Search for $B^0 \to \tau^+\tau^-$ in Belle and Belle II

GDR workshop 13/11/2025

C. Kim	Hanyang
G. Mancinelli	CPPM
M. Marfoli	CPPM-TYL
S. Park	KEK
V. S. Vobbilisetti	INFN Pisa
R. Tiwary	KEK-TYL
K. Trabelsi	CPPM-TYL

Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France

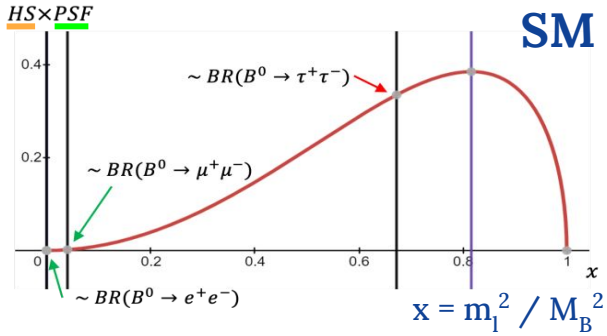
Motivations

- Expected Standard Model (SM) Branching Fraction (BF):

$B^0_d \to \tau^+\tau^-$ ~ O(10^{-8}) [1]

$$\mathcal{B}(B^0 \to \ell^+\ell^-) = \frac{G_F^4 M_W^4 M_B^3}{8\pi^5 \Gamma_B} \cdot \underbrace{f_B^2}_{\text{Decay constant}} \cdot \underbrace{|V_{tb}^* V_{td}|^2}_{\text{CKM elements}} \cdot \underbrace{\frac{4m_\ell^2}{M_B^2}}_{\text{Helicity suppression}} \cdot \underbrace{\sqrt{1-\frac{4m_\ell^2}{M_B^2}}}_{\text{Phase space factor}} \cdot |C_A(\mu)|^2$$

- Highest BF among $B^0 \to l^+l^-$:
  - helicity suppression
    - m(**e**): O(1MeV)
    - m(**μ**): O(100MeV)
    - m(***τ***): O(1000MeV)



- Observation of any signal with our current dataset would mean New Physics (NP)
- Important to follow up after searches in b→sll/**νν** to characterize any possible NP effect
- On experimental side:
  - 2***τ*** = 2 to 4 neutrinos in the final state

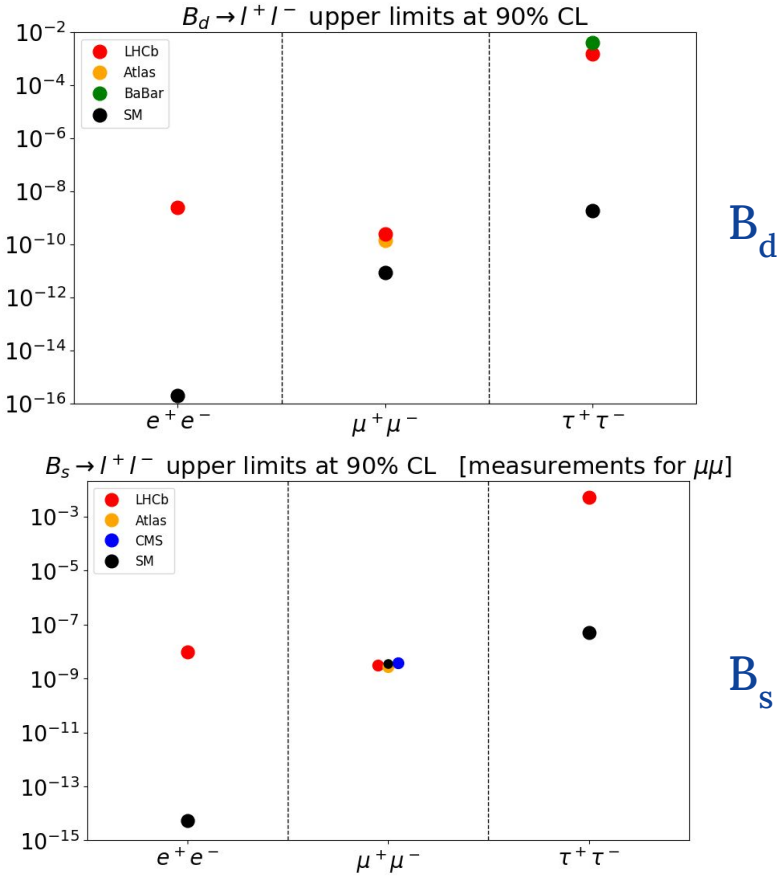
[1] Phys.Rev.Lett. 112 (2014) 2311.00020

Phys.Rev.D 93 (2016) Phys.Rev.D 86 (2012)

Experimental status

- Focus on $B^0_d \to \tau^+\tau^-$:
  - BaBar: [Phys.Rev.Lett. 96 (2006)]
    - 210 fb^{-1} (not full dataset)
    - $B^0_d \to \tau^+\tau^- < 4.1 \times 10^{-3}$ (@90% CL)
    - $\tau \to l\nu\nu$, $\pi\nu$, $\varrho\nu$ combinations
  - LHCb: [Phys.Rev.Lett. 118 (2017)]
    - 3 fb^{-1} (Run1, up to 2012)
    - $B^0_d \to \tau^+\tau^- < 1.6 \times 10^{-3}$ (@90% CL)
    - $\tau \to \pi\pi\pi\nu$

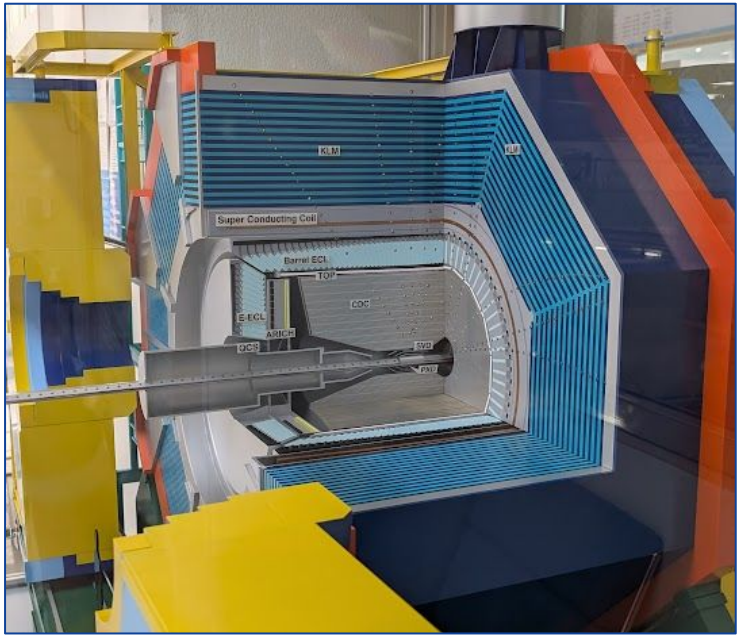
No results from Belle or Belle II yet!



The Belle (II) detector

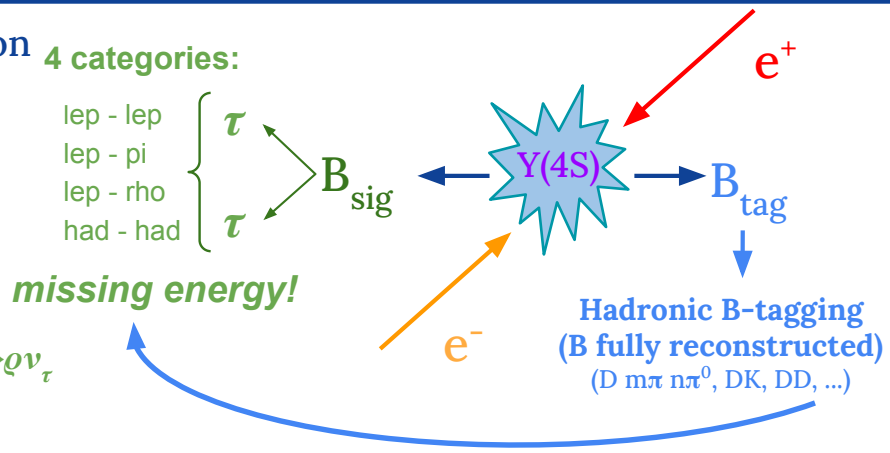
- SuperKEKB e^+e^- collider
  - nominal operation at Y(4s) resonance (~10.58 GeV)
  - world record luminosity: 5.1×10^{34} cm^{-2} s^{-1}
  - $>1 \times 10^9$ $B\bar{B}$ pairs per ab^{-1}
- Belle (II)
  - Both detectors are very similar
    - Belle II based on Belle with upgraded sub-detectors and new technologies
  - Perfect environment for missing energy searches
    - Almost hermetic detector
    - Well known initial state
    - Powerful PID
    - Excellent vertex resolution and tracking
  - Belle: 711 fb^{-1} @ Y(4s) resonance
  - BelleII: 365 fb^{-1} (Run1) + 130fb^{-1} (Run2) @ Y(4s) resonance

