Search for lepton flavour violation in $B \to \rho\tau\ell$ at Belle and Belle II

Seminar CPPM

Isaac Consigny

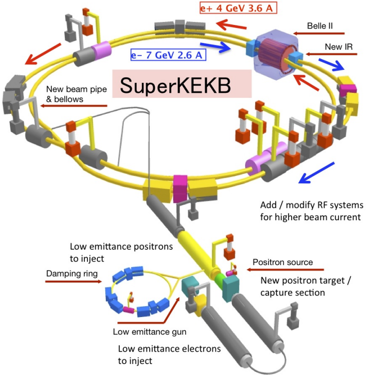
November 24, 2025

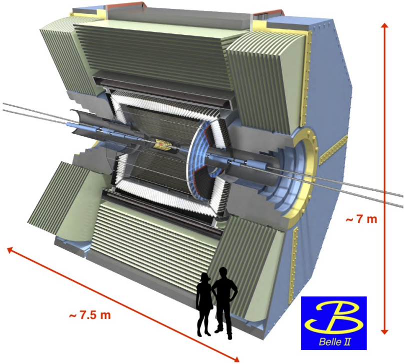
Outline

- Introduction
  - The Belle II experiment
  - Btagging at Belle II
  - Lepton Flavour violation at Belle II: state of the art
- Context
- Analysis strategy
- Selection
- Simultaneous fits on Belle and Belle II
- Control samples and BDT systematics
- Expected upper limits on the branching fractions
- GNN based alternative reconstruction: GraFEI

The Belle II experiment at SuperKEKB: e^+e^- collider

$\rightarrow$ Center of mass energy: 10.58 GeV (peak of the $\Upsilon(4S) \rightarrow B\bar{B} \approx 96\%$)
$\rightarrow$ Very clean B decays events
$\rightarrow$ Lepton Flavour Violation,CP violation, rare B decays, dark sector, tau factory ...
$\rightarrow$ Target luminosity: 50 ab^{-1}





Hadronic / Semi-leptonic tagging

$\rightarrow$ Full Event Interpretation (FEI)

$\rightarrow$ Reconstruct one B (tag side) in many modes using multivariate algorithms

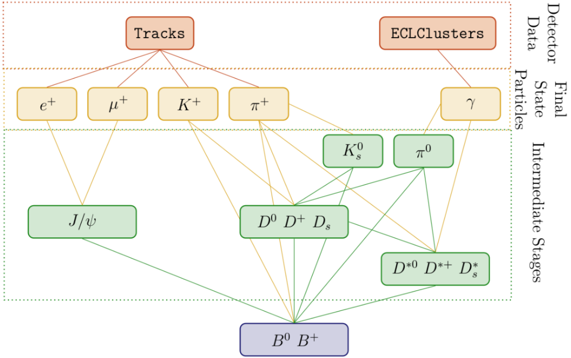
$\rightarrow$ Very clean reconstruction of the tag side

$\rightarrow$ low efficiency ≈ 1 %

Inclusive tagging

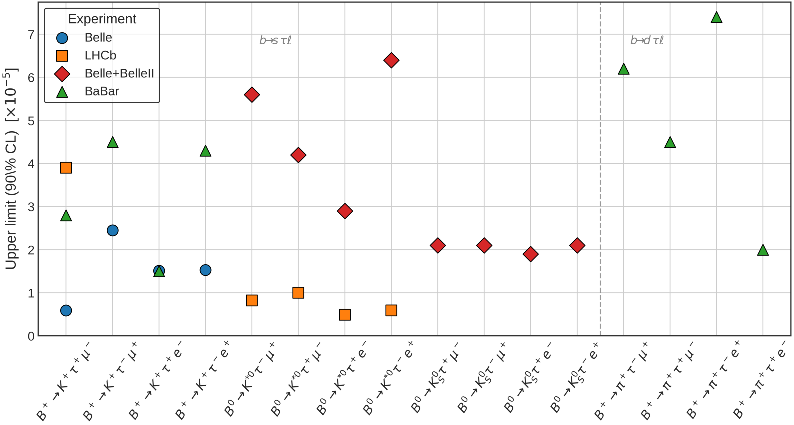
$\rightarrow$ Showed good results in $B \rightarrow K\nu\bar{\nu}$ for example

FEI diagram



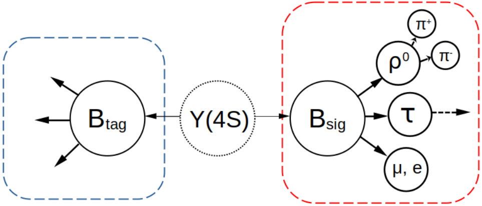
Lepton Flavour Violation at Belle II: state of the art

→ LFV involving τ leptons are good probes for new physics (enhanced BR in NP models)
→ Search for LFV in $b \to s\tau\ell$ decays largely studied at Belle II
→ $b \to d\tau\ell$ not studied yet at Belle II or LHCb



$B \to \rho\tau\ell$ - Context

- Search for lepton flavour violation $b \to d\tau\ell$: probe for new physics
- 4 modes to analyse:
  $B^0(\bar{b}d) \to \rho^0\tau^+\ell^-$: OSe and OSmu
  $B^0(\bar{b}d) \to \rho^0\tau^-\ell^+$: SSe and SSmu
  OS : charge of ℓ opposite sign of the b quark
  SS : charge of ℓ same sign as the b quark



- Upper limits on BR computed in Z' models (arxiv 2407.19060):
  $BR \approx 10^{-10}$
- $b \to d\tau\ell$ coupling measured with $B \to \pi\tau\ell$ at BaBar (arXiv:1204.2852):
  $BR \approx 10^{-5}$

$\rightarrow$ This analysis will be the **first direct measurement**

$\rightarrow$ Belle + Belle 2 samples

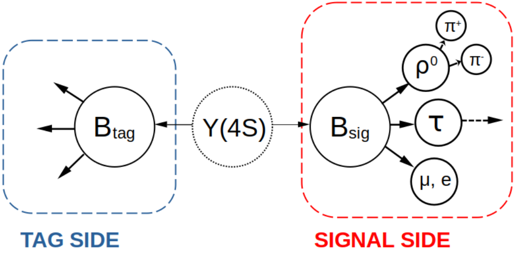
$B \to \rho\tau\ell$ - Analysis strategy

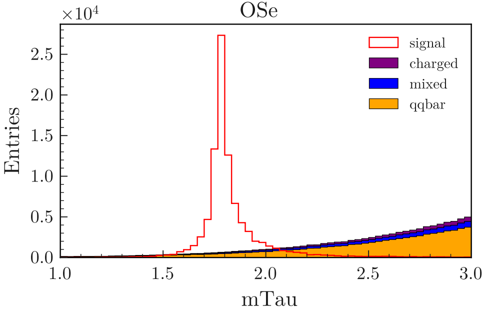
- Hadronic tagging
- The signal is extracted from a fit of the τ recoil mass:

$$\to M_\tau = \sqrt{m_B^2 + m_{\rho\ell}^2 - 2(E^*_{Btag}E^*_{\rho\ell} + p^*_{Btag}p^*_{\rho\ell}cos(\theta^*))} \quad (* : \text{center of mass})$$

- Compute the upper limit on the branching fraction:

$$\to Br^{UL} = \frac{N_{sig}^{UL}}{2 \times \epsilon_{sig} \times N_{B\bar{B}} \times BR_{\rho\to\pi\pi} \times f_{00}}$$





$B \to \rho\tau\ell$ - Selection

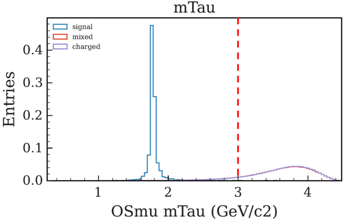
Same cut-based selection and vetos for Belle and Belle II

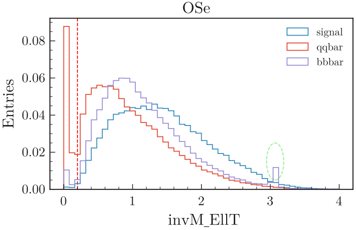
Cut-based selection: $mTau < 3$ GeV, ρ invariant mass ...

