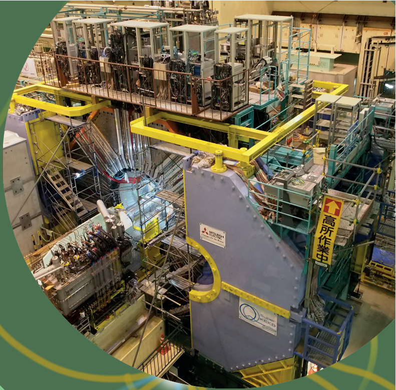
SEARCH FOR $\tau^- \rightarrow \mu^-\mu^+\mu^-$ LEPTON FLAVOUR VIOLATING DECAYS AND MEASUREMENT OF THE SVD SPATIAL RESOLUTION AT BELLE II EXPERIMENT

PhD Thesis Defense



Tuesday 31st October, 2023

Robin Leboucher

Supervisor:
Justine SERRANO

Jury members:
Sasha GLAZOV
Patrick ROBBE
Victor COCO
Cristinel DIACONU

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

1. The Lepton Flavour Violation

2. The Belle II experiment

3. Measurement of SVD spatial resolution

4. Search for $\tau \to \mu\mu\mu$ LFV decays
4.1 Untagged analysis strategy
4.2 Events reconstruction
4.3 Background suppression
4.4 Expected background yields
4.5 Study of the systematic uncertainties
4.6 Computation of the upper-limit on branching fraction

5. Conclusion and prospects

1. The Lepton Flavour Violation

2. The Belle II experiment

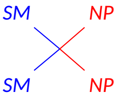
3. Measurement of SVD spatial resolution

4. Search for $\tau \to \mu\mu\mu$ LFV decays
4.1 Untagged analysis strategy
4.2 Events reconstruction
4.3 Background suppression
4.4 Expected background yields
4.5 Study of the systematic uncertainties
4.6 Computation of the upper-limit on branching fraction

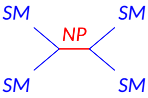
5. Conclusion and prospects

The Standard Model of Particle Physics

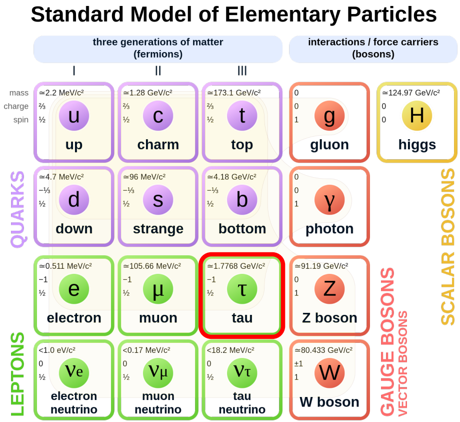
- The **Standard Model (SM)** is the best description of interactions between particles
- SM describes the particles by quantum fields in interaction
- Nevertheless, the **SM is still incomplete**:
  - Dark matter and energy
  - Matter-antimatter asymmetry
  - Neutrino oscillations and masses...
- All these phenomena are clues to the existence of **physics beyond the standard model**



Direct searches in **energy frontier**



Indirect searches in **intensity frontier**



10.1093/ptep/ptaa104. 2020

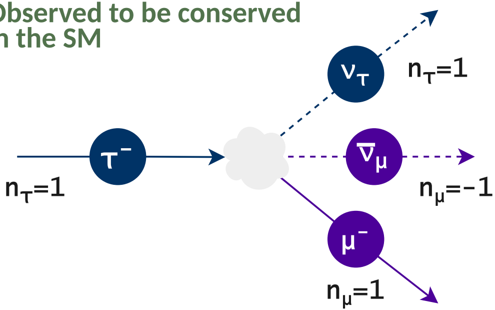
Lepton Flavours in the Standard Model

Lepton Numbers denotes **which particles are leptons** and which particles are not

$$L_e = 1 \begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \quad L_\mu = 1 \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \quad L_\tau = 1 \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}$$

