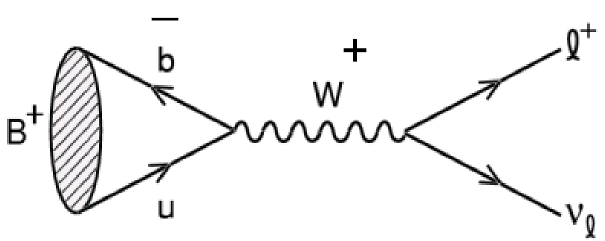
$B \to \tau\nu$ status update

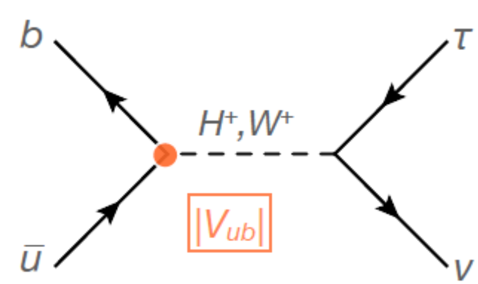
Mario Merola[1], *Elisa Manoni*[3,4], *Guglielmo De Nardo*[1,2]

[1] INFN Napoli, [2] Università di Napoli, [3]INFN Perugia, [4]Università di Perugia

Belle II Italia

Roma, 14/12/16

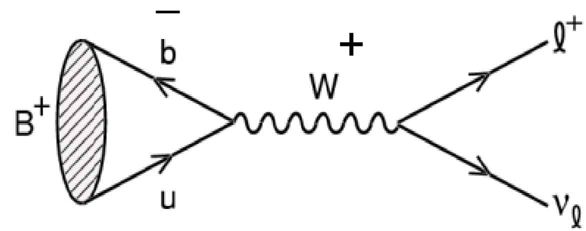




Outline

- Theory introduction and recent results overview
- B-tag reconstruction: Full Event Interpretation
- Selection and continuum rejection
- Sensitivity of the analysis with a luminosity of 1 ab^{-1}
- Conclusions and future plans

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



- Helicity suppressed

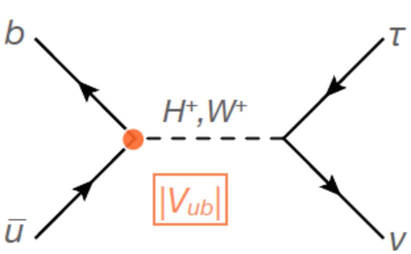
$$BR_{SM}\left(B \to \ell\, \nu\right) = \frac{G_F^2 m_B \tau_B}{8\pi} f_B^2\, |V_{ub}|^2\, m_\ell^2 \left[1 - \frac{m_\ell^2}{m_B^2}\right]^2$$

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

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

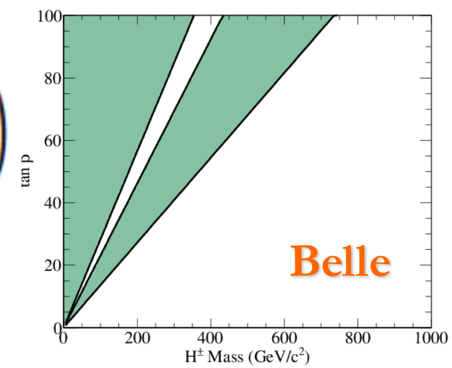
http://ckmfitter.in2p3.fr/

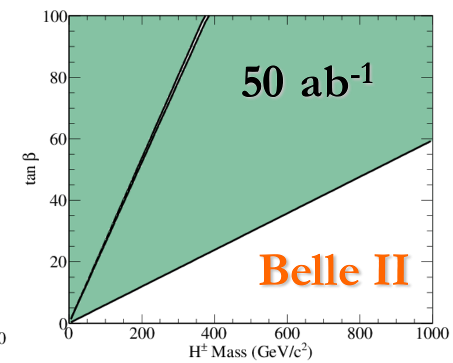
Higgs doublet models predict interference with SM decay with a modification of the branching ratio [PhysRevD.86.054014]



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

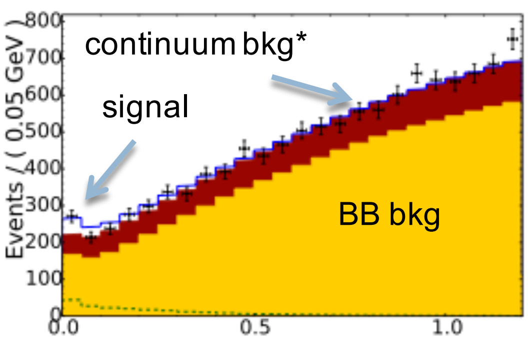
ratio of the two Higgs vacuum expectation values

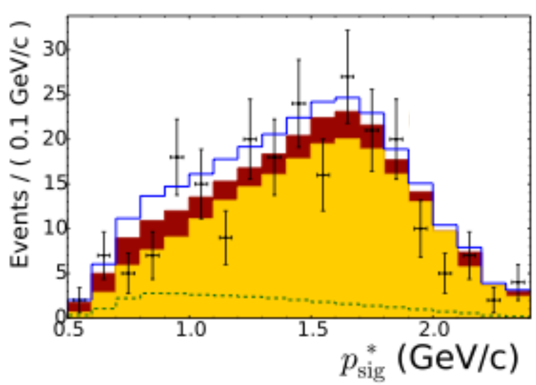




Recent results on $B \to \tau\nu$

- 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

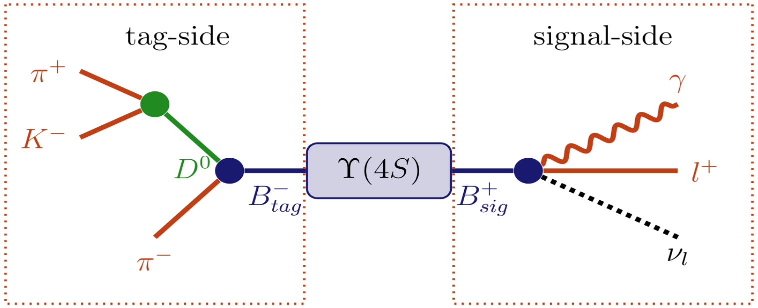
- E_{ECL} (later on called E_{extra}) is the sum of the energies of clusters in the ECL not associated to reconstructed B mesons
- p^*_{sig} is the momentum of the signal side particle in the CM