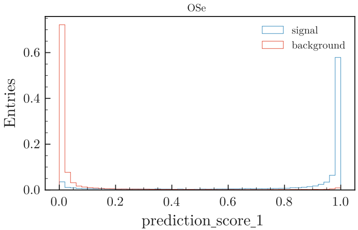
Vetos: J/Ψ veto: $M_{\ell\tau} \notin [3.056; 3.109]$ GeV, $\gamma \to e + e-$ veto: $M_{\ell\tau} > 0.2$ GeV ...

BDT selection: cut optimized with the Punzi figure of merit

Blind region: $M_\tau \in [1.6, 1.95]$GeV $\approx 95\%$ signal coverage







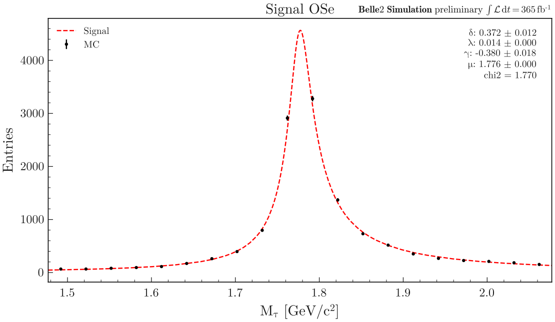
Selection results

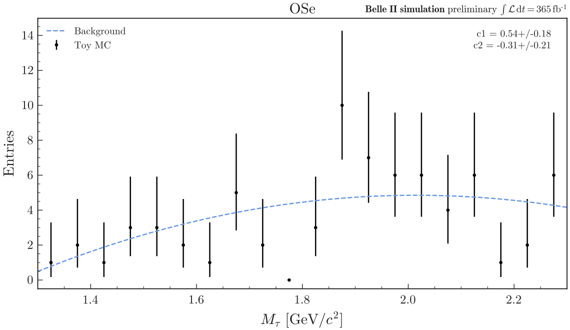
- Best candidates selection: random selection, multiplicity left: ≈ 1.25 for signal, ≈ 1.06 for background, ≈ 1.14 for data in full M_{tau} range

Belle 2	OSe	OSmu	SSe	SSmu
Efficiency before BDT	0.115 %	0.106 %	0.114 %	0.105 %
Efficiency after BDT	0.067 %	0.040 %	0.055 %	0.047 %
Background after BDT (scaled to data luminosity)	78	32	54	51
BDT cut	0.46	0.69	0.59	0.62
Belle				
Efficiency before BDT	0.182 %	0.157 %	0.180 %	0.147 %
Efficiency after BDT	0.051%	0.057 %	0.045%	0.037 %
Background after BDT (scaled to data luminosity)	194	403	164	226
BDT cut	0.83	0.76	0.85	0.86

Simultaneous fit of the τ recoil mass (Belle + Belle II) - I

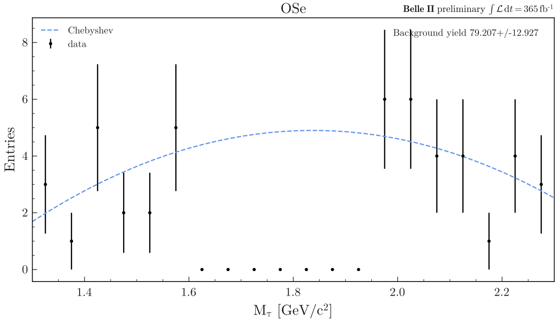
- Simultaneous fit Belle + Belle II
- Signal fit a Johnson PDF
- Background fit with a Chebychev polynomial (2nd order)
- Fit of the data sidebands with the Chebychev to make a toy



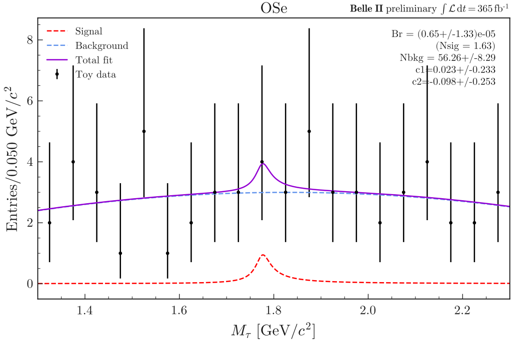


→ Fit of the data sidebands with the Chebychev to make a toy
→ Simultaneous fits of the toy with total PDF

Fit of the data sidebands



Fit of the toy



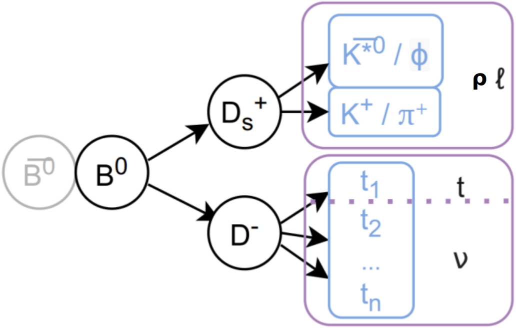
Control sample - $B^0 \to D^- D_s^+ (\to \bar{K}^{*0} K^+ / \phi\pi^+)$

- $\bar{K}^{*0} K^+ / \phi\pi^+$ mimics the $\rho\ell$ system
- D mimics τ: One track reconstructed, the rest of decay account for the missing energy of the τ decay
- Used for BDT and signal shape systematics

Selection:

- Same tag selection as nominal channels
- K^{*0} mass within $\pm 50 MeV/c^2$ window
  ϕ mass within $\pm 20 MeV/c^2$ window
- D_s^+ mass within $\pm 20 MeV/c^2$ window

$\rightarrow$ $B^0 \to D_s^+ D^- (\to K^+ \pi^- \pi^-)$ under study to increase statistics



Stolen from Clotilde Lemettais

Control sample - $B^0 \to D^- D_s^+ (\to \bar{K}^{*0} K^+ / \phi\pi^+)$ - Results