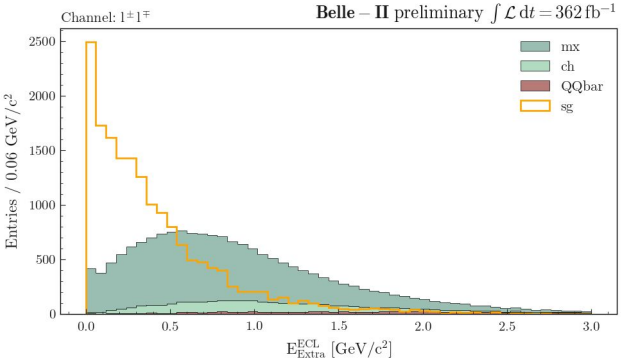
BelleII down-scaled version at the entrance of the Tsukuba Hall at KEK (Japan) (where the Interaction Point (IP) is located)



Analysis strategy

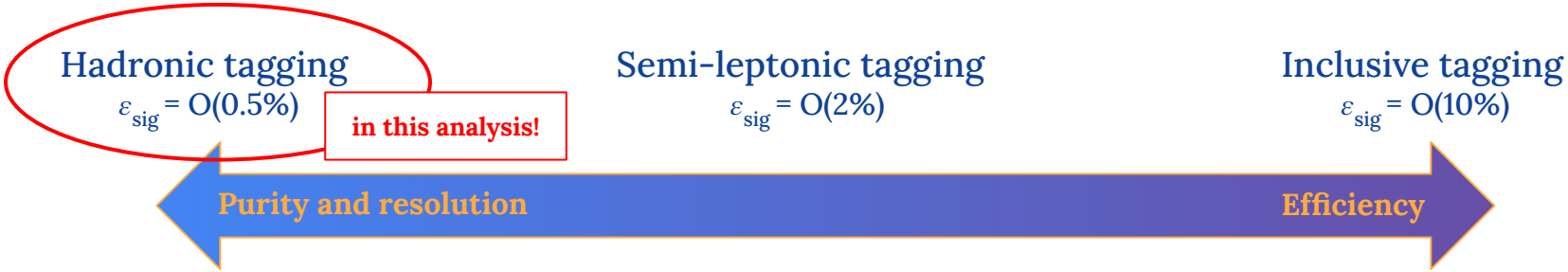
- Hadronic B-tagging with the Full Event Interpretation (FEI) algorithm
- Datasets:
  - Belle: 711 fb^{-1}
  - Belle II: Run1 (362 fb^{-1}) and Run2 (~130 fb^{-1})
- 1-prong τ decays:
  - Combination of: $\tau \to e \nu_e \nu_\tau$, $\tau \to \mu \nu_\mu \nu_\tau$, $\tau \to \pi \nu_\tau$ and $\tau \to \varrho \nu_\tau$
- BDT to suppress general/continuum background
- Use KLM and ECL clusters to suppress main background coming from missing K^0_L
- Fit in the residual energy in the Electromagnetic CaLorimeter (ECL)
  - Signal peaks at 0 since both sides are fully reconstructed (except neutrinos)
  - Background has a longer tail due to additional missing particles





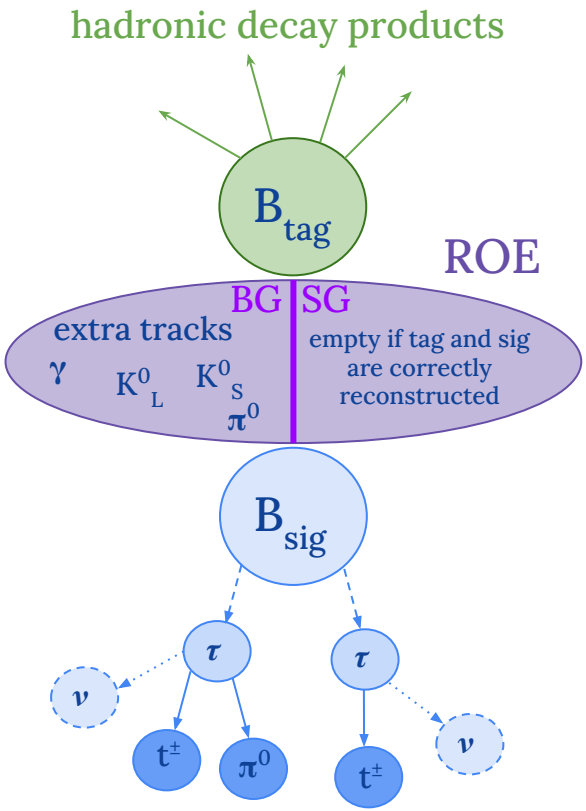
B tagging in Belle and Belle II

- There are 3 main B-tagging techniques used for Belle (II) analyses
  - Hadronic tagging — perfect kinematic constraints
  - Semi-leptonic tagging — good kinematic constraints
  - Inclusive tagging — only beam constraints available
- For both hadronic and semi-leptonic tagging the tag side is reconstructed firstly
  - Skims are usually centrally produced
  - Analysts can load directly prefilled lists with the Btag already reconstructed in different decay modes
  - Signal reconstruction happens in the Rest Of Event (ROE) [= Event - tag side particles]
- For the inclusive tagging the signal side is reconstructed firstly
  - multiple BDTs in series to skim the dataset and extract an high purity signal region to fit
  - Event properties to reduce background and multiple control channels to validate simulation



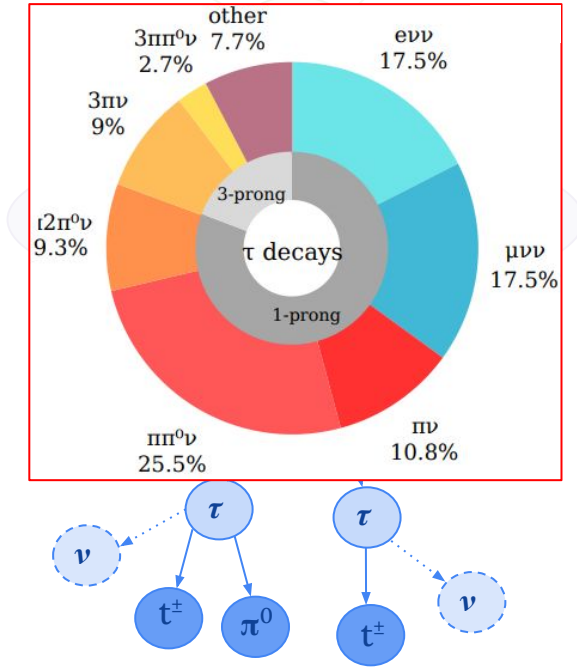
Signal selection

- Given the short lifetime of the two τ's we need to reconstruct them through their decay products
  - We target (some) 1-prong τ decays (~75% of τ decays)
  - We require two tracks that approach the interaction point
  - Additional PID selection to categorize the track for background rejection
    - lepton - lepton
    - lepton - pion
    - lepton - rho
    - hadron - hadron
  - The π^0 reconstruction has been optimized on the Upper Limit (UL) from MC
  - ROE mask (=selection) for the ECL neutral clusters still to be chosen
    - This mask determines whether or not a neutral ECL cluster will be counted towards the 'Extra energy in the ECL' definition (EECL)
  - Remaining cleanup of the ROE (no extra good tracks, K_S^0 ...)