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

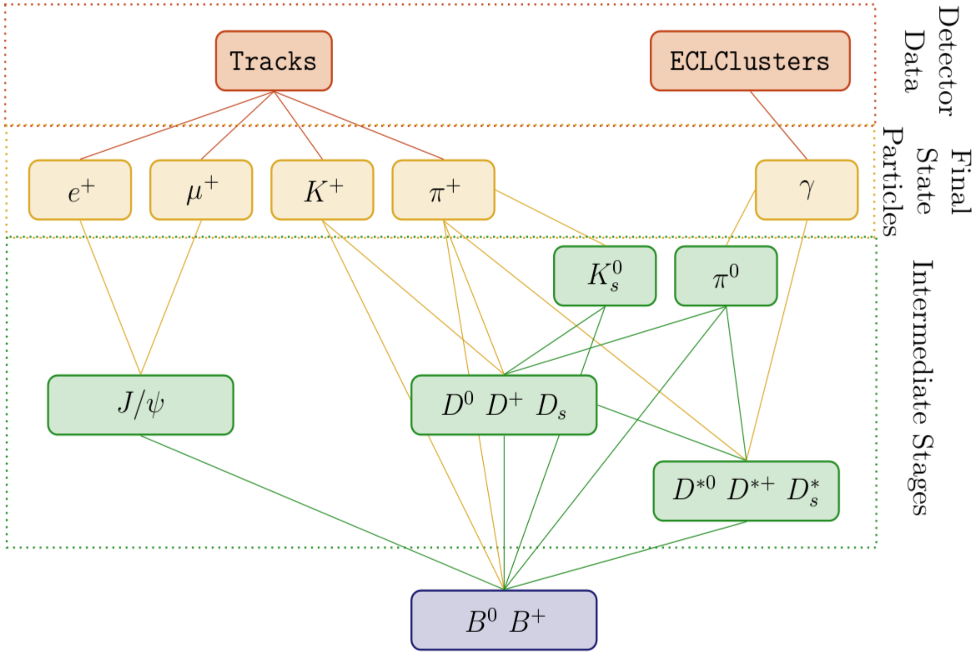
http://arxiv.org/abs/1503.05613v2

Tag side reconstruction in Belle II: 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 $100*10^6$ B^+B^- / $B^0\bar{B}^0$ events with beam background

- The result of the training is analysis independent.

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

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.

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: $M_{bc} = \sqrt{E^{*2}_{beam} - p^{*2}_{B}}$

* 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

Belle II

Selection (2)

INFN

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, E9/E25 and minimum distance between the cluster and tracks in the event (separately in forward, barrel and backward detector regions).
- Different cuts for "extra" clusters and for clusters reconstructing pi0s

Continuum rejection

- MVA with boosted decision trees (next slide)

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

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: $KSFW = \sum_{l=0}^{4} R_l^{so} + \sum_{l=0}^{4} R_l^{oo} + \gamma \sum_{n=1}^{N_t} |(P_t)_n|$

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

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

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

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

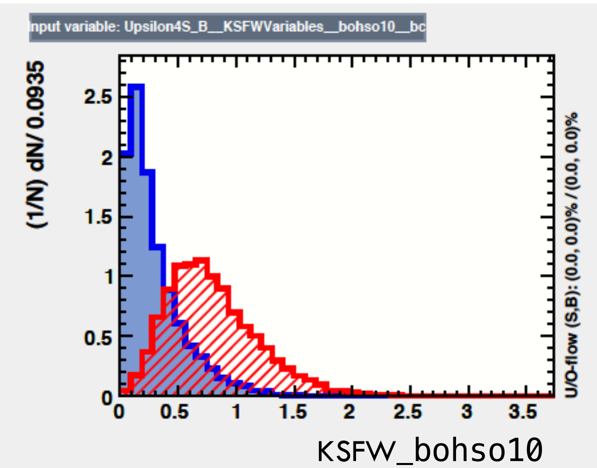
l odd $\quad R_l^{oo} = \sum_j \sum_k \beta_l Q_j Q_k |p_j||p_k| P_l(\cos\theta_{j,k})$

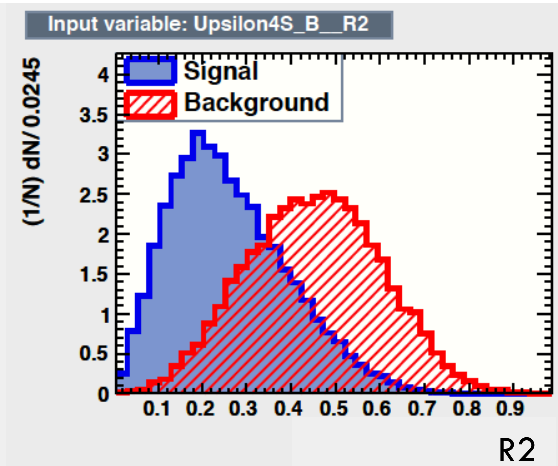
l even $\quad R_l^{oo} = \sum_j \sum_k \beta_l |p_j||p_k| P_l(\cos\theta_{j,k})$

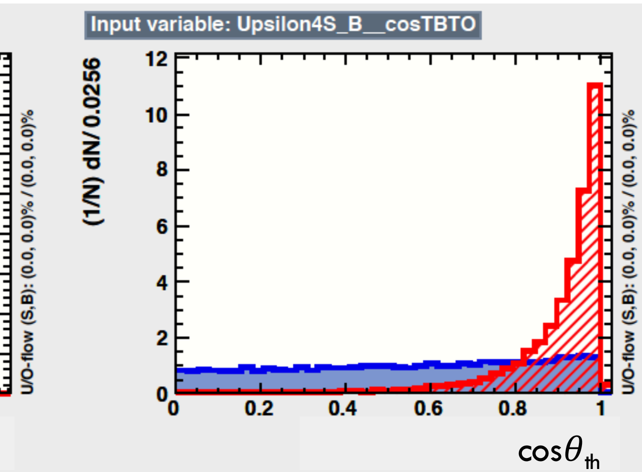
Continuum rejection - BDT

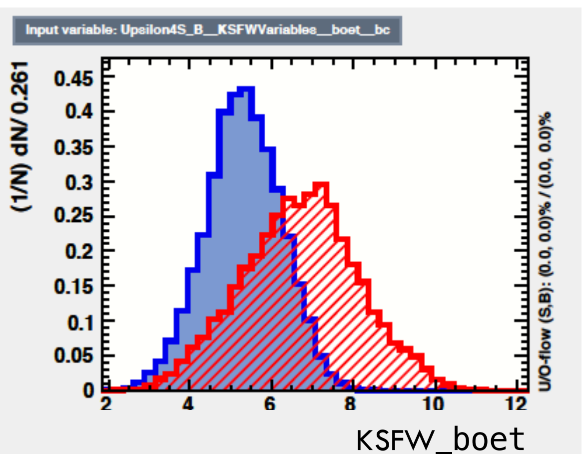
- BDT training
  - Preselection cuts on M_{bc} (5.27-5.29 GeV/c^2) and E_{extra} (< 0.3 GeV)
  - 20000/3000 events used for signal/background training (~3/10% of the entire samples)
  - Remove the “less powerful” (according to the BDT variable ranking) and highly correlated variables → 20 variables left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

