photons and pi0 presentation later on)
    - optimize FEI output discriminant cut
    - optimize the PID (problem related to running on the grid, see my talk on gbasf2)
  - Study the Eextra peaking background

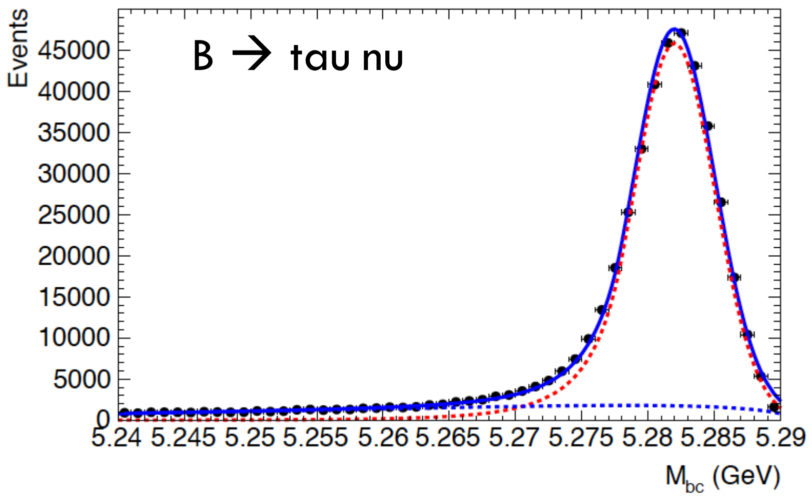
Backup

Purity after FEI reconstruction

• Purity of the background and signal samples is evaluated via fit to M_{bc} distribution with Argus (combinatorics) + double gaussian (true B candidates)