Signal selection

- Given the short lifetime of the two τ's we need to reconstruct them through their decay products
  - We target (some) 1-prong τ decays (~75% of τ decays)
  - We require two tracks that approach the interaction point
  - Additional PID selection to categorize the track for background rejection
    - lepton - lepton
    - lepton - pion
    - lepton - rho
    - hadron - hadron
  - The π^0 reconstruction has been optimized on the Upper Limit (UL) from MC
  - ROE mask (=selection) for the ECL neutral clusters still to be chosen
    - This mask determines whether or not a neutral ECL cluster will be counted towards the 'Extra energy in the ECL' definition (EECL)
  - Remaining cleanup of the ROE (no extra good tracks, K_S^0 ...)



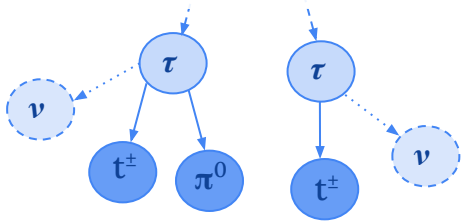
Signal selection

- Given the short lifetime of the two τ's we need to reconstruct them through their decay products
  - We target (some) 1-prong τ decays (~75% of τ decays)
  - We require two tracks that approach the interaction point
  - Additional PID selection to categorize the track for background rejection
    - lepton - lepton
    - lepton - pion
    - lepton - rho
    - hadron - hadron
  - The π^0 reconstruction has been optimized on the Upper Limit (UL) from MC
  - ROE mask (=selection) for the ECL neutral clusters still to be chosen
    - This mask determines whether or not a neutral ECL cluster will be counted towards the 'Extra energy in the ECL' definition (EECL)
  - Remaining cleanup of the ROE (no extra good tracks, K_S^0 ...)

Interactions parameters:
dr (transverse distance)
dz (beam direction)

CDC acceptance:
$17° < \theta < 150°$

PID requirements:
tight selection on e, μ and π ID



Signal selection

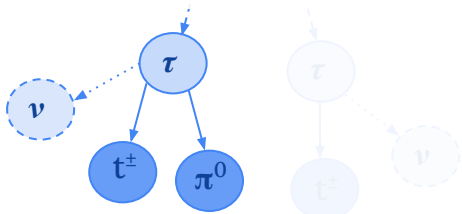
- Given the short lifetime of the two τ's we need to reconstruct them through their decay products
  - We target (some) 1-prong τ decays (~75% of τ decays)
  - We require two tracks that approach the interaction point
  - Additional PID selection to categorize the track for background rejection
    - lepton - lepton
    - lepton - pion
    - lepton - rho
    - hadron - hadron
  - The π^0 reconstruction has been optimized on the Upper Limit (UL) from MC
  - ROE mask (=selection) for the ECL neutral clusters still to be chosen
    - This mask determines whether or not a neutral ECL cluster will be counted towards the 'Extra energy in the ECL' definition (EECL)
  - Remaining cleanup of the ROE (no extra good tracks, K_S^0 ...)

π^0 are reconstructed through decay products of the ϱ ($\rightarrow\pi\pi^0$)

We can do multiple selections here:

- Invariant mass of $\pi\pi^0$
- Selection on photons coming from the π^0 ($\rightarrow\gamma\gamma$)

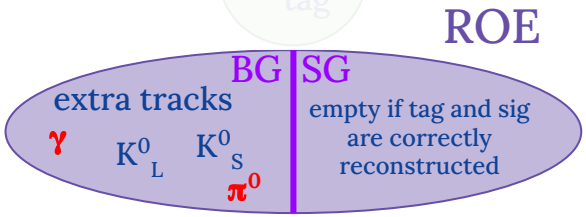
Especially for this last point it can be highly correlated with activity in the EM calorimeter..



Signal selection

- Given the short lifetime of the two τ's we need to reconstruct them through their decay products
  - We target (some) 1-prong τ decays (~75% of τ decays)
  - We require two tracks that approach the interaction point
  - Additional PID selection to categorize the track for background rejection
    - lepton - lepton
    - lepton - pion
    - lepton - rho
    - hadron - hadron
  - The π^0 reconstruction has been optimized on the Upper Limit (UL) from MC
  - ROE mask (=selection) for the ECL neutral clusters still to be chosen
    - This mask determines whether or not a neutral ECL cluster will be counted towards the 'Extra energy in the ECL' definition (EECL)
  - Remaining cleanup of the ROE (no extra good tracks, K_S^0 ...)

.. and it can happen that **γ's** or **π^0's** will end up with in the "ROE" and count towards the ECL total activity

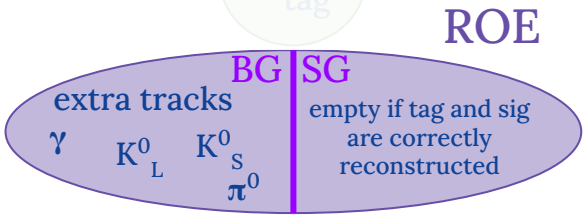


So for this reason we decided to optimize their selection on the final UL from MC (varying the selections among the ones I showed before)

Signal selection

- Given the short lifetime of the two τ's we need to reconstruct them through their decay products
  - We target (some) 1-prong τ decays (~75% of τ decays)
  - We require two tracks that approach the interaction point
  - Additional PID selection to categorize the track for background rejection
    - lepton - lepton
    - lepton - pion
    - lepton - rho
    - hadron - hadron
  - The π^0 reconstruction has been optimized on the Upper Limit (UL) from MC
  - ROE mask (=selection) for the ECL neutral clusters still to be chosen
    - This mask determines whether or not a neutral ECL cluster will be counted towards the 'Extra energy in the ECL' definition (EECL)
  - Remaining cleanup of the ROE (no extra good tracks, K_S^0 ...)

EECL definition:
"sum of all clusters from ECL not used in the reconstruction (ie ending up in the ROE)"

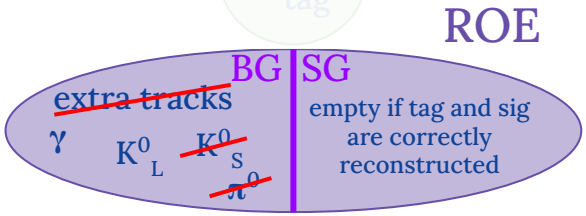


By using a mask here we select which clusters count towards the EECL distribution

More on this later!

Signal selection

- Given the short lifetime of the two τ's we need to reconstruct them through their decay products
  - We target (some) 1-prong τ decays (~75% of τ decays)
  - We require two tracks that approach the interaction point
  - Additional PID selection to categorize the track for background rejection
    - lepton - lepton
    - lepton - pion
    - lepton - rho
    - hadron - hadron
  - The π^0 reconstruction has been optimized on the Upper Limit (UL) from MC
  - ROE mask (=selection) for the ECL neutral clusters still to be chosen
    - This mask determines whether or not a neutral ECL cluster will be counted towards the 'Extra energy in the ECL' definition (EECL)
  - Remaining cleanup of the ROE (no extra good tracks, K_S^0 ...)

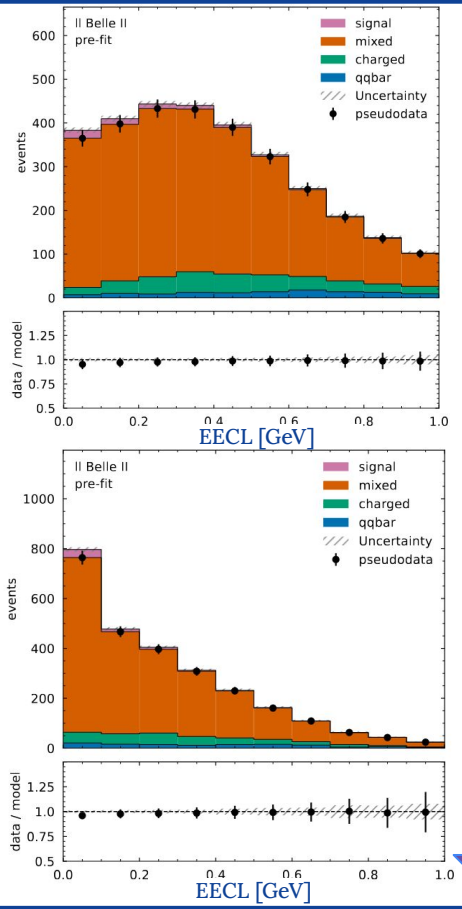


γ's are selected with the ROE mask and
K^0_L requires different treatment (main background)