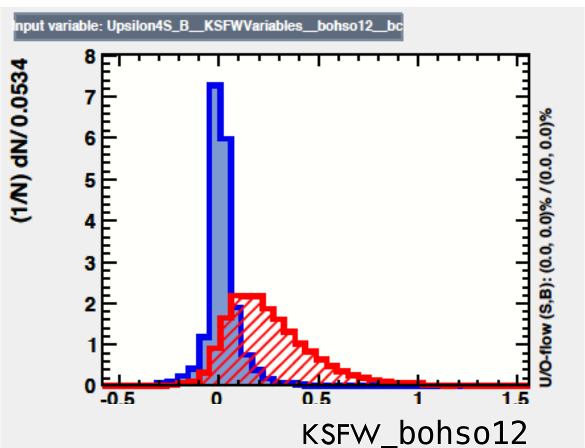
Highest ranked variables







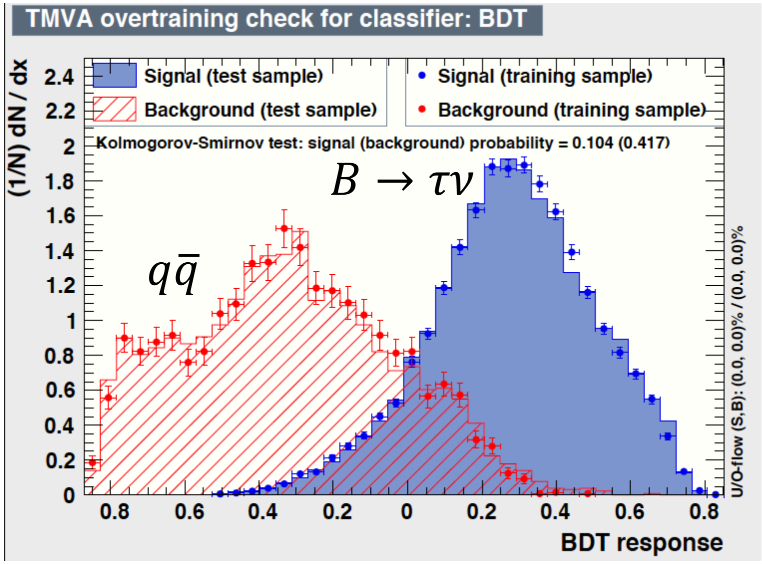


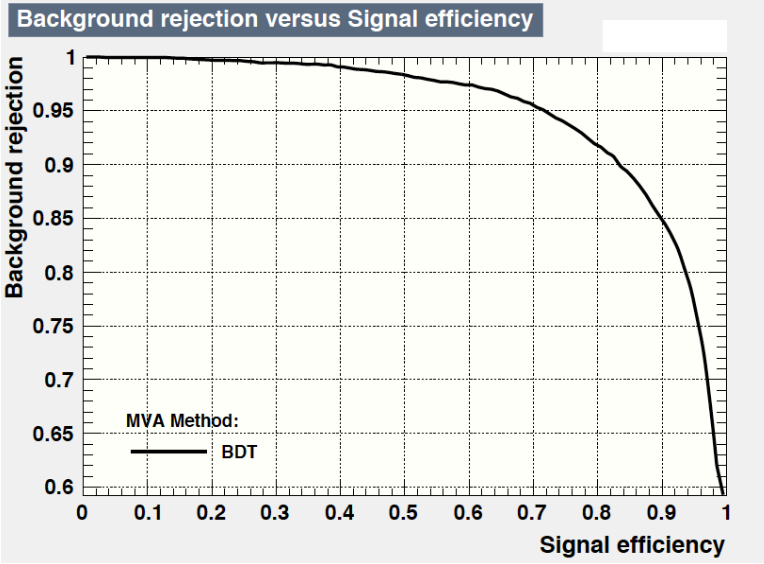


Signal: $B \to \tau\nu$

Background: $q\bar{q}$

Training results



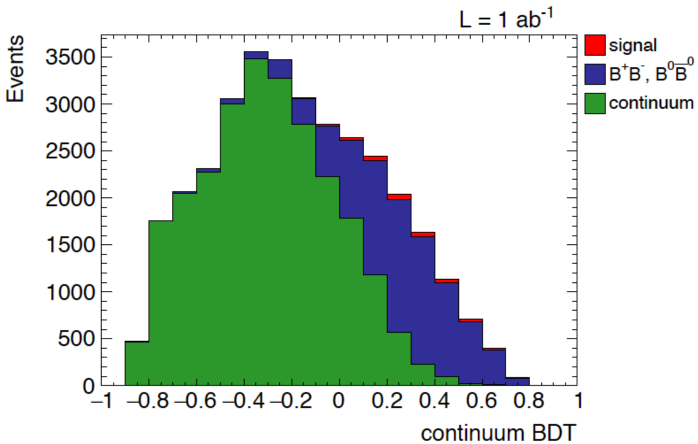


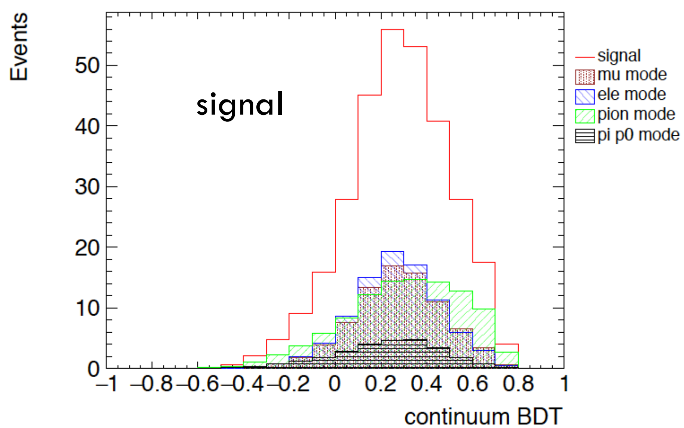
Overtraining under control
Limited statistics for the backgrounds