Purity ~50%



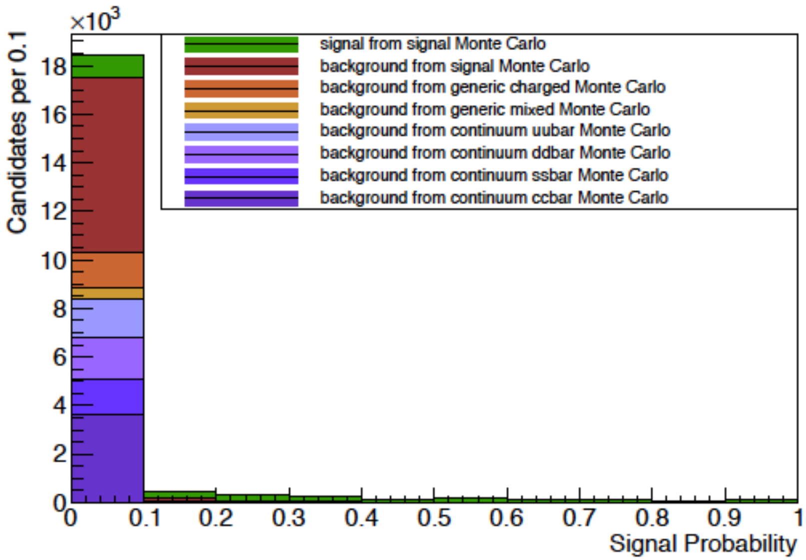
Purity ~93%

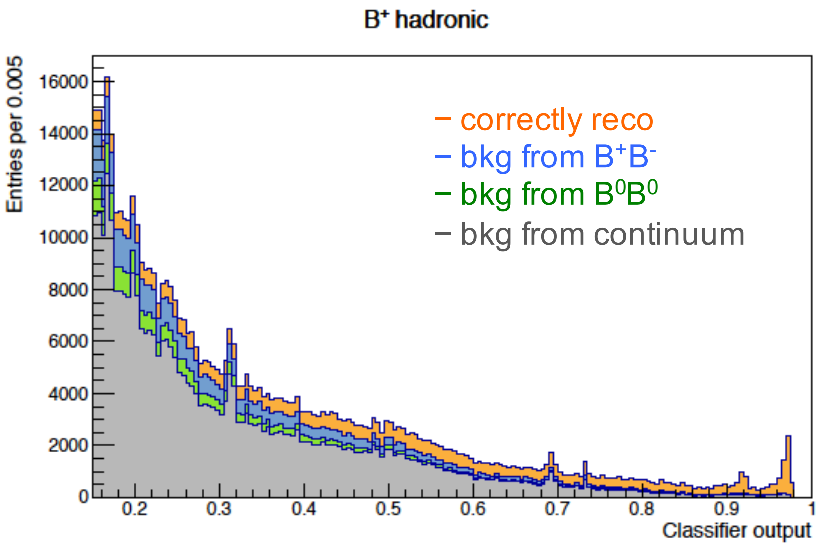
continuum rejection applied

Full Event Interpretation: variables

- Input variables used to train the multivariate classifiers:
  - PID, tracks momenta, impact parameters (charged FS particles);
  - cluster info, energy and direction (photons);
  - invariant mass, angle between photons, energy and direction (π^0);
  - released energy, invariant mass, daughter momenta and vertex quality ($D^{(*)}_{(s)}$, J/ψ);
  - the same as previous step plus vertex position, ΔE (B);
  - additionally, for each particle the classifier output of the daughters are also used as discriminating variables.

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)

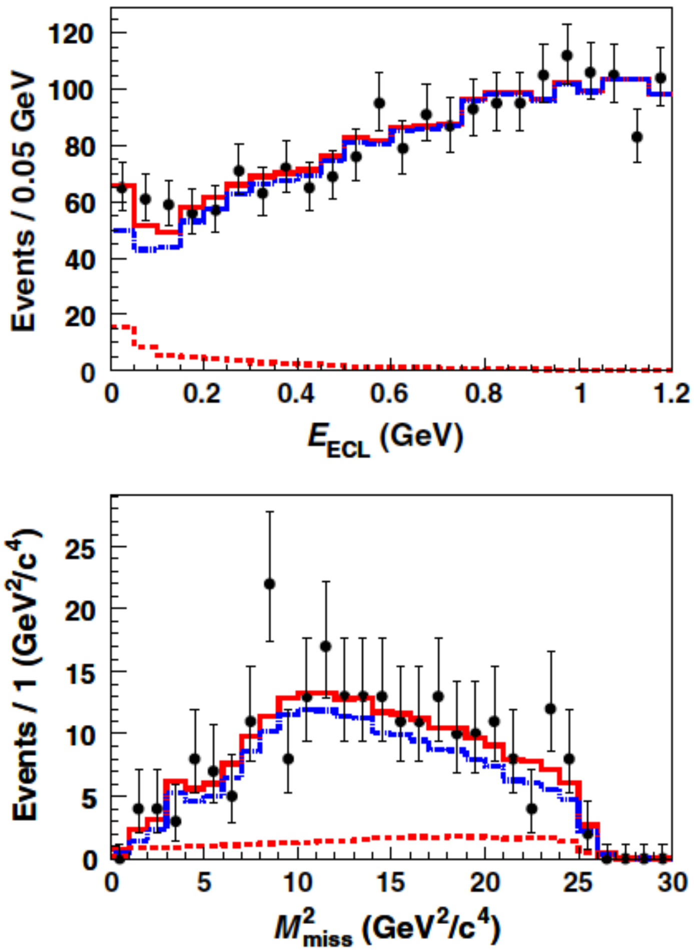
Entire Belle data sample ~ 700 fb^{-1}

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

Submode	N_{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



Photon cuts

Y4S photons

- E > 72 MeV, -114 < clusterTiming < -46, E9E25>0.800, minC2HDist>39 cm – forward
- E > 71 MeV, -112 < clusterTiming < -48, E9E25>0.805, minC2HDist>29 cm – barrel
- E > 66 MeV, -142 < clusterTiming < -18, E9E25>0.710, minC2HDist>23 cm – backward