Extra ECL energy sources and ROE masks

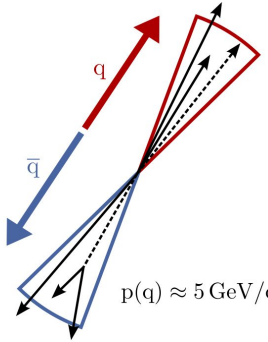
- ECL sources
  - Mis-reconstructed photons
    - ..from charged hadronic showers
    - ..from charged particles with high transverse momentum
    - ..from neutrons or K_L^0 → (dedicated study for K_L^0)

    distance between cluster and track helps reducing this component!
  - Beam induced photons
    - ..from beam particles showering due to Coulomb scattering or Bremsstrahlung and interacting with accelerator or residual gas molecule in the beam → can be reduced with timing difference between ECL cluster and event
  - Properly reconstructed photons
    - ..from real gammas coming from the e^+e^- cascade interaction but not associated to the Y(4s) (mis-reconstructed event)
- ROE masks are used to suppress some of these contributions in the EECL
  - looser masks lead to mis-modeling
    - beam gammas are not simulated but overlaid in MC from data
    - fake photons simulations are also not well modeled
  - tighter masks lead to less signal-background separation power
    - Both signal and background are shifted towards 0 in the EECL

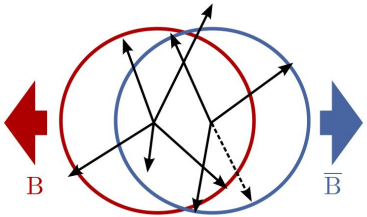


Background characterization at Belle (II)

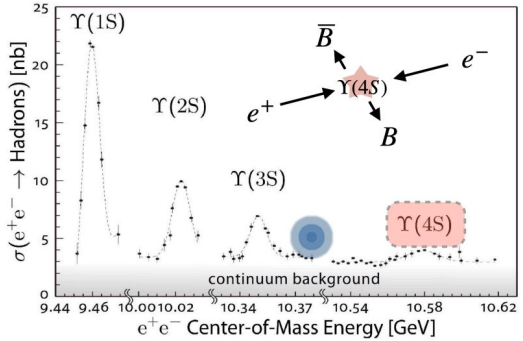
- Background for B analyses comes from 2 main sources
  - continuum (or qqbar): udsc-bar from e^+e^-
  - BB sample (mixed B^0B^0 or charged B^+B^-)
- Other sources are also simulated but they can be easily rejected because of triggers + very different event topology
  - low multiplicity events
  - $e^+e^- \to \tau^+\tau^-$



$e^+e^- \to q\bar{q}$ $(q \in \{u, d, s, c\})$



$e^+e^- \to \Upsilon(4S) \to B\bar{B}$



[Belle II physics book]

Physics process	Cross section [nb]
$\Upsilon(4S)$	1.110 ± 0.008
$u\bar{u}(\gamma)$	1.61
$d\bar{d}(\gamma)$	0.40
$s\bar{s}(\gamma)$	0.38
$c\bar{c}(\gamma)$	1.30
$e^+e^-(\gamma)$	300 ± 3 (MC stat.)
$e^+e^-(\gamma)$	74.4
$\gamma\gamma(\gamma)$	4.99 ± 0.05 (MC stat.)
$\gamma\gamma(\gamma)$	3.30
$\mu^+\mu^-(\gamma)$	1.148
$\mu^+\mu^-(\gamma)$	0.831
$\mu^+\mu^-\gamma(\gamma)$	0.242
$\tau^+\tau^-(\gamma)$	0.919
$\nu\bar{\nu}(\gamma)$	0.25×10^{-3}
$e^+e^-e^+e^-$	39.7 ± 0.1 (MC stat.)
$e^+e^-\mu^+\mu^-$	18.9 ± 0.1 (MC stat.)

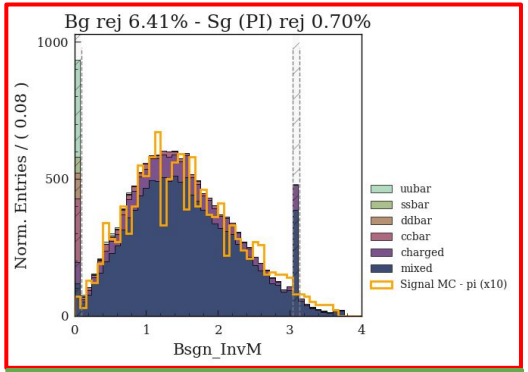
$E_{HER} = 7.004\,\text{GeV}$ and $E_{LER} = 4.002\,\text{GeV}$.

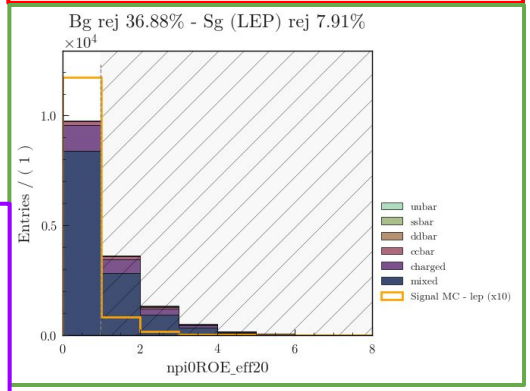
Background suppression

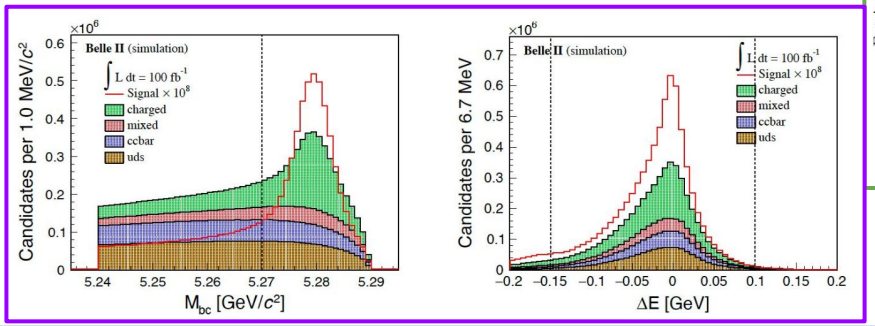
- There are 3 main steps of background suppression:
  - preselection
    - Recommendations from performance group
    - retain high signal efficiency while trying to remove obvious background
      - Invariant mass cut for ee (J/ψ and photon conversion)
    - selections dictated by physics
      - no extra good tracks in the ROE
      - no K_S^0, Λ, π^0 in the ROE

      } decreases also tag mis-reconstruction
  - Boosted Decision Tree (BDT)
    - targeting general background (both qqbar and BB)
  - Main background treatment
    - mainly K_L^0