ROC curve

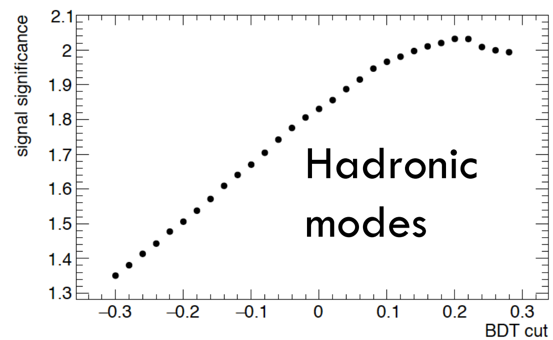
- Good separation power

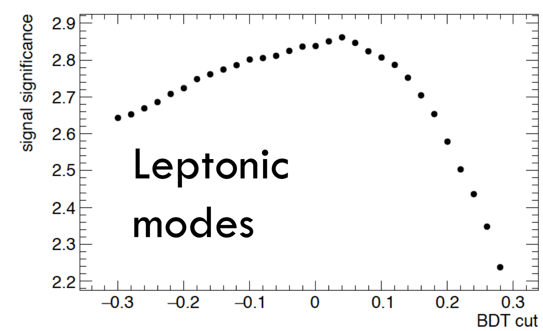
Continuum rejection: BDT





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}

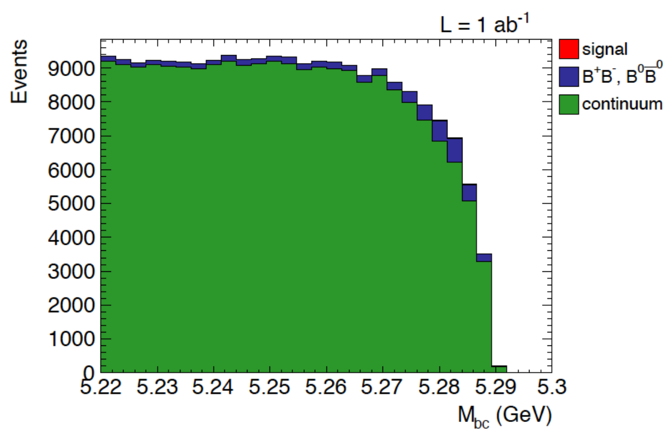
*estimated as S/sqrt(S+B) where **S** is $B \to \tau\nu$ and **B** is BB+continuum bkg, normalized to 1 ab^{-1}

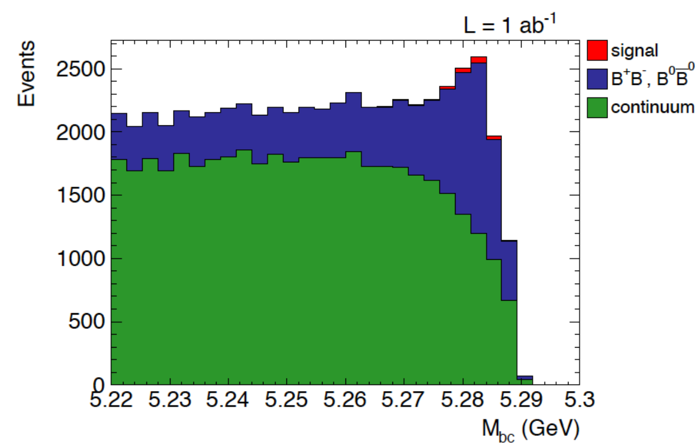
M_{bc} distribution

hadronic modes

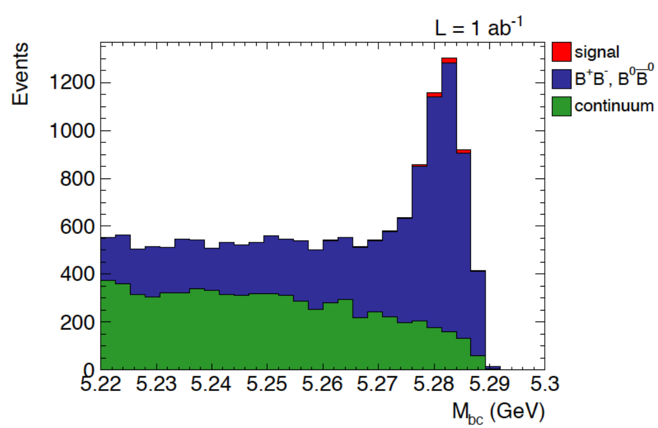
leptonic modes

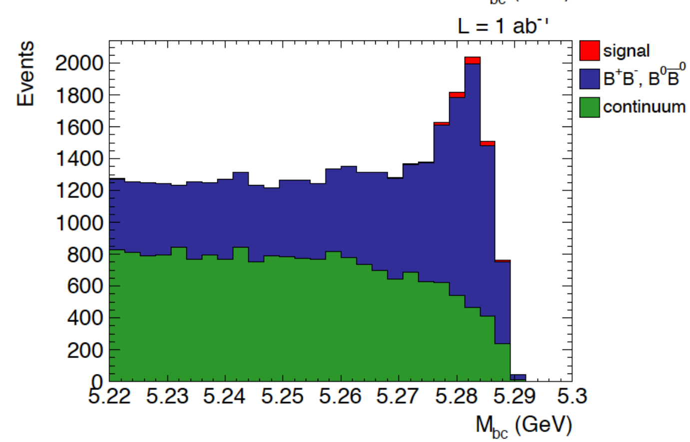
Before BDT cuts





After BDT cuts



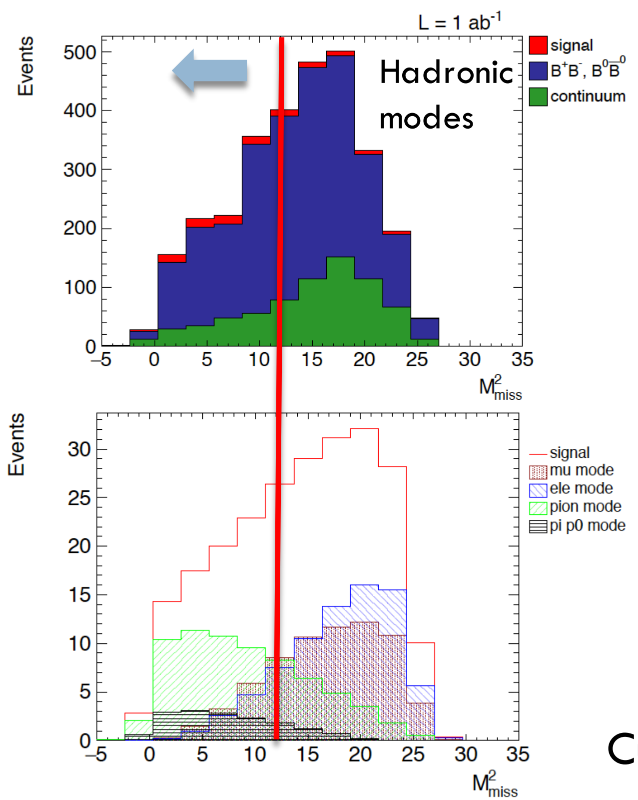


$5.275 < M_{bc} < 5.29$ GeV/c^2

Missing mass squared

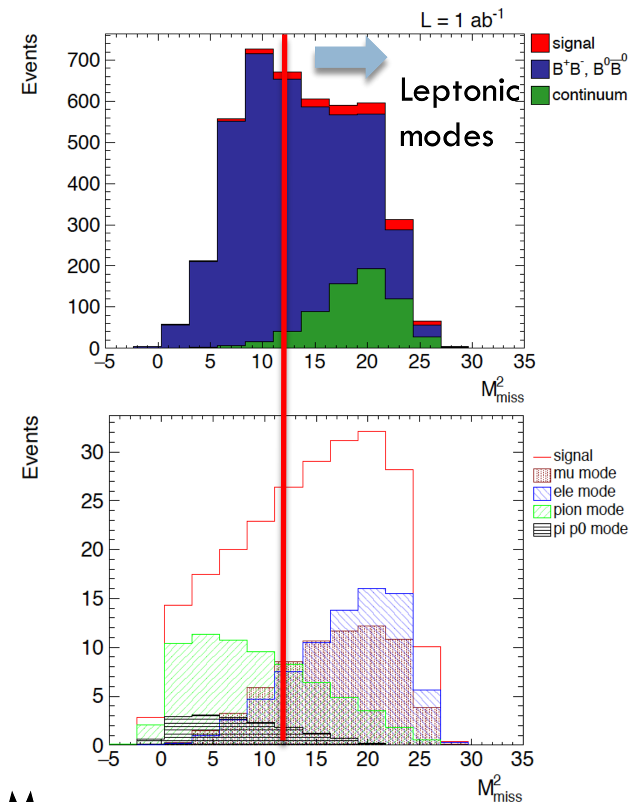
Missing mass squared

$$M_{miss}^2 = \left(2E_{beam} - E_{B_{tag}} - E_{B_{sig}}\right)^2 - \left|\vec{p}_{B_{tag}} + \vec{p}_{B_{sig}}\right|^2$$