$B^0 \to D^{*-} D_s^+$ is also reconstructed and fitted

$\to$ DDs yields before BDT vs after BDT

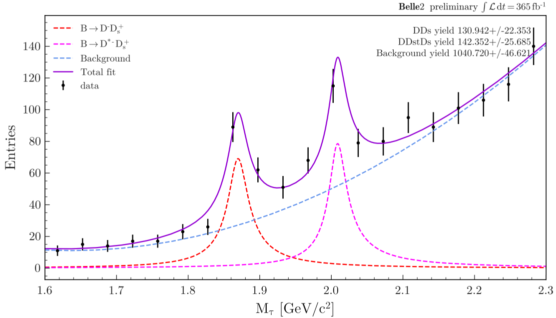
$\to$ Define the BDT cut with $\epsilon_{MC}^{CC} = \epsilon_{MC}^{nominal}$

$\to$ ϵ_{data}, ϵ_{MC} : relative efficiency of the BDT

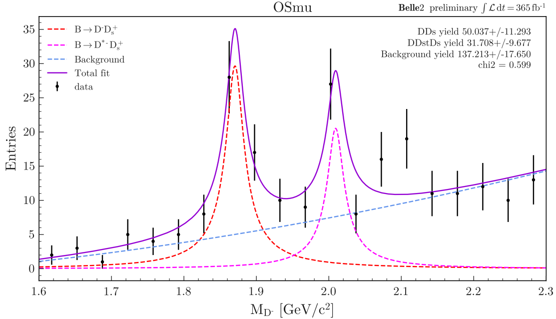
$\to$ $R = \epsilon_{data}^{CC} / \epsilon_{MC}^{CC}$

$\to$ Apply: $max(|1 - R|, \sigma_R)$

before BDT



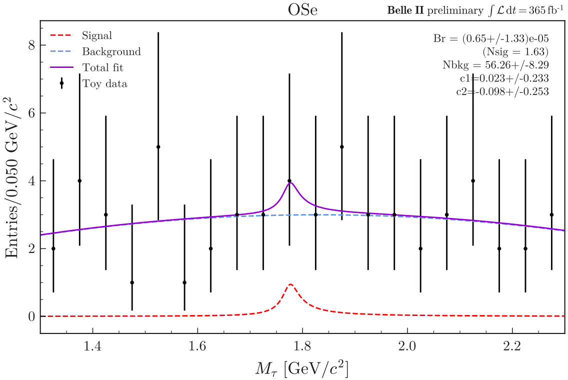
after BDT



$\to$ Systematics on BDT efficiency $\approx$ 30-40% due to low statistics $\to$ $B^0 \to D_s^+ D^- (\to K^+ \pi^- \pi^-)$

Upper limits expected on the branching fractions

Measurement of the expected upper limit on the branching ratio using the CLs asymptotic method in the toy generated from data sidebands



	OSe	OSmu	SSe	SSmu
Br^{UL} expected ($\times 10^{-5}$) - (with systs)	2.4 (2.6)	2.9 (3.2)	2.8 (3.1)	3.4 (3.9)

Conclusion - FEI analysis

- Reconstruction, selection and simultaneous fit procedure ready
- The systematics are highly dominated by the BDT efficiency and the signal shape parameters

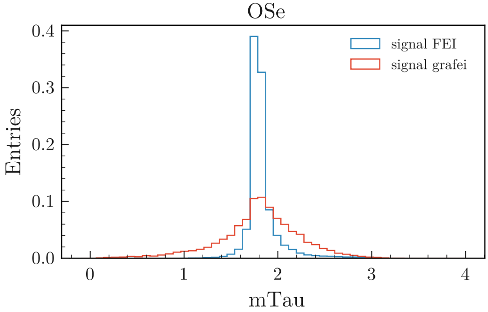
$\rightarrow$ Next steps: New control samples and unblinding

$B \to \rho\tau\ell$ - Alternative reconstruction method: GraFEI

Reconstruction using FEI

- Hadronic tagging
- Reconstruction efficiency : $\approx 0.2\%$

$\to$ The constraint comes from the tag



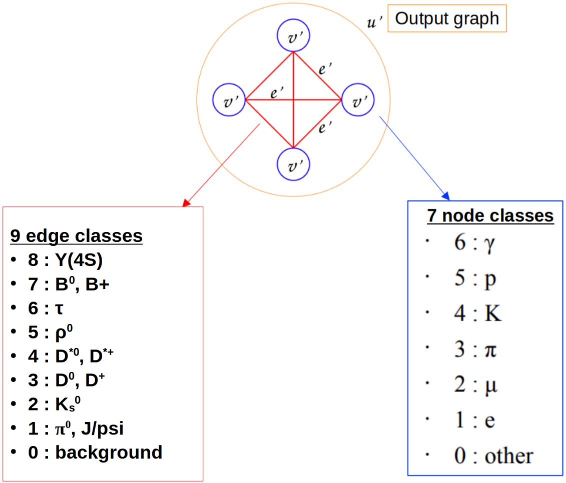
Reconstruction using GraFEI

- **GraFEI** : GNN (graph neural network) trained with MC signal events
- The signal side is reconstructed by the GNN
- The tag side is defined with the rest of particles (inclusive tagging)
- Only on the OSe mode
- Reconstruction efficiency : $\approx 6\%$

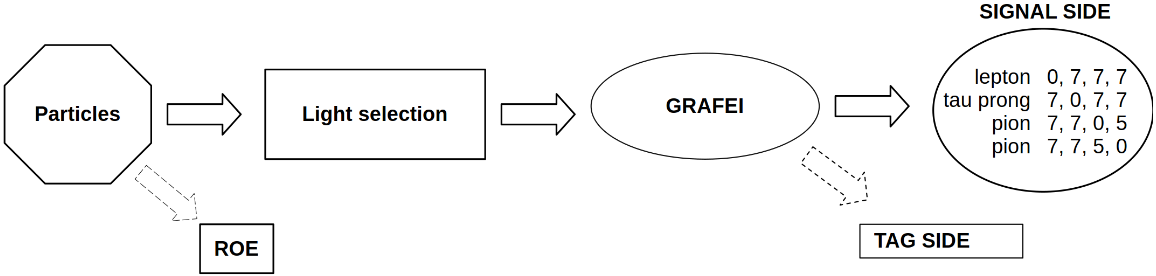
$\to$ The constraint comes from the signal

GraFEI : Output

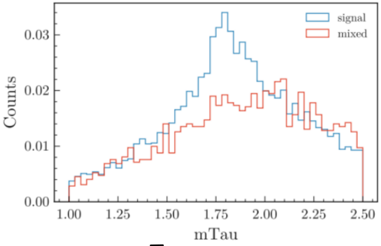
- Nodes = mass hypothesis
- Edges = latest common ancestor
- Nodes and edges are compared to MC truth
- Loss function calculation and GNN parameters update