next slides



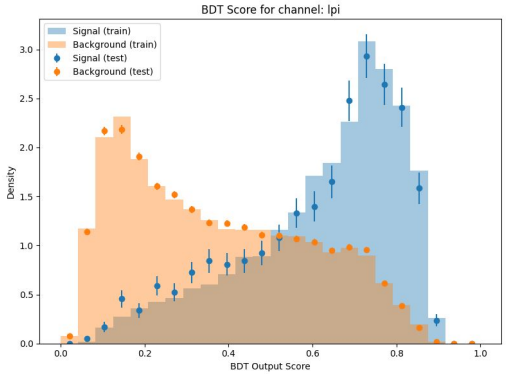


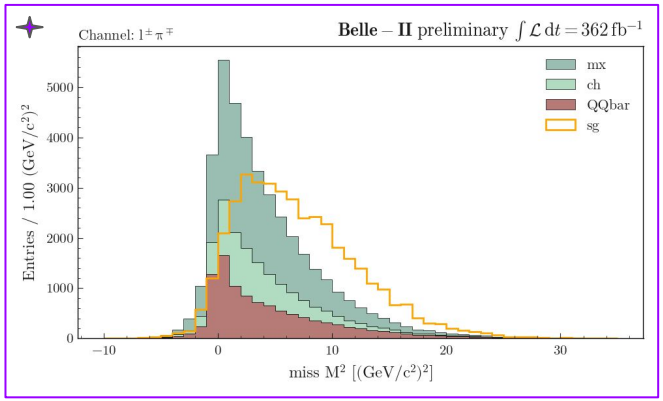


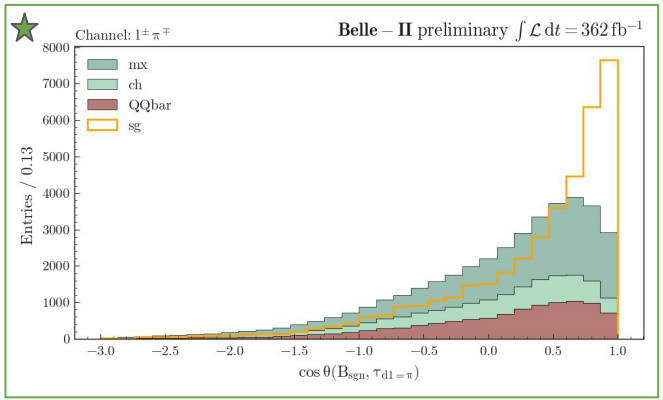
BDT

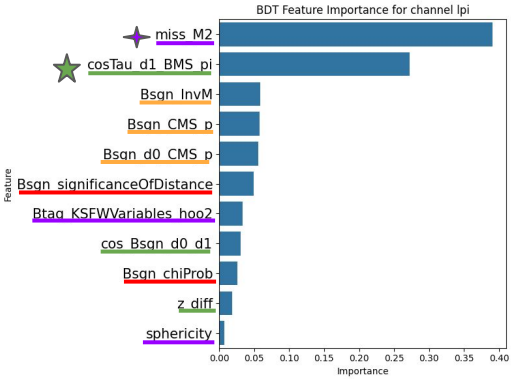
- We will train 4 x 2 BDTs (categories x Belle/BelleII) targeting general background
  - The 4 categories are chosen based on similarities of the input distributions
  - Input features target different background types
    - Continuum suppression
    - Topology
    - Vertexing
    - Kinematic

most sensitive channel: lepton-pion



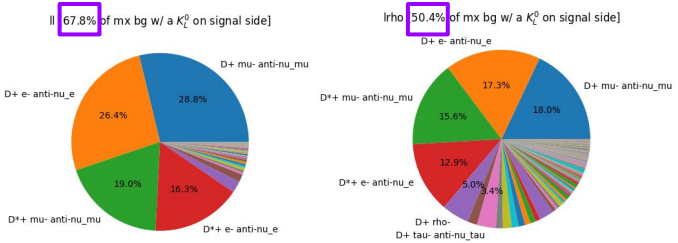




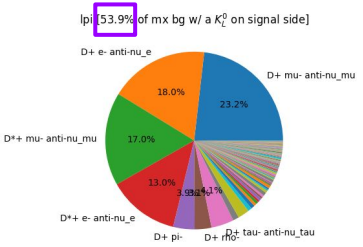


Main background

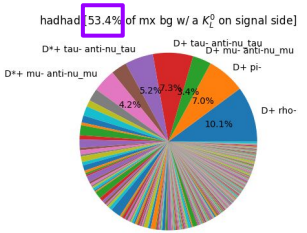
- Main background source is coming from B had/SL decay followed by a D had/SL decay containing a K_L^0
  - for lepton-lepton: $B \to D\,[\to K^0\,(\to K_L^0)\,l\,\nu\,]\,l\,\nu$
  - for lepton-pi: $B \to D\,[\to K^0\,(\to K_L^0)\,\pi\,]\,l\,\nu$
  - for lepton-rho: $B \to D\,[\to K^0\,(\to K_L^0)\,\varrho\,]\,l\,\nu$
  - for hadron-hadron: $B \to D\,[\to K^0\,(\to K_L^0)\,h\,]\,h$



- It consists between 50% to 70% of the mixed (B^0B^0) background
  - When the K_L^0 is missed, it mimics completely the signal
  - Difficult to detect the missing K_L^0 from missing mass in the event due to neutrinos

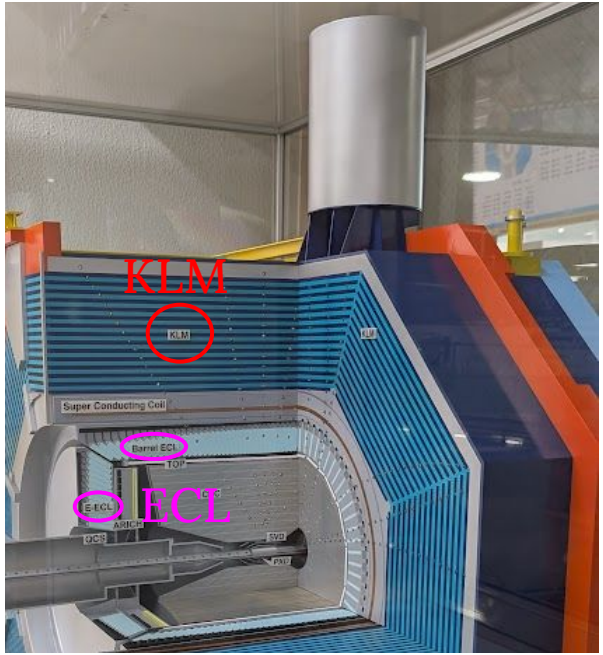


- Thorough validation is needed
  - We can exploit the productions of K_S^0 in our main background to control the K_L^0 contributions
    - Same background as above but with $K^0 \to K_S^0$
    - We can revert the cut on the extra good tracks in the ROE and the number of K_S^0



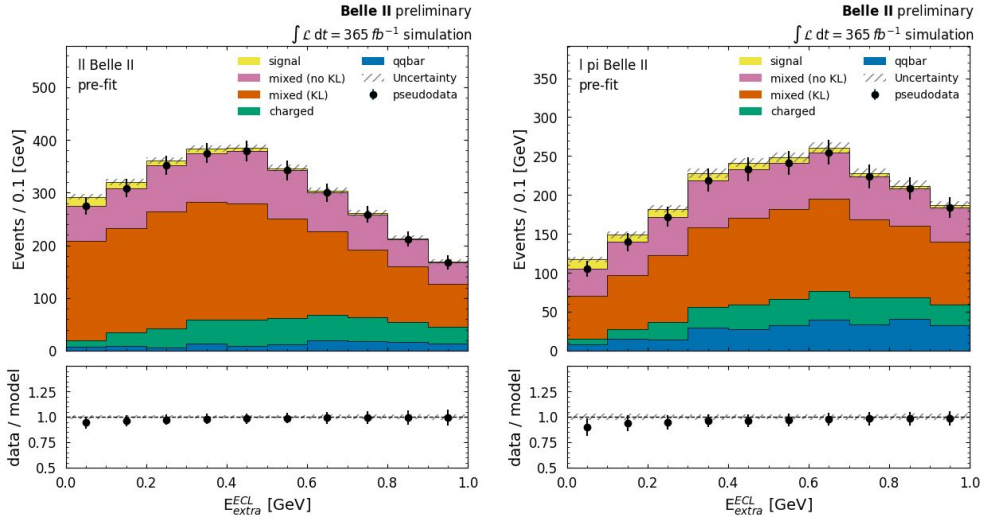
The ever-st(r)anding task of dealing with K_L^0's in Belle (II)

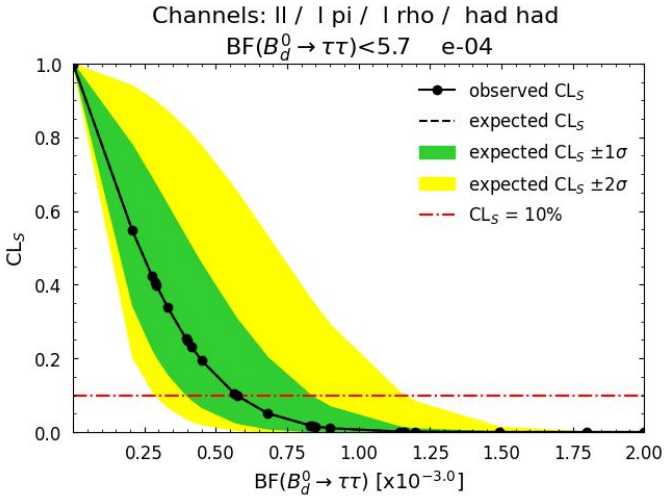
- K_L^0 interacts with two sub-detectors in Belle (II):
  - **KLM** **KL**ong and **M**uon detector
  - **ECL** **E**lectromagnetic **CaL**orimeter
- The interaction with either of the two depends on the momentum of the particle
  - It is possible to have a K_L^0 reconstructed from either of the two sub-detectors…
  - ..or both of them
    - in this case a matching based on the angle between the ECL and KLM cluster is needed
- We will try to construct a K_L^0 veto specific for this analysis
  - Probable approach: BDT trained on ECL and/or KLM cluster properties
  - Some possible variables:
    - KLM variables (number of layers hitten, cluster energy, distance to tracks…)
    - ECL variables (cluster timings, cluster position in the ECL, cluster energy…)