M_{miss}^2 < 12 GeV²

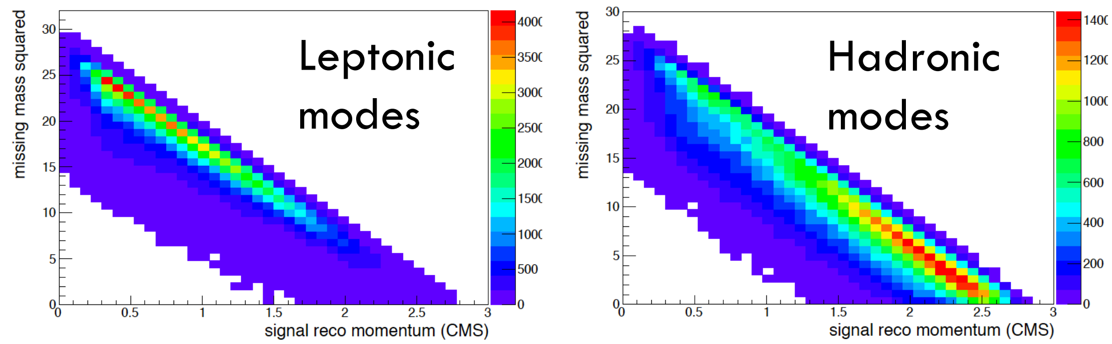
Cut points chosen maximizing the FOM: S/sqrt(S+B)



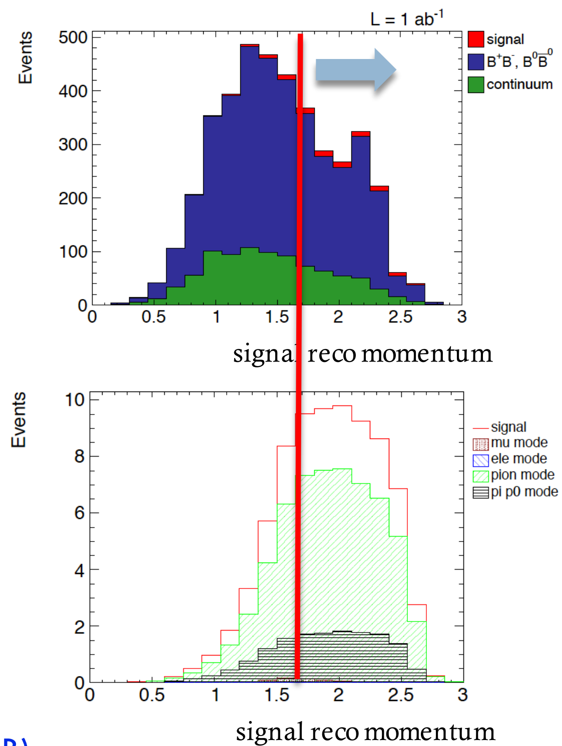
M_{miss}^2 > 12 GeV²

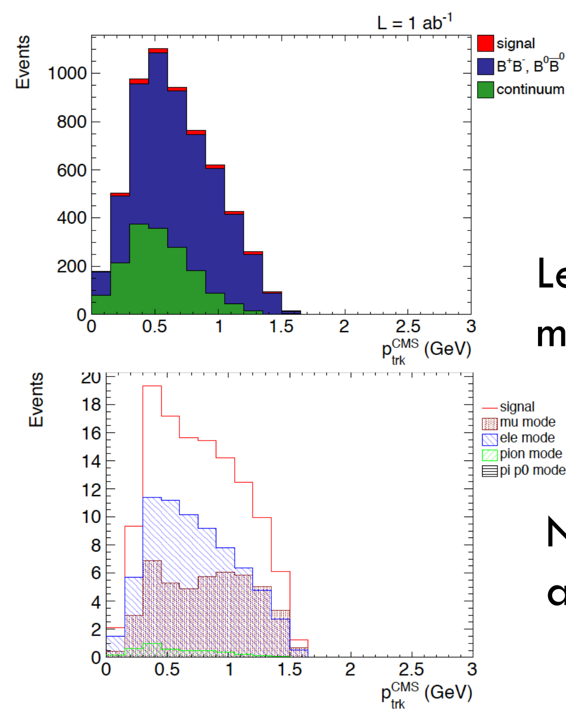
Signal side reconstructed momentum

Highly correlated with the missing mass



Hadronic modes





Leptonic modes

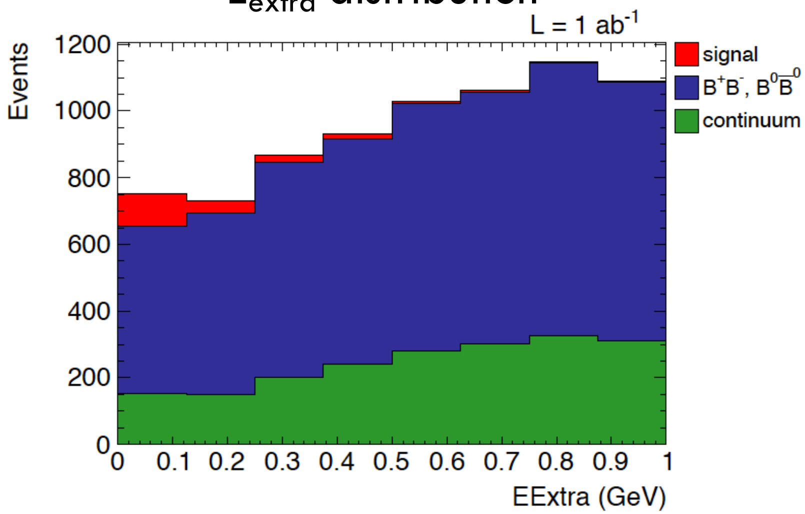
No cut applied

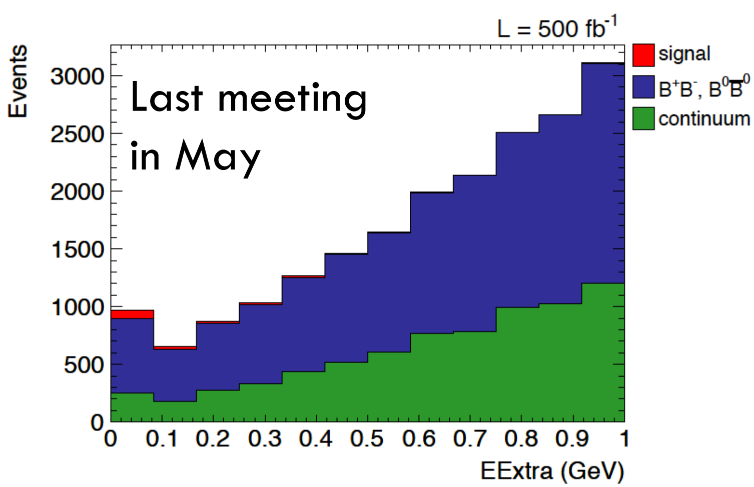
FOM: S/sqrt(S+B)