Each cut corresponds to an efficiency of photons form physics of 95%

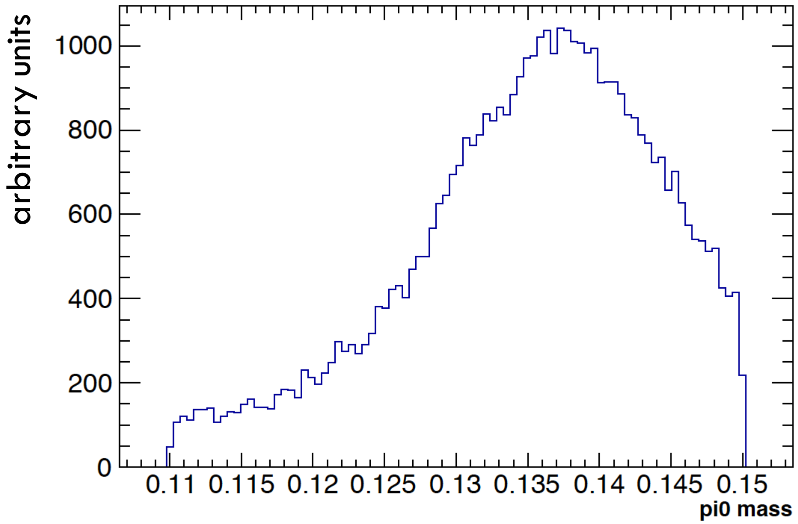
Extra photons

- E > 48 MeV, -121 < clusterTiming < -39, E9E25>0.665, minC2HDist>32 cm – forward
- E > 51 MeV, -123 < clusterTiming < -37, E9E25>0.685, minC2HDist>22 cm – barrel
- E > 49 MeV, -151 < clusterTiming < -9, E9E25>0.650, minC2HDist>24 cm – backward

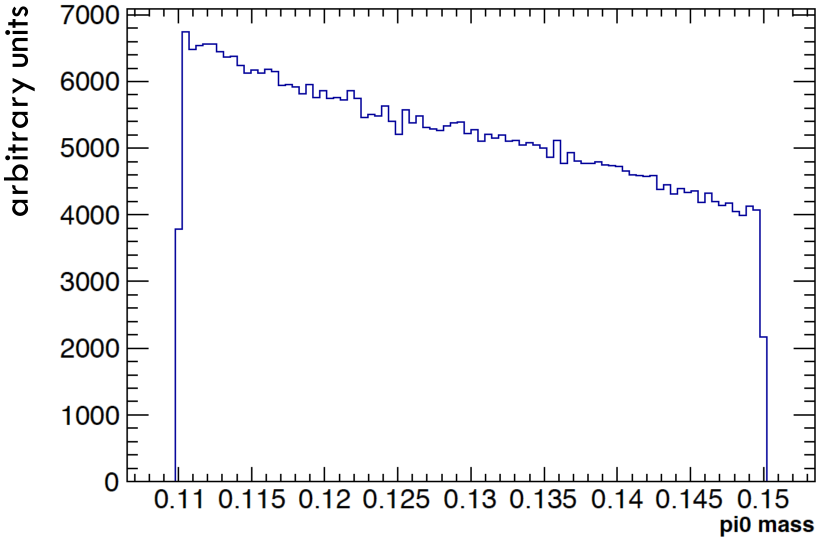
Each cut corresponds to an efficiency of photons form physics of 90%