Upper limit estimation

- We will extract signal yield through a template fit
  - Fitting directly the EECL distribution
- For now we have an UL estimation with Asimov data rescaled to BelleII Run1 luminosity (362 fb^{-1})
  - $\mathbf{B^0_d \to \tau^+\tau^- \lesssim 6 \times 10^{-4}}$ **(@90% CL)**





Summary and next steps

- The main path for the analysis is in place
  - Need to check data/MC agreement in sidebands and for off-resonance data
    - We need to be sure that the EECL distribution is well modeled
    - As well as the input features in the BDT
    - Tighter ROE masks can help
  - Belle will follow (very) shortly
  - Systematics & corrections
    - Main part of centrally produced corrections still need to be computed for the new MC version
    - The analysis is *of course* statistically limited

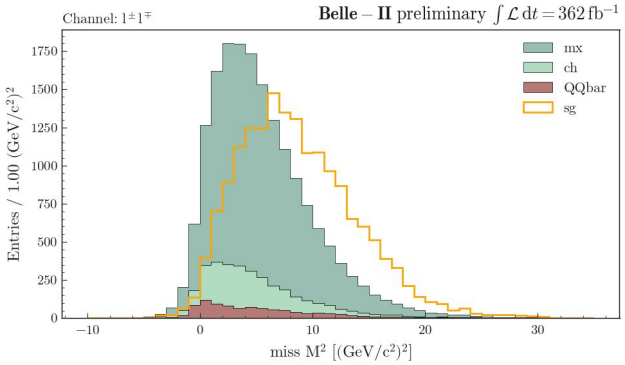
With the current Belle (II) dataset we will provide a competitive result for $B^0 \to \tau^+\tau^-$

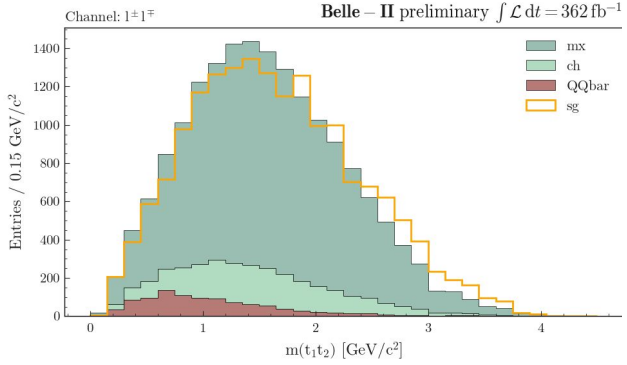
BACKUP

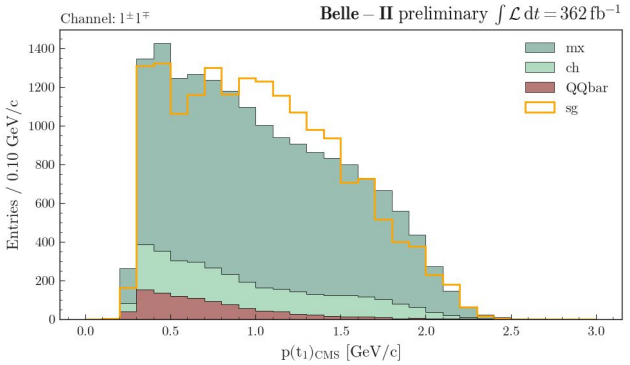
Belle status

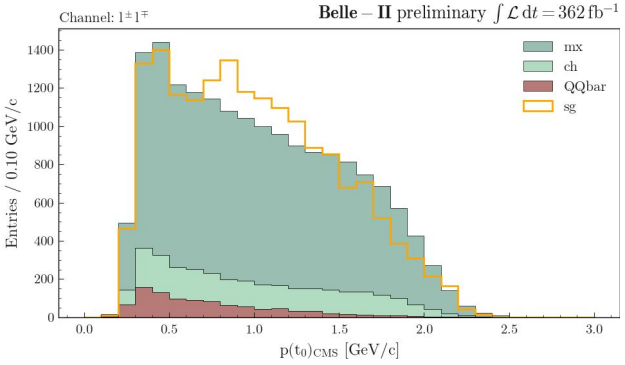
- We re-generated signal for Belle recently
  - Moved from PHSP to SLL in evtGen
- We tried to unify all the selections at reconstruction level with Belle II
  - There are some intrinsic differences in PID
  - Some cuts for tracking and photons differ
- We will use the same pre-selections as of BelleII and retrain the BDTs specifically on Belle dataset
- Last step is the simultaneous fit for both Belle and Belle II (8 fits in total)
- Ntuples production is ongoing