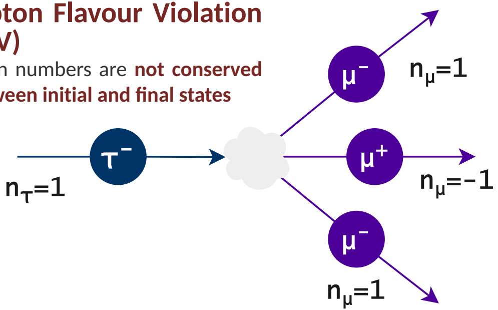
$$L_e = -1 \begin{pmatrix} \overline{\nu}_e \\ e^+ \end{pmatrix} \quad L_\mu = -1 \begin{pmatrix} \overline{\nu}_\mu \\ \mu^+ \end{pmatrix} \quad L_\tau = -1 \begin{pmatrix} \overline{\nu}_\tau \\ \tau^+ \end{pmatrix}$$

Observed to be conserved in the SM



Lepton Flavour Violation (LFV)

when numbers are **not conserved between initial and final states**



Lepton Flavours in the Standard Model – *Neutrino sector*

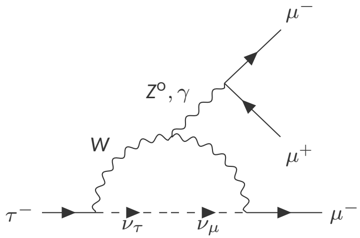
Neutrino oscillations

- **The only source of lepton flavour violation in the Standard Model.**
- First evidence in 1998 by **Super-Kamiokande**.
- Described by the Pontecorvo–Maki–Nakagawa–Sakata **PMNS matrix**.
- **Consequences:**
  - neutrinos have masses,
  - **neutrinos flavour oscillates**.

$$\underbrace{\begin{pmatrix}\nu_e\\ \nu_\mu\\ \nu_\tau\end{pmatrix}}_{\text{Flavours}} = \underbrace{\begin{pmatrix}U_{e1} & U_{e2} & U_{e3}\\ U_{\mu1} & U_{\mu2} & U_{\mu3}\\ U_{\tau1} & U_{\tau2} & U_{\tau3}\end{pmatrix}}_{\text{PMNS matrix}} \cdot \underbrace{\begin{pmatrix}\nu_1\\ \nu_2\\ \nu_3\end{pmatrix}}_{\text{Masses}}$$

Charged Lepton Flavour violation in the SM

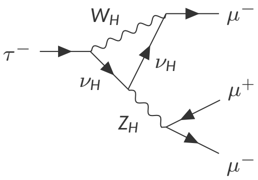
- **Neutrino oscillations introduce Flavour-Changing Neutral Currents** via $W^{\pm}$ boson loops.
- **Processes heavily suppressed**, decay rates below $\mathcal{O}(10^{-50})$.



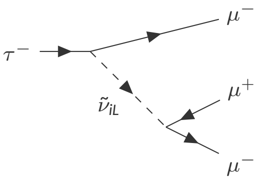
The $\tau^- \to \mu^-\mu^+\mu^-$ LFV decays Beyond the Standard Model

The $\tau^- \to \mu^-\mu^+\mu^-$ **LFV decay appears in many New Physics** models as:

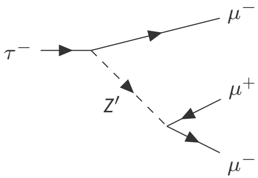
New Physics models	Largest allowed branching fraction
Littlest Higgs with T-parity	10^{-8}
R-parity violating SUSY	10^{-8}
Non-universal Z'	10^{-8}
MSSM + seesaw	10^{-9}
SUSY SO(10)	10^{-10}
SUSY Higgs	10^{-10}
SM + seesaw	10^{-10}
V(1) Leptoquarks	10^{-12}



Littlest Higgs with T-parity



R-parity violating SUSY



Non-universal Z'

Strength: purely muonic final state $\Rightarrow$ **small background**

10.1103/PhysRevD.77.073010. 2008

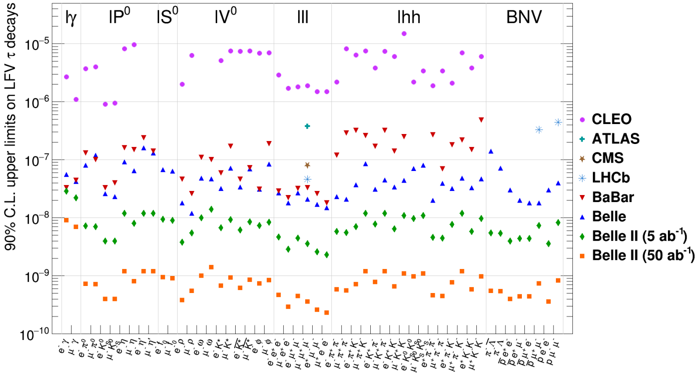
Searches for LFV in τ decays at colliders

Current status

Around 50 τ LFV channels have been studied in the past two decades; **the best upper limits on branching fractions are set by the Belle experiment** (10^{-8} — 10^{-7} range).