Grafei - signal reconstruction



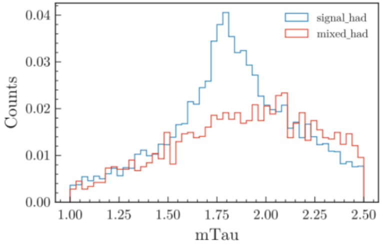
Inclusive tag



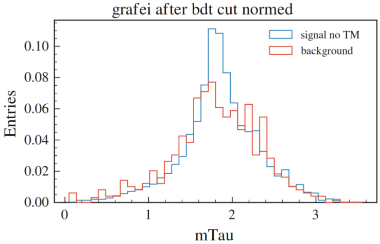
-30 % efficiency

-15 % mixed

« hadronic » tag

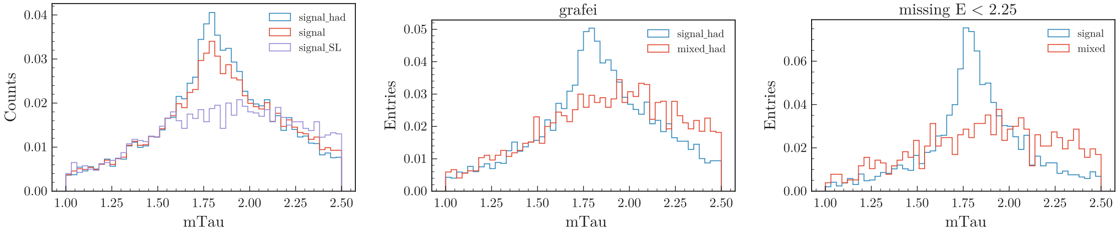


BDT



GraFEI "hadronic" reconstruction

→ Trying to remove semileptonics of "1st level" $B \to \ell X$ from tag side using the LCA matrix
→ Selection on the missing energy



→ BDT output was correlated to the τ mass
→ Using an inclusive reconstruction for this analysis was probably not the best idea

GraFEI new ideas

- Training on generic to reconstruct hadronic decays of B
  - reconstruct the same modes as the FEI ($\approx$10000 modes = time challenging)
  - try an "inclusive hadronic" reconstruction
- Training on generic B decays and add a global target "isSemiLeptonic"
  - Showing better training results overall
- Trying other global targets (momentum of the B, beam-constrained mass, "isSignal"...)

BACK UP

BDT systematics

→ Define the BDT cut with $\epsilon_{MC}^{CC} = \epsilon_{MC}^{nominal}$

→ ϵ_{data}, ϵ_{MC} : relative efficiency of the BDT

→ $R = \epsilon_{data}^{CC} / \epsilon_{MC}^{CC}$

→ Apply: $max(|1 - R|, \sigma_R)$

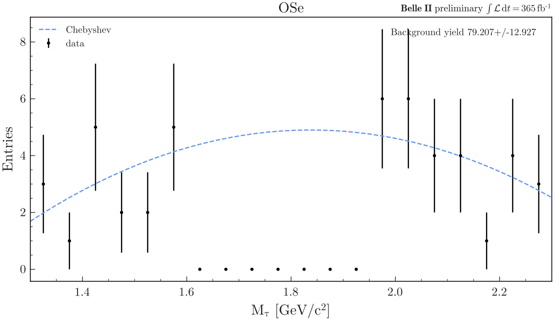
Belle 2	ϵ_{data}^{CC}	ϵ_{MC}^{CC}	R	Systematics on BDT
OSe	$18 \pm 5\%$	$22 \pm 7\%$	0.84 ± 0.36	0.36
OSmu	$25 \pm 6\%$	$22 \pm 8\%$	0.92 ± 0.45	0.45
SSe	$13 \pm 4\%$	$19 \pm 7\%$	0.70 ± 0.36	0.36
SSmu	$14 \pm 4\%$	$17 \pm 4\%$	0.80 ± 0.37	0.37
Belle				
OSe	$10 \pm 3\%$	$12 \pm 4\%$	0.83 ± 0.34	0.34
OSmu	$14 \pm 3\%$	$17 \pm 4\%$	0.84 ± 0.28	0.28
SSe	$11 \pm 3\%$	$12 \pm 4\%$	0.93 ± 0.35	0.35
SSmu	$15 \pm 3\%$	$16 \pm 4\%$	0.92 ± 0.31	0.31

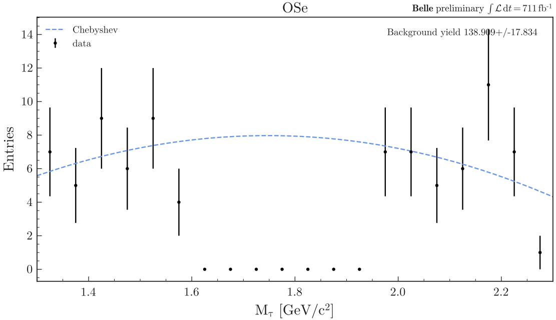
Systematics summary

- Systematics on the efficiency and signal shape parameters are dominant

Belle	OSe	SSe	OSmu	SSmu
LID	2.084%	2.242%	2.184%	2.153%
HadID	1.041%	0.943%	1.045%	1.017%
FEI	4.872%	4.872%	4.872%	4.872%
BDT	33.7%	35.4%	27.8%	30.6%
Total efficiency	33.72 %	35.42 %	27.82 %	30.62 %
Signal shape λ	13 %	13 %	13 %	13 %
Signal shape μ	0.3 %	0.3 %	0.3 %	0.3 %
Belle 2	OSe	SSe	OSmu	SSmu
LID	0.888%	0.947%	0.735%	0.700%
HadID	5.542%	5.712%	5.497%	5.551%
FEI	6.205%	6.123%	6.194%	6.171%
BDT	36.1%	35.9%	44.9%	37.1%
Total efficiency	36.12 %	35.92 %	44.91 %	37.12 %
Signal shape λ	59 %	59 %	59 %	59 %
Signal shape μ	0.2 %	0.2 %	0.2 %	0.2 %
Background shape $\times 10^{-5}$	0.01	0.2	0.1	0.2

Toy from the data sidebands





Samples

Belle 2

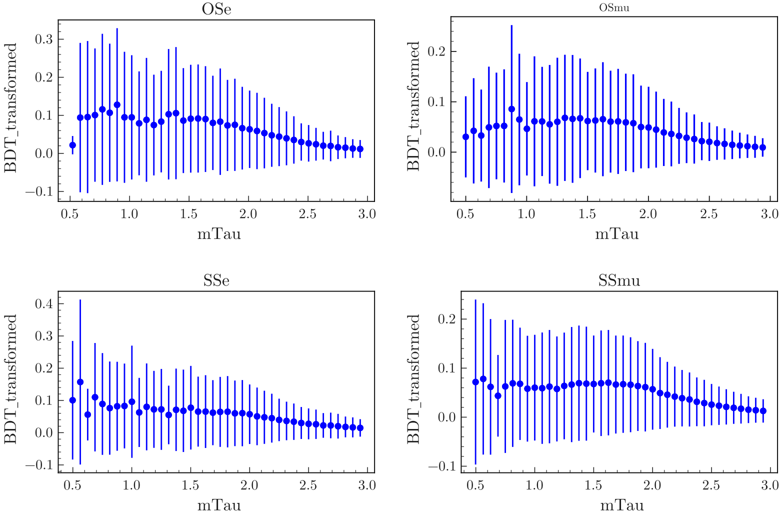
- Signal MC15rd: 20 M events for each mode.
- MC Background: generic MC15rd processed with release-06-01-12 (1444 $\mathrm{fb}^{-1} \approx 4\times$ data luminosity) (Hadronic FEI skim)
- Data: run 1, proc13 (exp 7-18) $+$ prompt (exp 20-26) (364 fb^{-1}) (Hadronic FEI skim)

Belle:

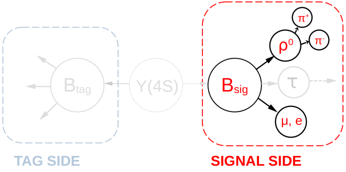
- Signal: 20 M events for each mode
- MC: generic case B MC simulations, 4 x data luminosity
- Data: full dataset (exp 7-65), 711 fb^{-1}

Profile plots - background (Belle II)

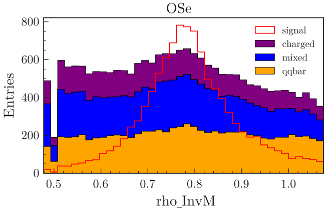
No sculpting in the background



Signal side reconstruction - $\rho\ \ell$

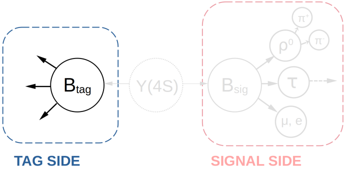


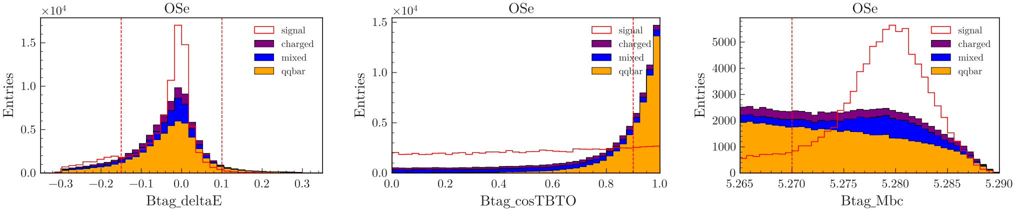
- Reconstruction of the $B \to \rho\ell$ system
- Track selection: $dr < 0.5$ cm and $|dz| < 5$ cm, in CDC acceptance and $nCDCHits > 20$
- Particle identification:
  $\to$ pion: $pionID > 0.6$ and $muonID < 0.9$
  $\to$ electron: $electronID > 0.9$ and $muonID < 0.9$
  $\to$ muon: $muonID > 0.9$
- ρ mass: $\pi\pi$ invariant mass in $\pm 298 MeV$ around ρ nominal mass ($= 775 MeV$)



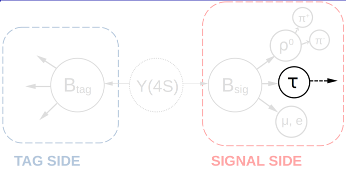
All the track selection was done based on $K^*\tau\ell$ analysis by C.Lemettais

- Hadronic tagging FEI
- FEI skims are used
- Recommended cuts by the FEI group:
  $\rightarrow -0.15 < \Delta E < 0.1$ GeV
  $\rightarrow cosTBTO < 0.9$
  $\rightarrow Mbc > 5.27$ GeV
  $\rightarrow$ SignalProbability $> 10^{-3}$





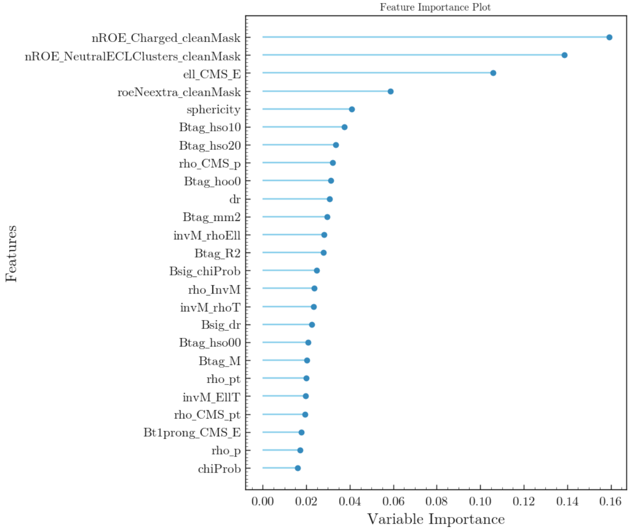
Signal side reconstruction - τ prong



- Reconstruction of one track (τ prong) from the τ for background rejection only.
- Not used in the τ recoil mass calculation.
- $dr < 4$ and $|dz| < 20$ cm
- 4 τ decay modes:
  $\rightarrow$ e: $electronID > 0.9$ and $muonID < 0.9$
  $\rightarrow$ π: $pionID > 0.6$ and $muonID < 0.9$
  $\rightarrow$ μ: $muonID > 0.9$
  $\rightarrow$ $\rho^{\pm} \rightarrow \pi^{\pm}\pi^{0}$ resonance: $pionID > 0.6$
  $|dM| < 0.02$ GeV and π^0 daughter has $|dM| < 0.01$ GeV
  $\rightarrow$ Missing cut: N pi0 in ROE $= 0$

The τ modes are not separated in the analysis.

BDT variable importance

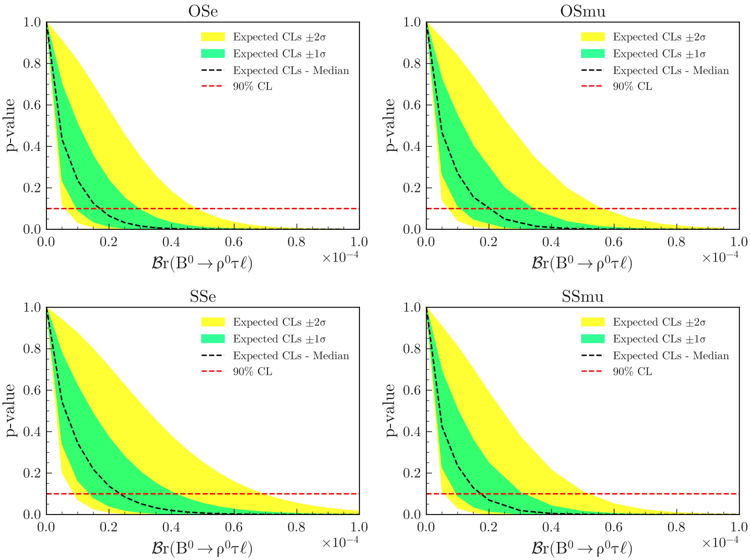


Multiplicity

	OSmu	OSe	SSmu	SSe
Signal	1.22	1.28	1.27	1.22
Background	1.04	1.08	1.06	1.04
Data	1.20	1.16	1.13	1.07

Table 1: Multiplicity of the signal and background events

Asymptotic upper limit

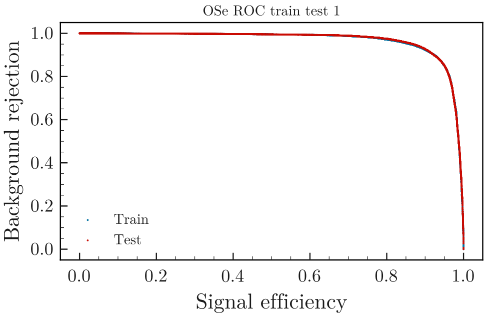


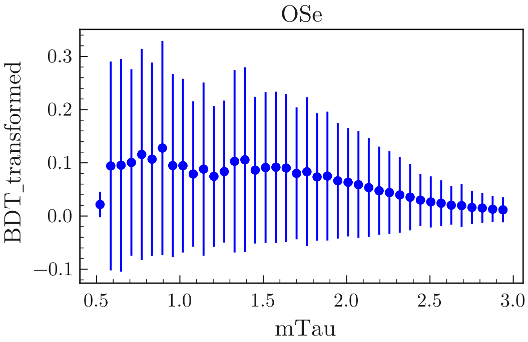
BDT selection

Different BDTs for Belle and Belle II

Optimization of the BDT cut with the Punzi f.o.m $\left(= \frac{\epsilon_{sig}}{3/2+\sqrt{N_{bkg}}}\right)$