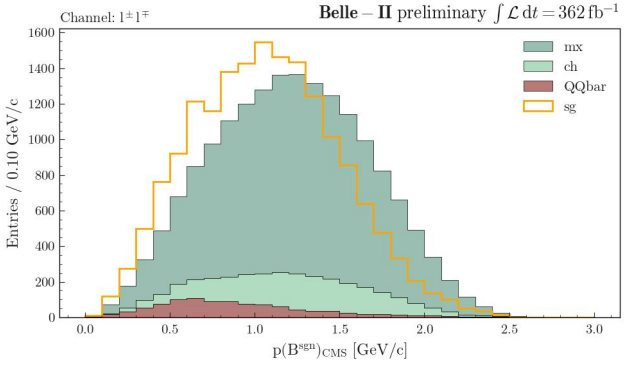
BDT variables - *ll*

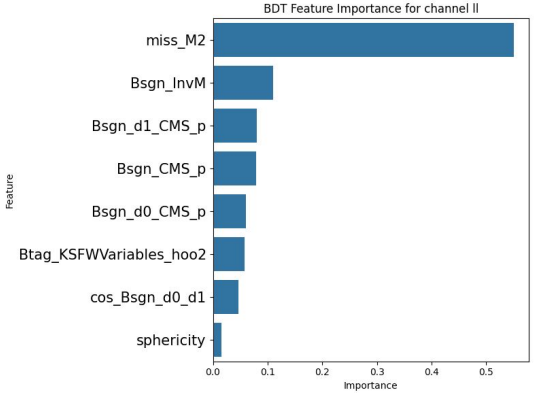




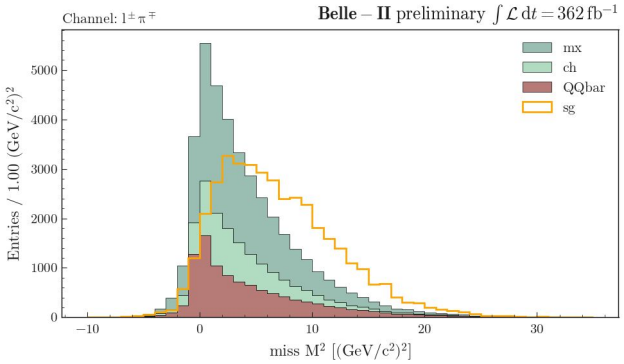


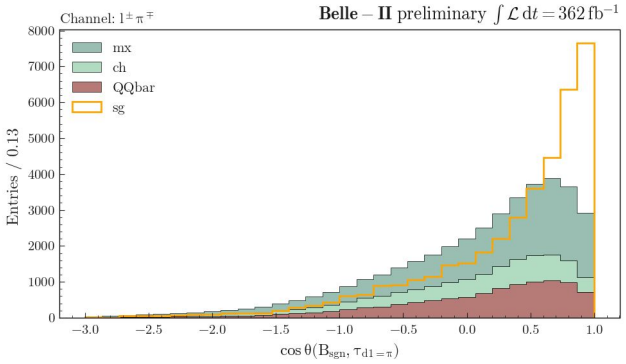


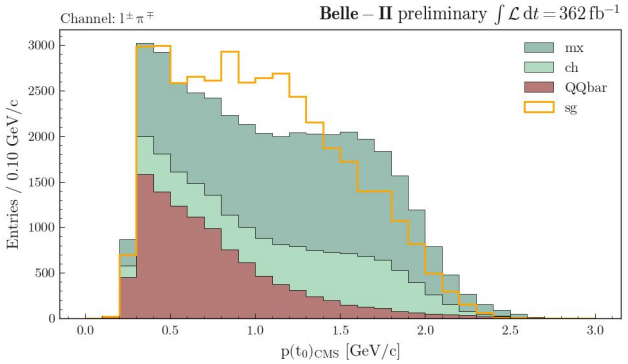


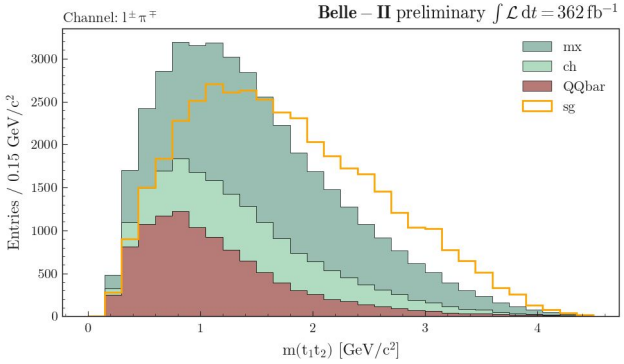


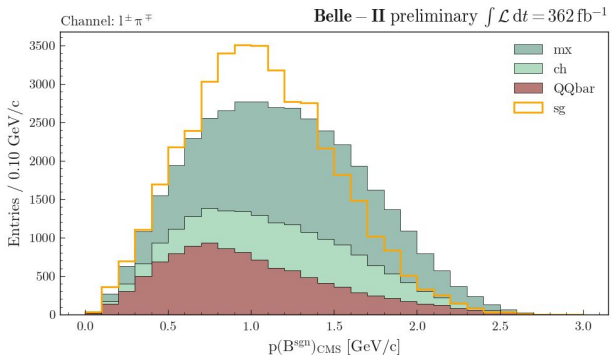
BDT variables - $l\pi$

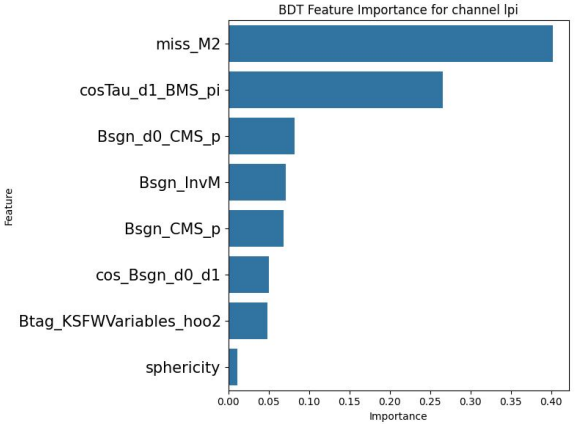




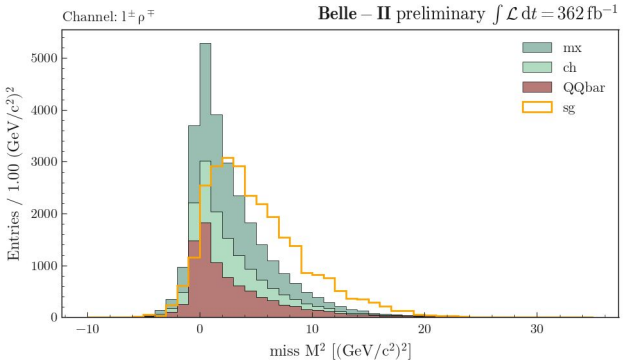


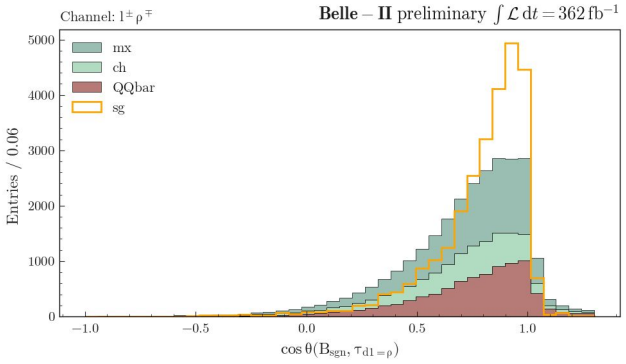


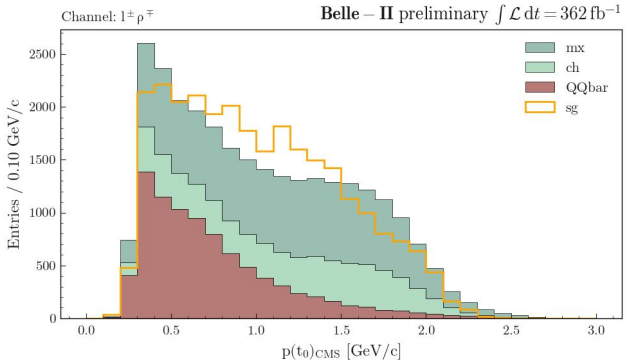


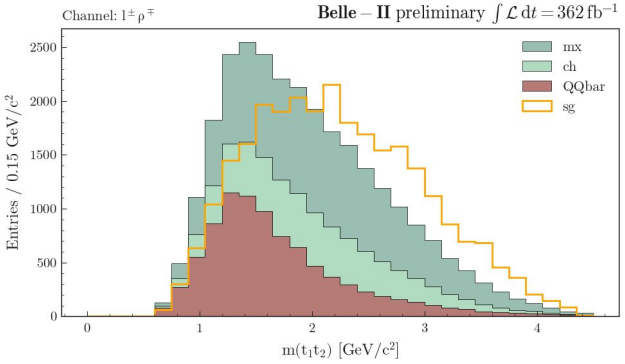


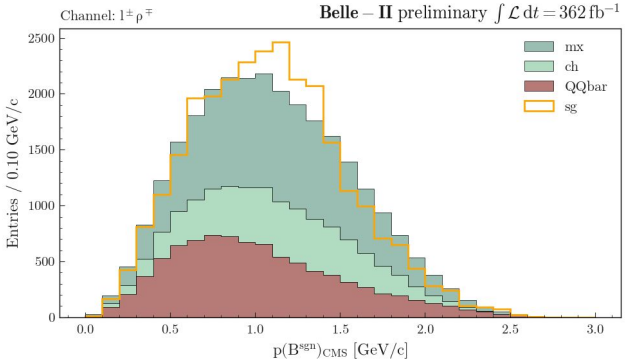
BDT variables - $l\rho$

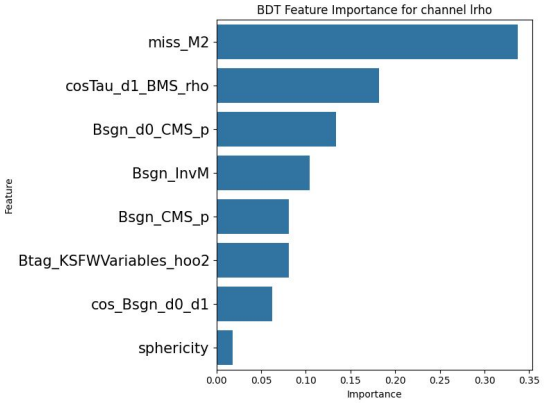




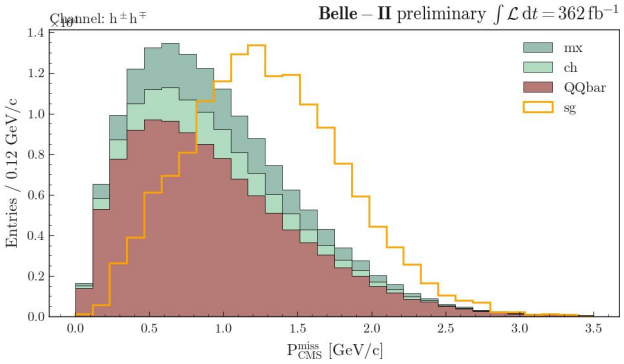


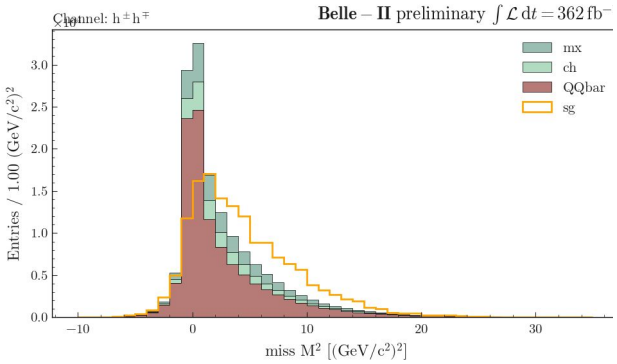