Upper limits on $\tau^- \to \mu^-\mu^+\mu^-$ at 90% CL ($\times 10^{-8}$)

Belle	Babar	LHCb	ATLAS	CMS
782 fb^{-1}	468 fb^{-1}	3 fb^{-1}	20.3 fb^{-1}	131 fb^{-1}
2.1	3.3	4.6	38	2.9



Banerjee et al., 2022a; Kou et al., 2019a

Belle II Prospects

Belle II is expected to **improve current limits by at least 1** order of magnitude $\Rightarrow$ **sensitive to some NP models.**

1. The Lepton Flavour Violation

2. The Belle II experiment

3. Measurement of SVD spatial resolution

4. Search for $\tau \to \mu\mu\mu$ LFV decays
4.1 Untagged analysis strategy
4.2 Events reconstruction
4.3 Background suppression
4.4 Expected background yields
4.5 Study of the systematic uncertainties
4.6 Computation of the upper-limit on branching fraction

5. Conclusion and prospects

SuperKEKB status

SuperKEKB:

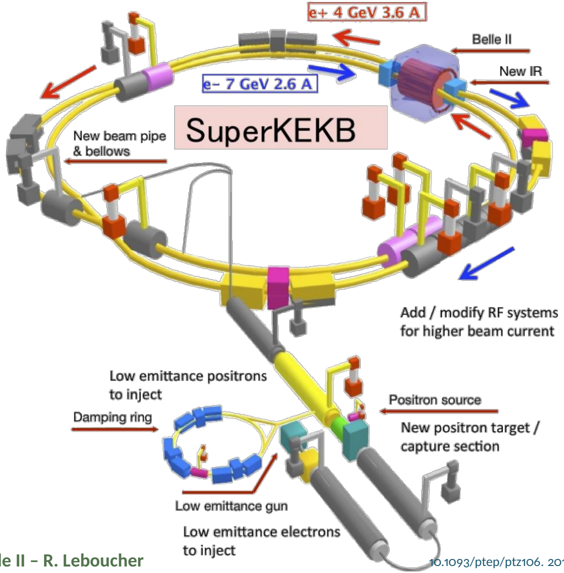
- e^+e^- collider with $E_{CM} = \sqrt{s} = 10.58$ GeV, $\Upsilon(4S)$ resonance
- Currently hold **world highest instantaneous luminosity** $4.7 \times 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$

Status: First long shutdown since July 2022, **424** fb^{-1} collected by Belle II since 2019
$\rightarrow$ including **362** fb^{-1} @ $\Upsilon(4S)$

SuperKEKB advantages for τ physics:

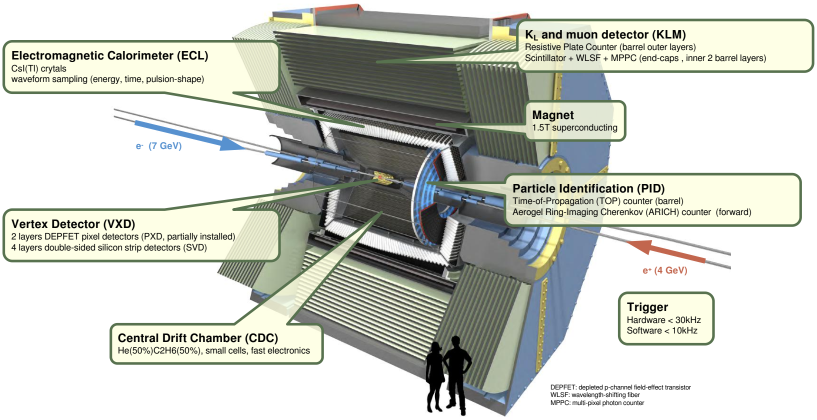
- e^+e^- collision:
  - Well-defined kinematics of initial state
  - Clean environment; small pile-up
  - High tau-pair production:

$$\sigma(e^+e^- \to \tau^+\tau^-) = 0.92\ \text{nb}$$
$$\sigma(e^+e^- \to B\bar{B}) = 1.05\ \text{nb}$$