Optimizing the cuts with the combined upper limit in MC was also tried

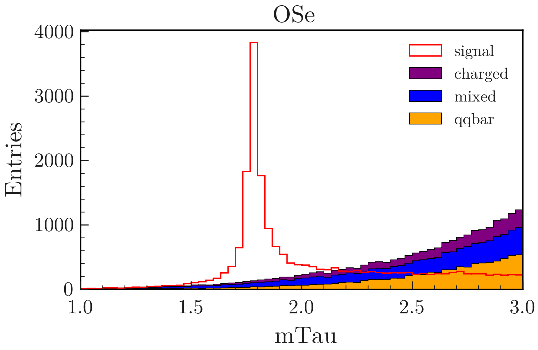




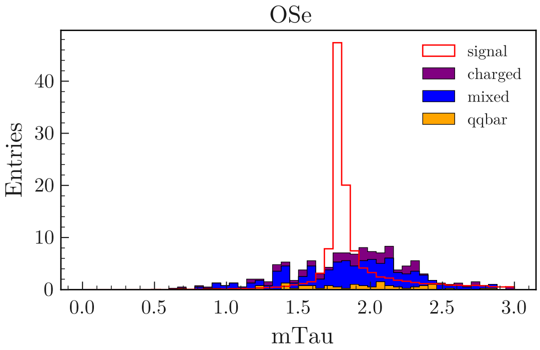
Belle II plots: ROC curve and profile plot in background MC

Tau mass

Tau mass before BDT

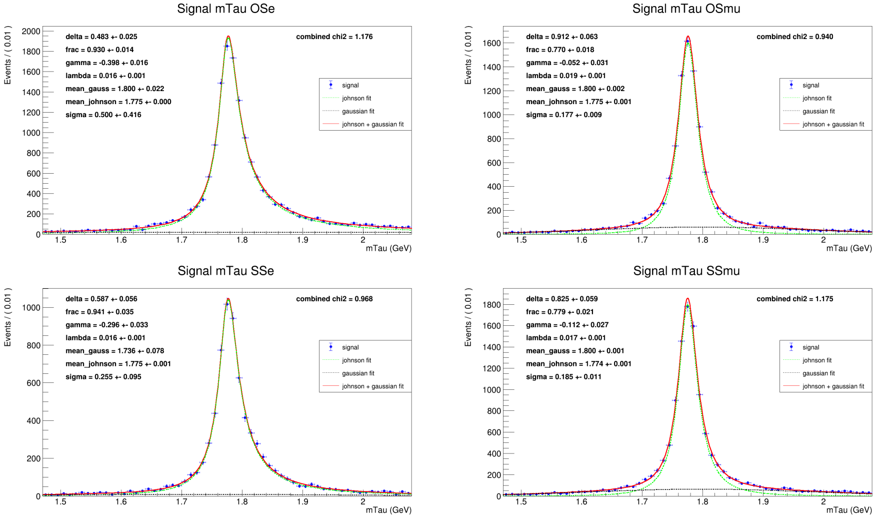


Tau mass after BDT

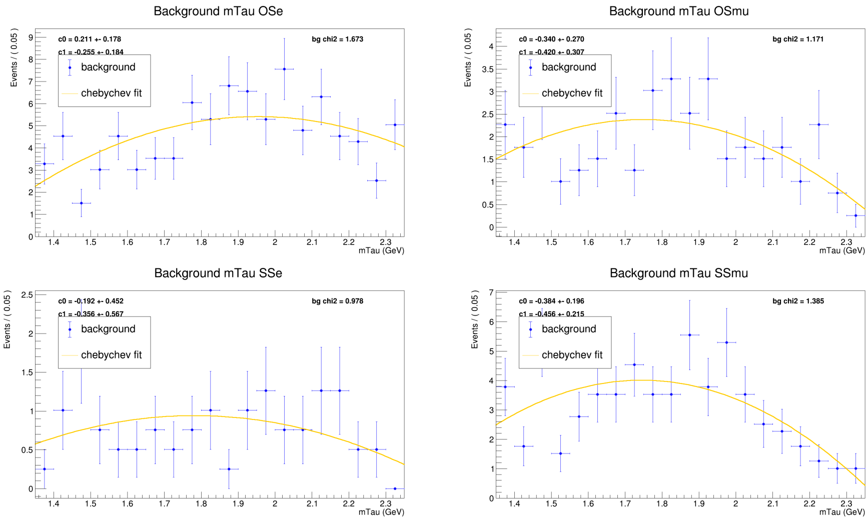


Signal not MC matched

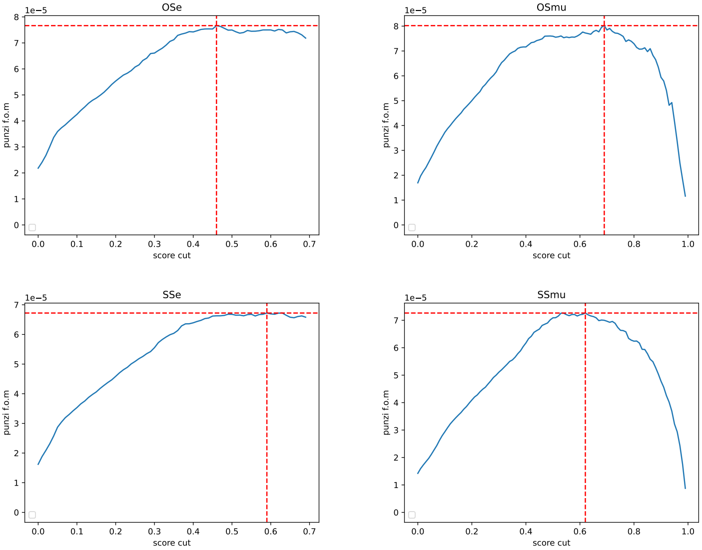
Fit of the τ recoil mass - signal



Fit of the τ recoil mass - bakcground



BDT results - Punzi



Track selection - summary I

Signal side (ρ and lepton):

- $dr < 0.5$ cm and $|dz| < 5$ cm
- particleID ≥ 0.9 (0.6 for pions)
- Bremsstrahlung correction for e
- ρ: $|dM| < 0.298$ GeV

Tag side (FEI) :

- hadronic tag
- B_{tag} Mbc > 5.272 GeV
- cosTBTO < 0.9
- -0.15 GeV $<$ *deltaE* $<$ 0.1 GeV
- SignalProbability $> 10^{-3}$

τ **prong**:

e, μ, π:

- $dr < 4cm$ and $|dz| < 20$ cm
- particleID ≥ 0.9 (0.6 for pions)