π^0 mass

After selection



Before selection

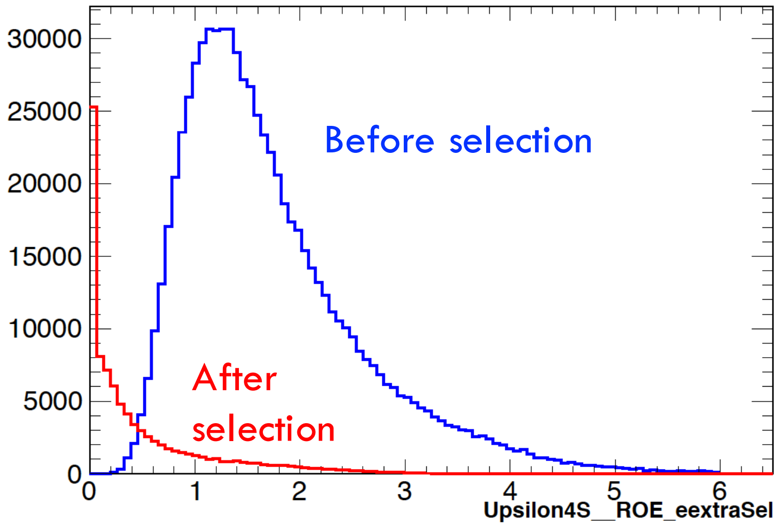


Signal $B \to \tau\nu$ sample

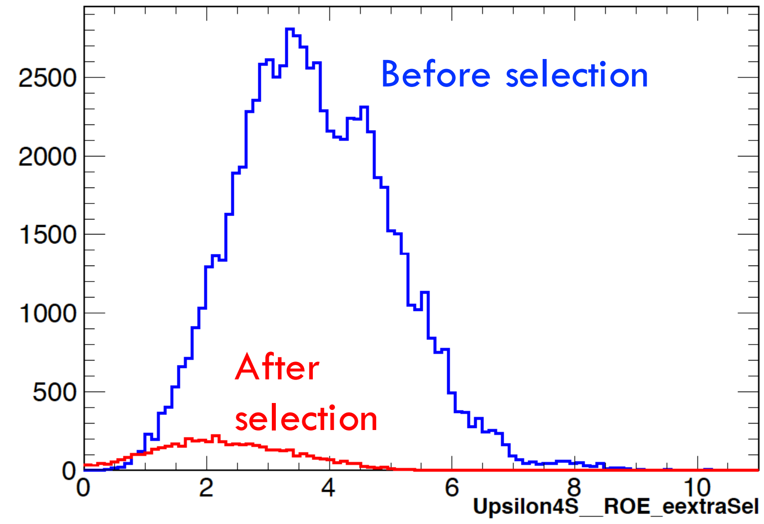
Eextra

Extra cluster cleaning selection

signal $B \to \tau\nu$



B^+B^- bkg



Photon and PID selection eff: 12.2 %

N.B. before PID selection we have a lot of multiple candidates (particle reconstructed as mu and ele and pi)

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} : $|\cos(\vartheta_{thrust})| = \frac{|\vec{p}_B \cdot \hat{T}|}{|\vec{p}_B|}$ 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: $KSFW = \sum_{l=0}^{4} R_l^{so} + \sum_{l=0}^{4} R_l^{oo} + \gamma \sum_{n=1}^{N_t} |(P_t)_n|$

so: particles from b-tag and ROE are considered
oo: particles from ROE only are considered

scalar sum of the transverse momentum of each particle

c: charged,
n: neutral,
m: missing

$$R_l^{so} = \frac{\alpha_{cl} H_{cl}^{so} + \alpha_{nl} H_{nl}^{so} + \alpha_{ml} H_{ml}^{so}}{E_{beam}^* - \Delta E}$$

l odd $H_{cl}^{so} = \sum_i \sum_{jx} Q_i Q_{jx} |p_{jx}| P_l(\cos\theta_{i,jx})$