10.1093/ptep/ptz106. 2019

Belle II Detector



Belle II important features for τ physics:

- Hermetic detector: **good missing energy reconstruction**
- **Special triggers** dedicated dedicated to low track multiplicity events
- Excellent vertexing and tracking capabilities

10.1093/ptep/pts083. 2013

1. The Lepton Flavour Violation

2. The Belle II experiment

3. Measurement of SVD spatial resolution

4. Search for $\tau \to \mu\mu\mu$ LFV decays

4.1 Untagged analysis strategy

4.2 Events reconstruction

4.3 Background suppression

4.4 Expected background yields

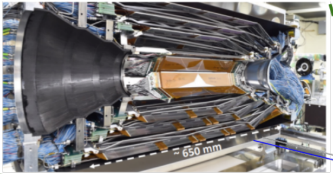
4.5 Study of the systematic uncertainties

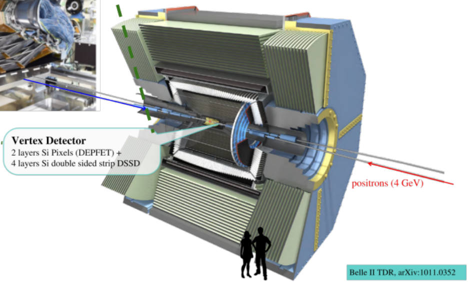
4.6 Computation of the upper-limit on branching fraction

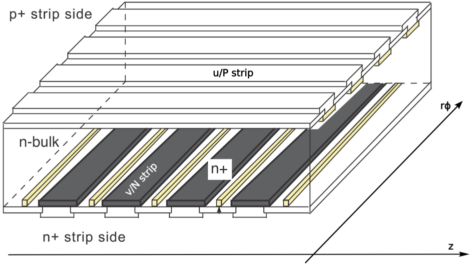
5. Conclusion and prospects

The Silicon Vertex Detector (SVD)

The Belle II Vertex Detector is composed of 2 layers (1 to 2) of pixels and 4 layers (3 to 6) of **double-sided silicon strip sensors** with radii from 39 mm to 135 mm







u/P side along $r\phi$-direction
v/N side along z-direction

Spatial resolution is crucial to:

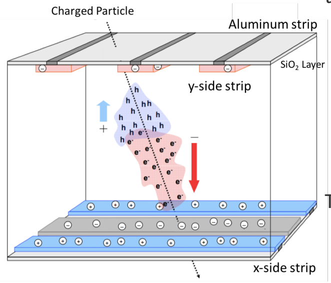
Provide best quality track reconstruction

Correctly propagate uncertainty on hit's position to track parameters

10.1088/1748-0221/17/11/P11042. 2022

Cluster position and resolution

- **Charged particles** passing cross sensors and **create electron-hole pairs**
- Electrons (holes) are **collected by a cluster** of strips



Digital resolution (design) obtained from **pitch** (distance between two strips):

$$\sigma_{digital} = \frac{Pitch}{2\sqrt{12}}$$

The spatial resolution is defined as:

$$\sigma^2_{true} \equiv E[(x - m)^2]$$

where x is the **true hit position**

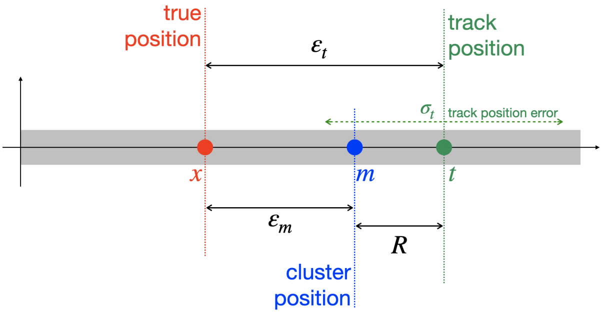
Define the **cluster position** (m) by:

$$m = \frac{\sum_{i=1}^{N} X_i S_i}{\sum_{i=1}^{N} S_i},$$

- N is the number of strips in the cluster
- S_i is the collected charge by a strip
- X_i is the position of the strip