p_{sig} (CMS) > 1.6 GeV

E_{extra} and selection efficiency

E_{extra} distribution





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

Eextra < 1 GeV
- sig: 188 events
- bkg: 7420 events

(1965 qq + 5455 BB)

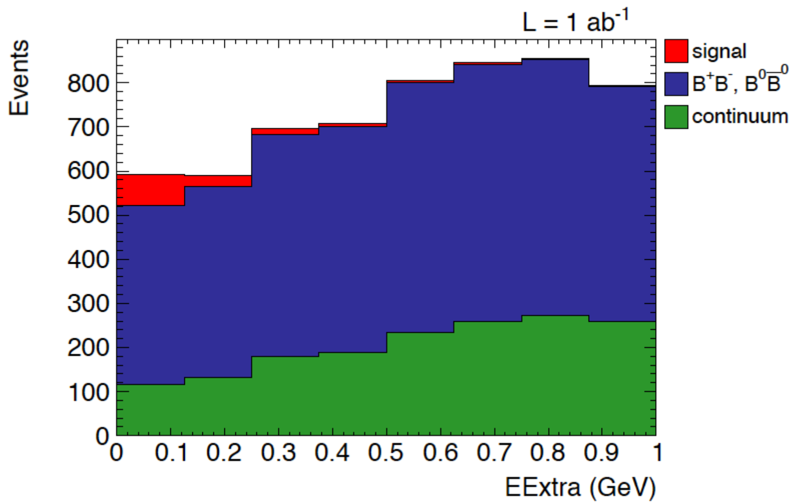
Eextra < 0.2 GeV
- sig: 123 events
- bkg: 1013 events

E_{extra} < 1 GeV	Babar PRD 88, 031102 (2013)	Belle PRL 110, 131801 (2013)	Belle II (this analysis)
Signal Efficiency (‰)	**0.72**	**1.1**	**2.2**

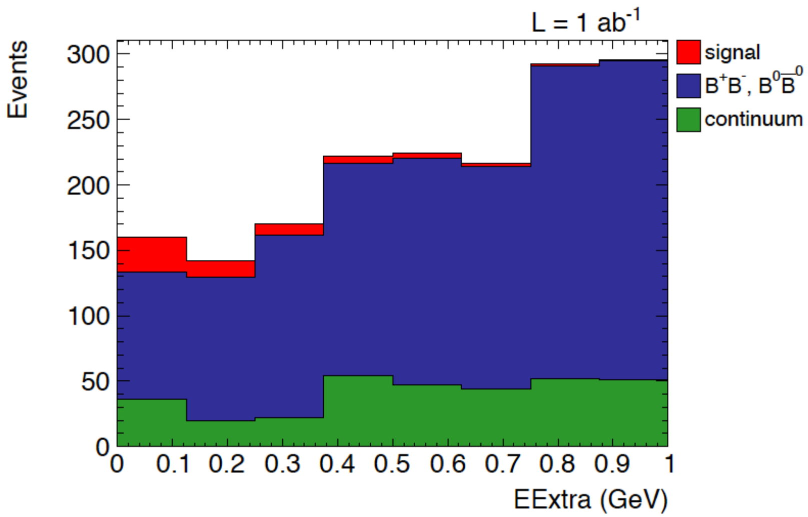
N.B. “offline” rough correction of the branching ratios applied: in MC5 many B decay modes were modelled with wrong BR → mean weight of 0.62 applied to correctly reconstructed B candidates (for both signal and background).

E_{extra} and event yields in the different channels

Leptonic channels



Hadronic channels



#events in (0 - 1.0) GeV	Lep channels	Had channels	Total
Signal	126	62	188
Background	5758	1662	7420

#events in (0 - 0.2) GeV	Lep channels	Had channels	Total
Signal	88	35	123
Background	817	196	1013

Toy MC study 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 two components: signal and background pdfs built from the expected MC distributions

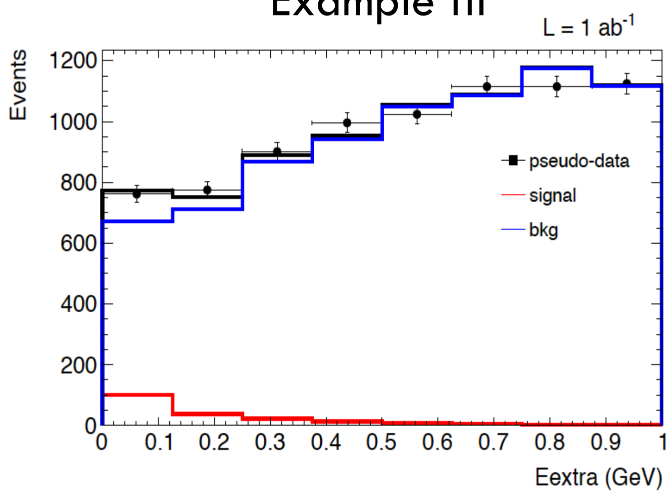
- Toy MC with 20000 pseudo-datasets:

mean fitted signal yield = 188 evts → no bias
mean uncertainty = 55 evts

$$BR(B \to \tau\nu) = 0.83 \pm 0.24 \times 10^{-4}$$

~30% precision

Example fit

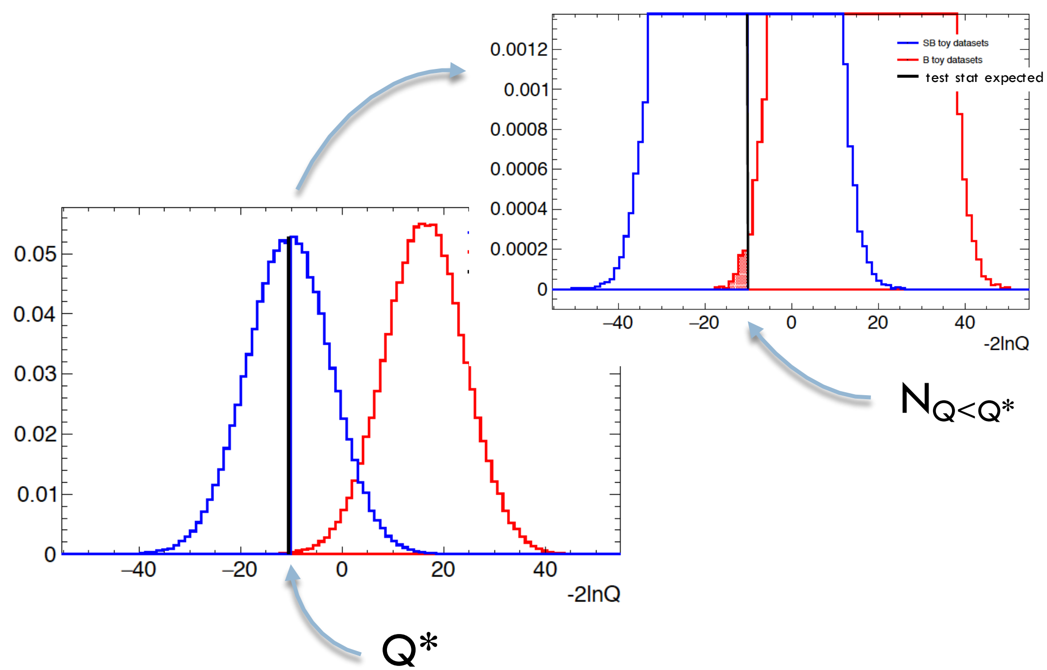


Expected sensitivity of the measurement with CLb method

- Define the test statistics $Q = -2\ln[L(s+b)/L(b)]$ and perform 200000 pseudo-experiments generating pseudo-datasets sampled from S+B and B only E_{extra} distributions.

- Evaluate the expected p-value of the null hypohesis on the toys background samples as $1\text{-}CL_b = N_{Q<Q^*}/N$, where $N_{Q<Q^*}$ is the number of pseudo-experiments with Q lower than the mean of the test statistics distributions on the S+B toy samples Q*, and N is the total number of pseudo-experiments.

blue hist distribution of Q evaluated on S+B toy datasets
red hist: distribution of Q evaluated on B only toy datasets
Black line: expected value of Q in the S+B hypothesis



p-val = 0.000385 → significance: 3.4 σ

Summary and plans

- 1 ab^{-1} of MC5 production analysed

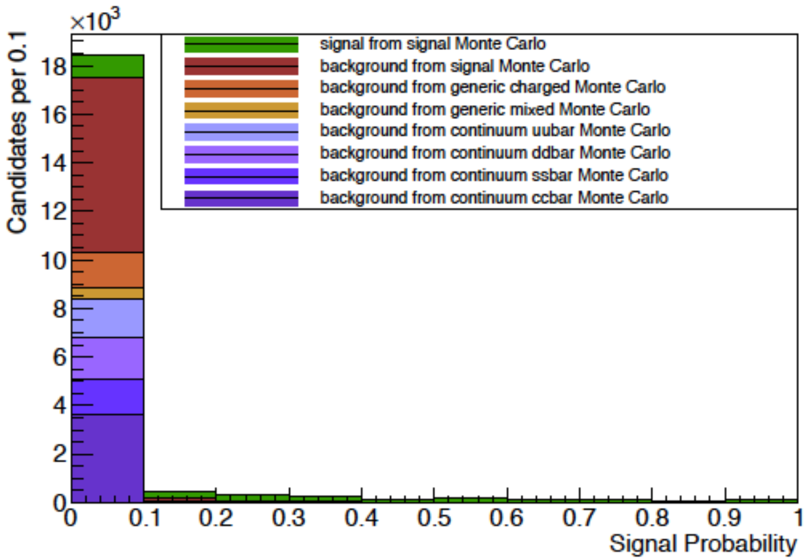
- The signal efficiency in the region Eextra < 1 GeV is found to be 2.2 ‰ (bkg efficiency ~$10^{-6} \div 10^{-7}$)

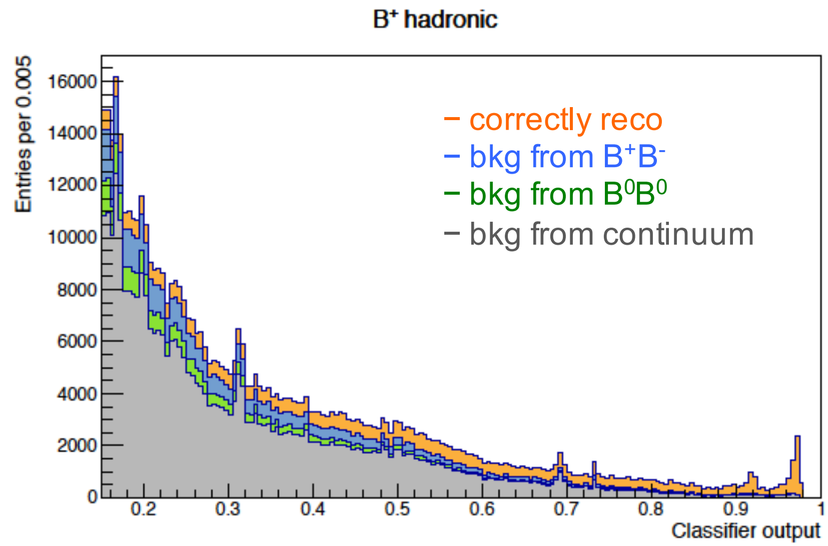
- Sensitivity/precision of the analysis estimated with toy MC performing maximum likelihood fit to E_{extra} distribution. In 1 ab^{-1}, assuming a branching ratio of 0.83×10^{-4}, the expected statistical uncertainty is ~30%, corresponding to an expected sensitivity of 3.4 σ.

- To do list:
  - Run on MC7 with correct branching ratios (MC7 complete by the end of the month)
  - Set up a simultaneous fit assuming the leptonic and hadronic tau BRs
  - Perform a realistic estimation of the systematic uncertainties (B tag efficiency, signal efficiency)
  - Study the Eextra peaking background (mainly semileptonic B decays with KL)

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)