l even $H_{xl}^{so} = \sum_i \sum_{jx} |p_{jx}| P_l(\cos\theta_{i,jx})$

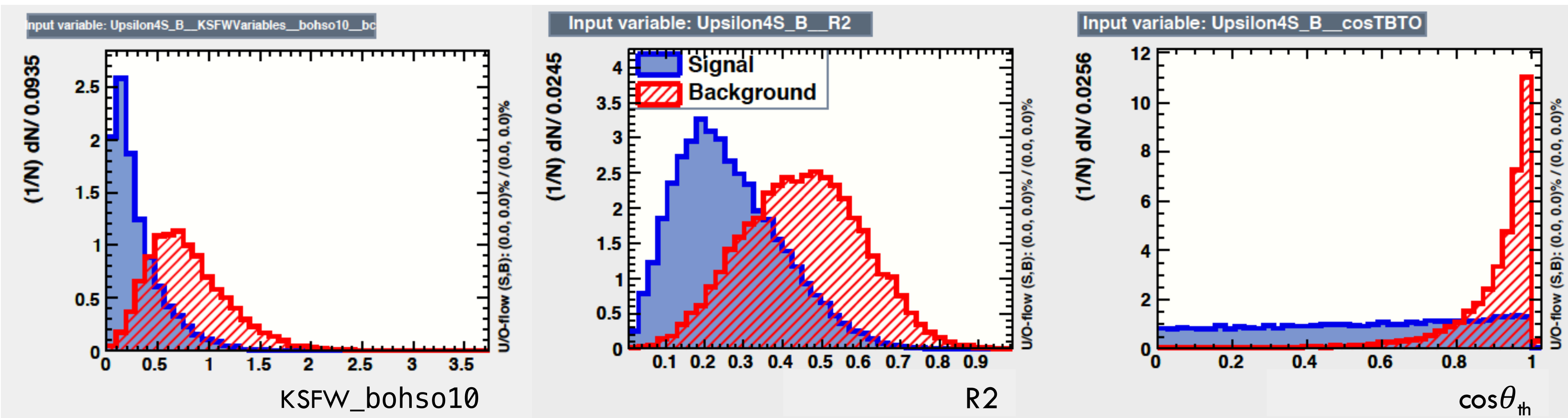
l odd $R_l^{oo} = \sum_j \sum_k \beta_l Q_j Q_k |p_j||p_k| P_l(\cos\theta_{j,k})$

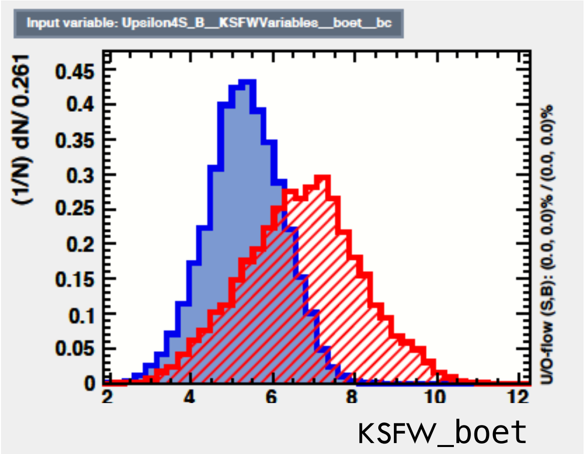
l even $R_l^{oo} = \sum_j \sum_k \beta_l |p_j||p_k| P_l(\cos\theta_{j,k})$

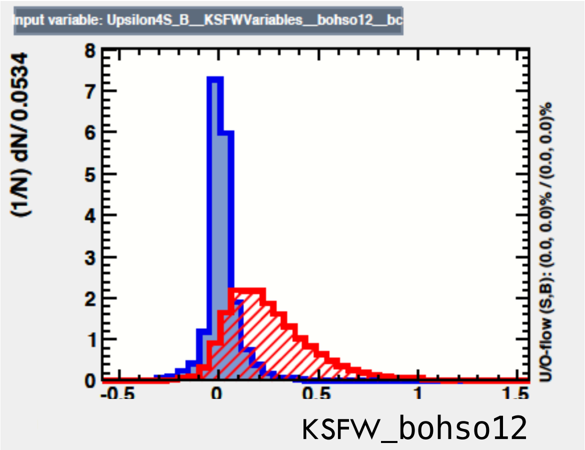
Continuum rejection - BDT

- BDT training
  - Preselection cuts on M_{bc} (5.27-5.29 GeV/c^2) and E_{extra} (< 0.3 GeV)
  - 20000/3000 events used for signal/background training (~3/10% of the entire samples)
  - Remove the "less powerful" (according to the BDT variable ranking) and highly correlated variables → 20 variables left with a negligible degradation of the BDT performances (i.e. ROC curve integral)