π^0:

- $|dM| < 0.02$ GeV
- daughters are goodGammas and have fakePhotonSuppression > 0.3 and beamBackgroundSuppression > 0.3

$\rho+$:

- $|dM| < 0.02$ GeV and π^0 daughter has $|dM| < 0.01$ GeV

Track selection - summary III

Photons:

- GoodGamma:
  clusterReg $= 1$ and clusterE > 0.075 GeV or clusterReg $= 2$ and clusterE > 0.05 GeV or clusterReg $= 3$ and clusterE > 0.1 GeV
- E > 0.05 GeV
- beamBackgroundSuppression > 0.3 and fakePhotonSuppression > 0.3

Good track:

- $dr < 10$ cm and $|dz| < 20$ cm
- theta in CDC Acceptance

Continuum suppression:

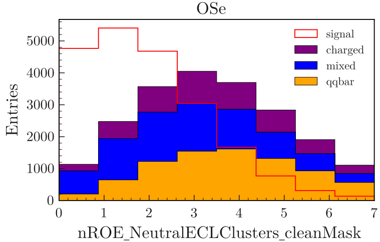
- sphericity > 0.2

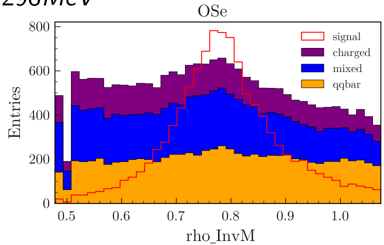
ROE: e,μ, K:

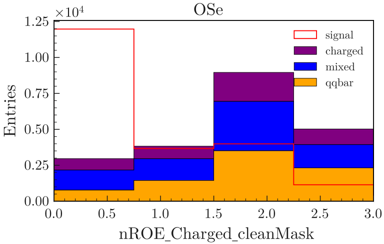
- $dr < 0.5$ cm and $|dz| < 5$ cm
- particleID ≥ 0.9 (0.6 for pions)
- No charged tracks in the ROE

Event level selection

- ROE tracks $dr < 5$ and $|dz| < 5$ cm in CDC acceptance and $nCDCHits > 20$
- ECL clusters, goodGamma, $fakePhotonSuppression > 0.3$, $beamBackgroundSuppression > 0.3$
- Number of charged particles in ROE ≤ 3
- Number of neutrals ECL clusters in ROE ≤ 7
- sphericity > 0.2
- ρ invariant mass in $\pm 298 MeV$ around nominal mass $(= 775 MeV)$

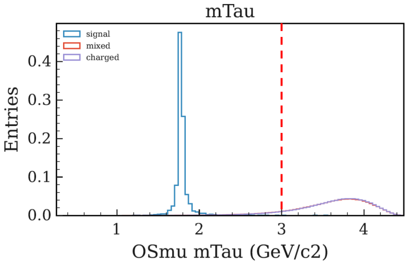


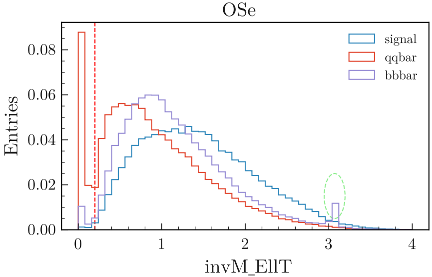


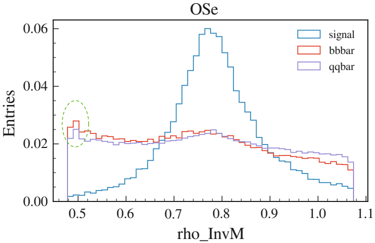


τ mass and vetos

- $M_\tau < 3$ GeV
- Ks veto: $M_\rho \notin [0.490; 0.503]$ GeV
- J/Ψ veto: $M_{\ell\tau} \notin [3.056; 3.109]$ GeV
- electron modes only:
  $\gamma \to e + e-$ veto: $M_{\ell\tau} > 0.2$ GeV

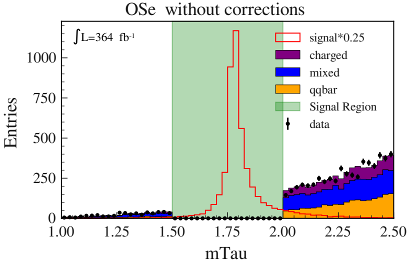


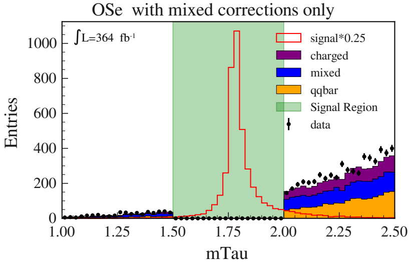




DATA/MC Corrections

PID corrections applied for the two pions, the lepton and the τ prong.
FEI corrections applied on signal and mixed background events

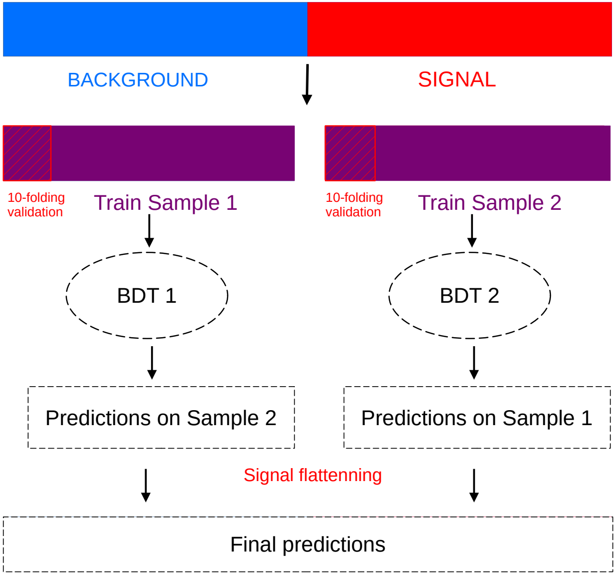




Mode	Data/MC	Data/MC (no corrections)
OSe	1.11±0.01	0.99 ± 0.01
SSe	1.10±0.01	0.96 ± 0.01
OSmu	0.97 ±0.01	0.81 ± 0.01
SSmu	0.98 ±0.01	0.82 ± 0.01

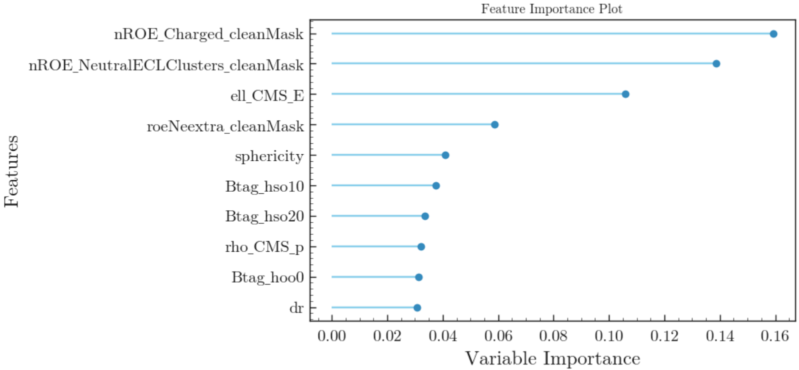
BDT strategy

- Pycaret library + xGboost optimized with Optuna
- 2 BDTs with same signal/background ratio
- Train samples: $\approx$ 25 000 signal events and $\approx$ 130 000 background events
- signal flattenning (variable change) to assure performances of both BDTs are the same
- Cut optimization with Punzi f.o.m