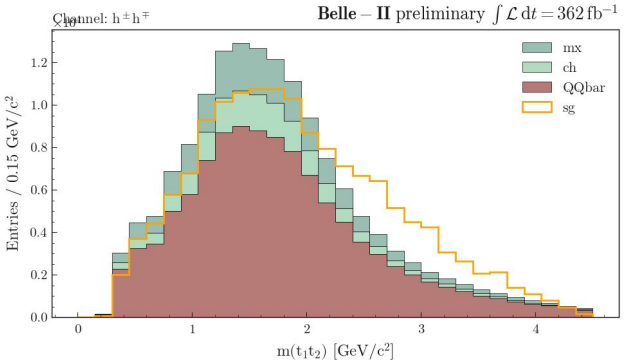


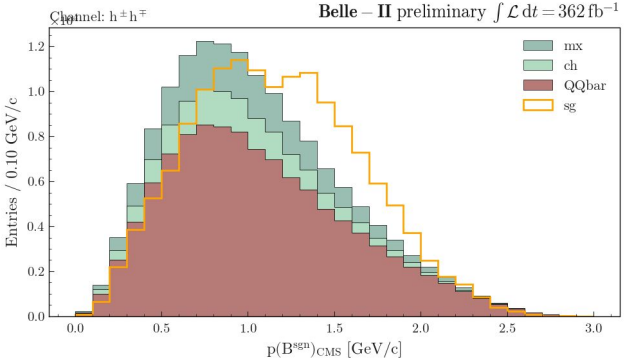


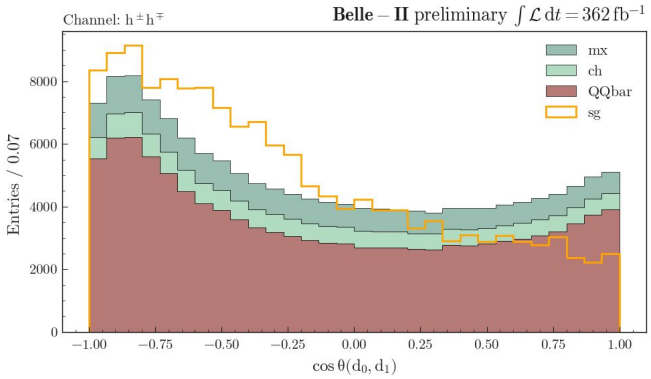
BDT variables - *hh*

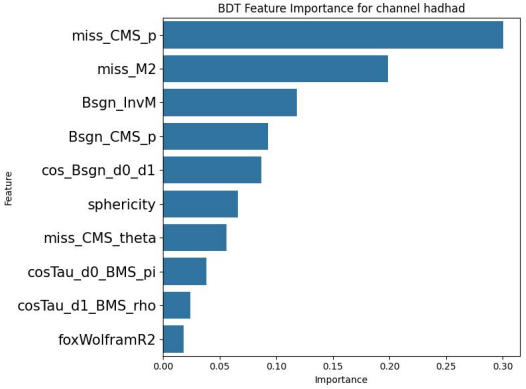




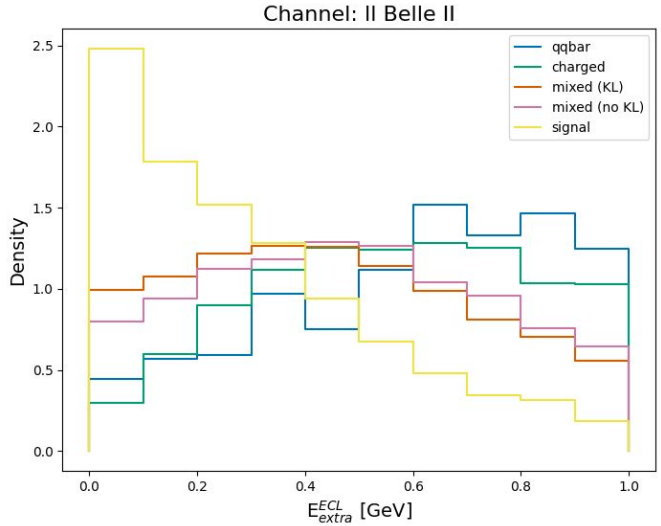


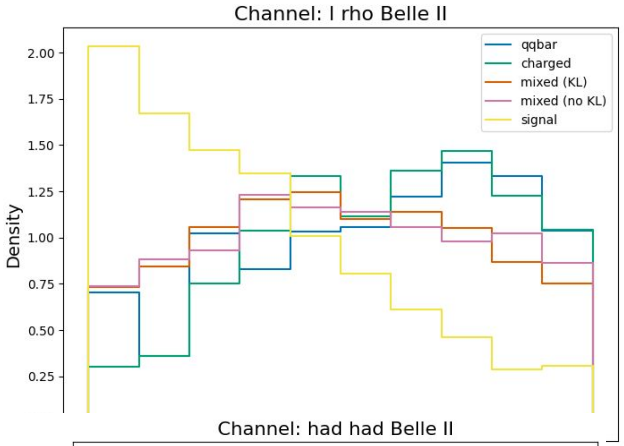


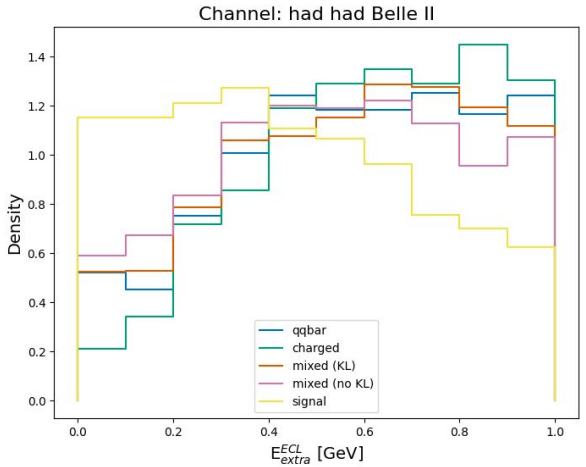


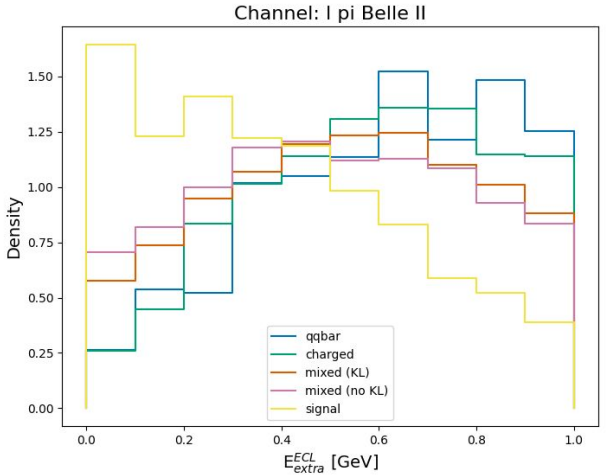


Extra energy ECL distributions









Belle status

- Same preselection as BelleII is applied
- The order of magnitude is similar to BelleII and we see also the same behaviour for channels with $t_1 \neq t_2$

Channel (tau tau ->)	Bkg events (per 100fb^{-1})	Signal eff [x10^{-4}]
mu-pi	6'969	3.40
el-pi	8'188	4.87
el-el	976	0.89
el-mu	1'444	1.27
mu-mu	488	0.42

general agreement in the distributions with BelleII