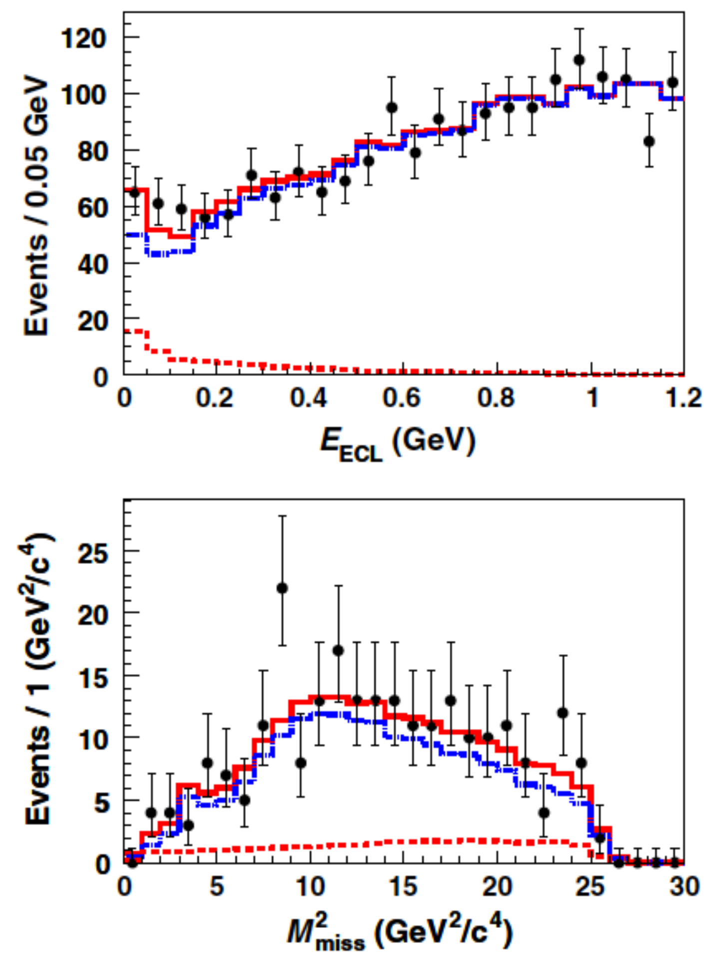
Entire Belle data sample ~ 700 fb^{-1}

TABLE I. Results of the fit for $B^- \to \tau^- \bar{\nu}_\tau$ yields (N_{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%

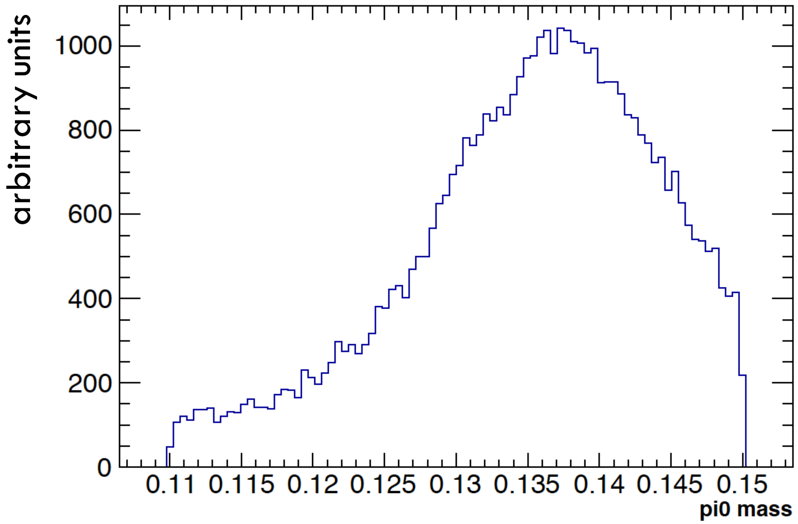
Extra photons

- 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

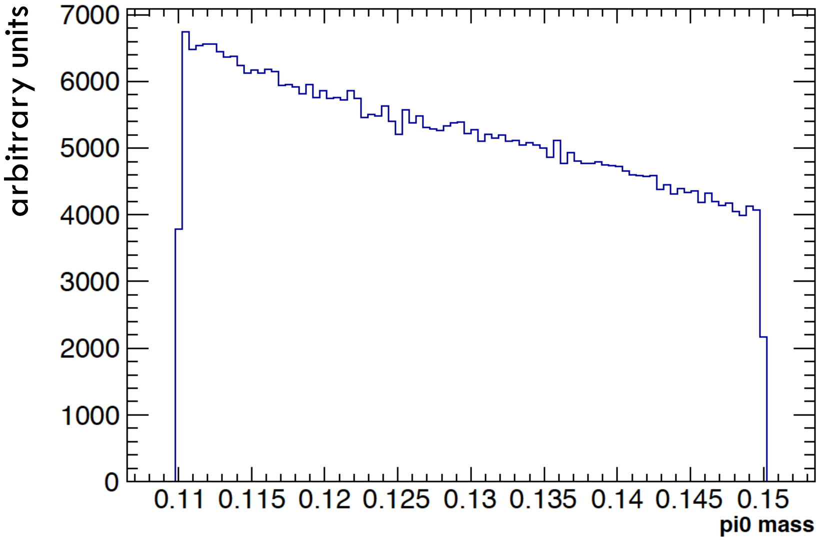
Each cut corresponds to an efficiency of photons form physics of 90%

π^0 mass

After selection



Before selection

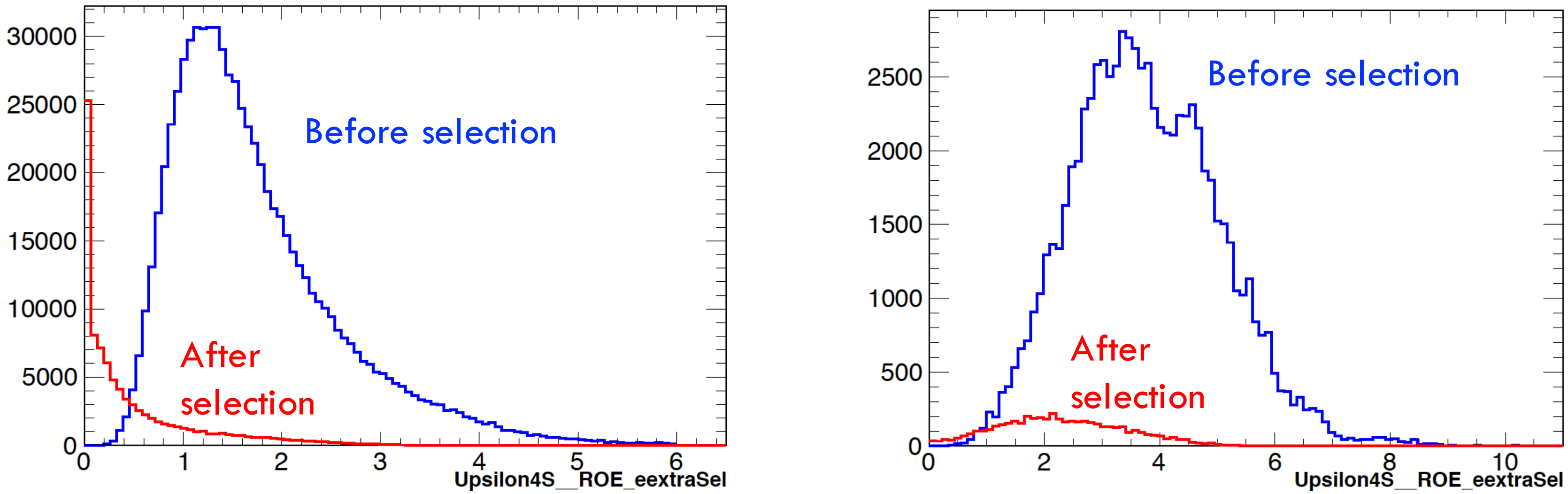


Signal $B \to \tau\nu$ sample

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)