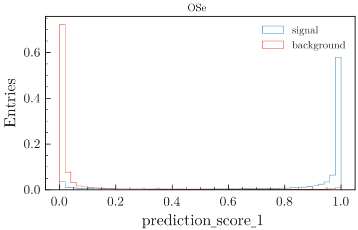
BDT variables

BDT variables: number of charged tracks, neutral ECL clusters in ROE, sphericity, tracks variables, invariant masses (ρ, $\rho\tau - prong$, $\ell\tau - prong$, $\rho\ell$)... (full list in back-up)

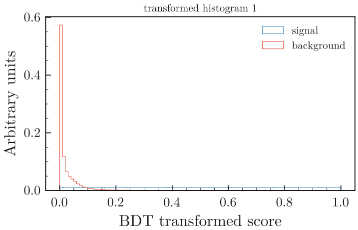


BDT results - I

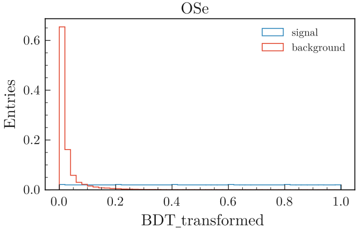
BDT 1



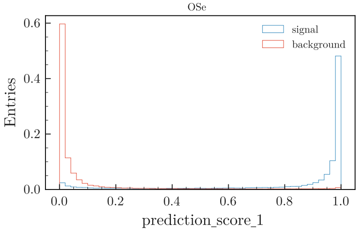
FLATTEN SIGNAL BDT 1



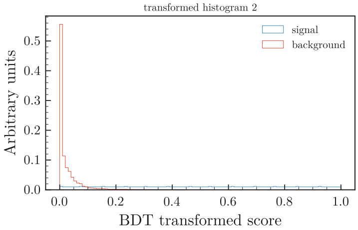
FINAL OUTPUT



BDT 2

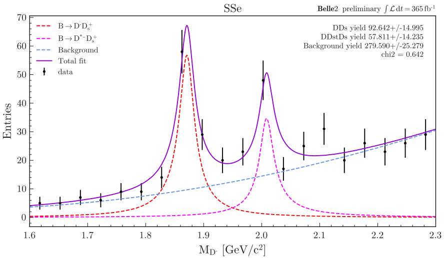


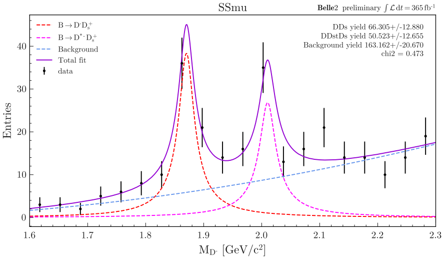
FLATTEN SIGNAL BDT 2



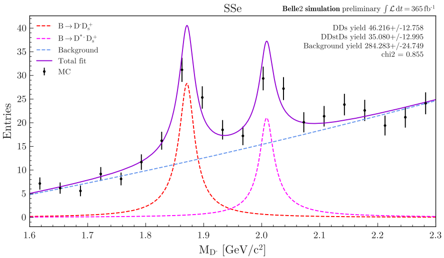
after BDT SSe and SSmu

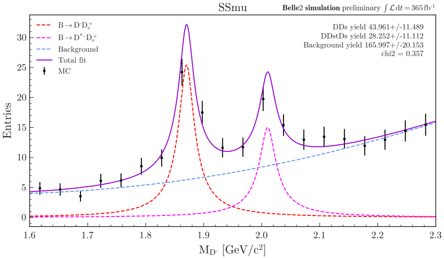
Data



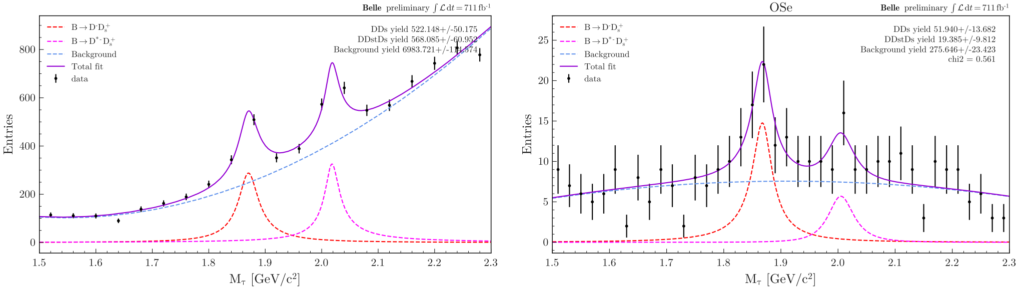


MC

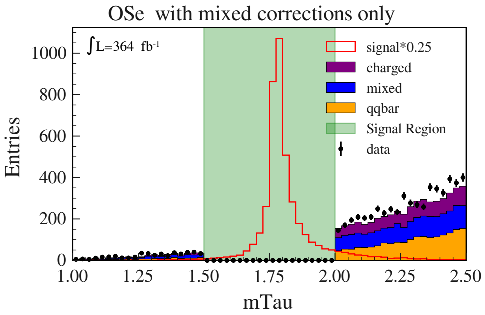


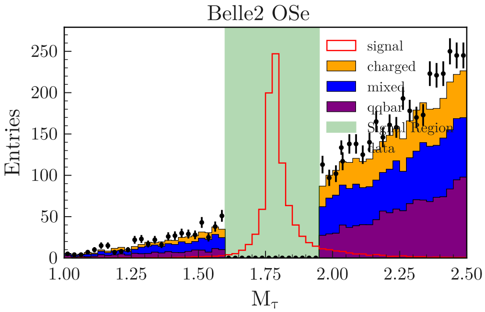


Belle control samples



DATA/MC after BDT - Belle II





	DATA/MC before BDT	DATA/MC after BDT
OSe	1.11 ± 0.01	1.23 ± 0.17
OSmu	0.97 ± 0.01	1.19 ± 0.25
SSe	1.10 ± 0.01	1.18 ± 0.30
SSmu	0.98 ± 0.01	1.23 ± 0.21

Grafei analysis - Track selection

Charged tracks:

- $dr < 0.5$ cm and $|dz| < 5$ cm
- particleID most likely ¿ 0
- $p_t > 0.2$ GeV and $p < 5$ GeV
- *theta* in CDC acceptance
- $N_{CDCHits} > 20$

Photons:

- $beamBackgroundSuppression > 0.4$
- $fakePhotonSuppression > 0.3$
- $|clusterTiming| < 100$
- $\left|\frac{clusterTiming}{clusterErrorTiming}\right| < 2.0$
- ($clusterReg = 1$ and $E > 0.09$ GeV) or (clusterReg $= 2$ and $E > 0.09$ GeV) or (clusterReg $= 3$ and $E > 0.14$ GeV)

ROE number of tracks - before and after BDT