Highest ranked variables



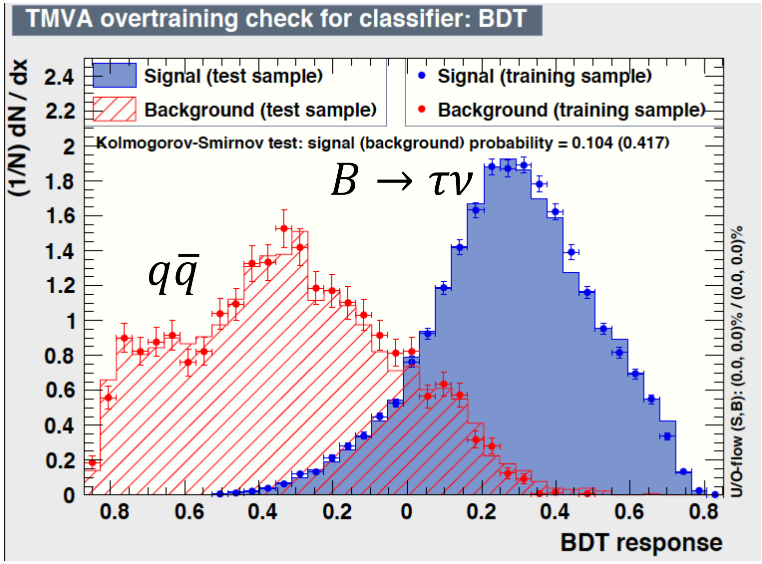


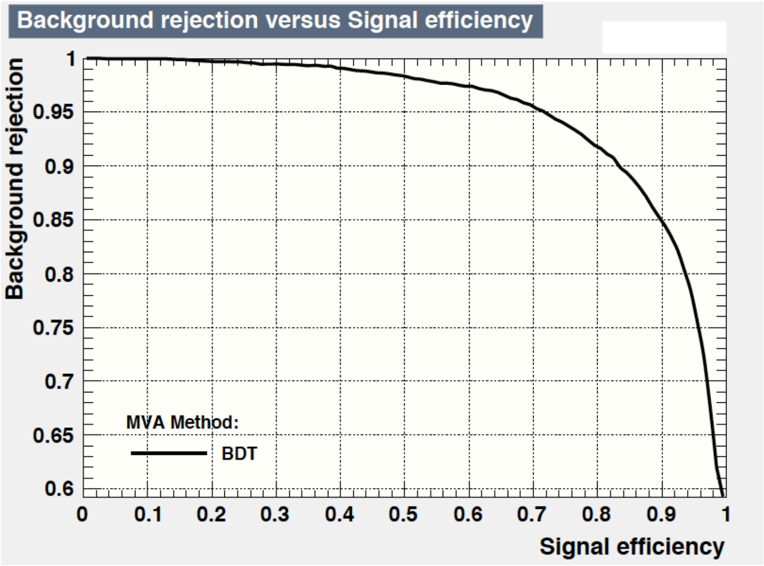


Signal: $B \to \tau\nu$

Background: $q\bar{q}$

Training results





Overtraining under control
Limited statistics for the backgrounds

ROC curve

- Good separation power

B → tau nu projections

Hadronic tag

Integrated Luminosity (ab^{-1})	1	5	50
statistical uncertainty (%)	29.2	13.0	4.1
systematic uncertainty (%)	12.6	6.8	4.6
total uncertainty (%)	31.6	14.7	6.2

Semileptonic tag

Integrated Luminosity (ab^{-1})	1	5	50
statistical uncertainty (%)	19.0	8.5	2.7
systematic uncertainty (%)	17.9	8.7	4.5
total uncertainty (%)	26.1	12.2	5.3