True hit position is **only known in simulation**

Cluster position and resolution



Quantities linked to SVD hits used to measure the spatial resolution σ_{cl} in data:

- x the true position **(only available in simulations)**
- m the position of a cluster
- t the **unbiased extrapolated track position**

Defined **residual** R:

$$\mathbf{R = m - t}$$

Overlapped sensor method

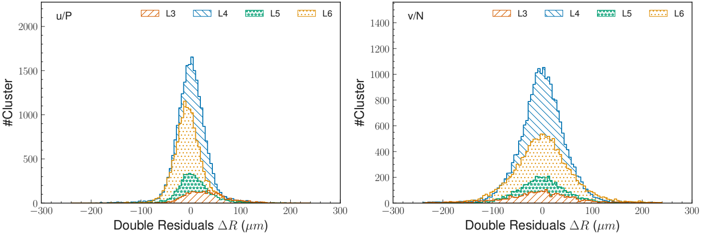
Overlapped sensor method (adapted from CMS):

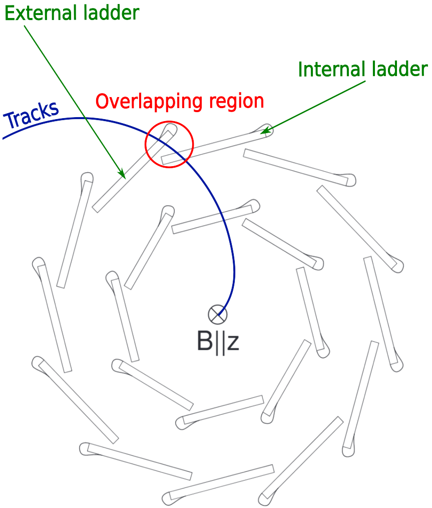
- **Select tracks with two hits on the same layer and consecutive ladders**
- **Compare residuals** computed for the pair of overlapping ladders, double residuals:

$$\Delta R = R_{int} - R_{ext}.$$

- **Spatial resolution** (σ_{cl}):

$$\sigma_{cl}^2 = Var(\Delta R),$$





- Apply geometrical correction due to non-parallel sensors
- Extract the variance with `sigma-68` estimator of a Student-T distribution fit

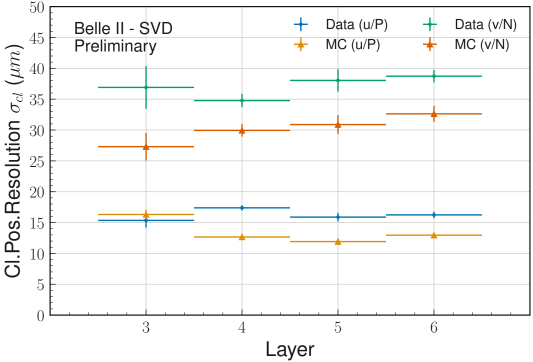
10.1088/1748-0221/4/05/P05004. 2009

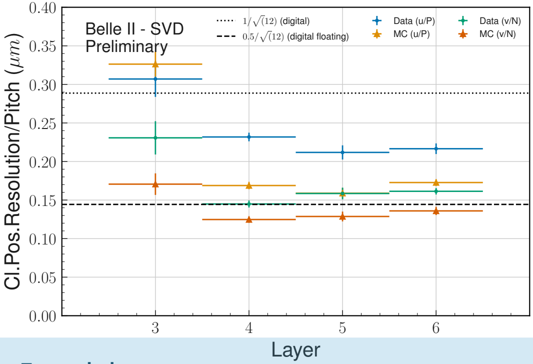
Results on the spatial resolution

10.1016/j.nima.2022.166746. 2022

Measured resolutions on data

≈15 μm for the u/P side **≈35 μm for the v/N side**





To conclude:

- **First try of overlapped sensor method in Belle II** in complement to other methods (with lower estimated resolution)
- Not sensitive to Coulomb scattering and tracking uncertainty
- Limited range of incident angle

1. The Lepton Flavour Violation

2. The Belle II experiment

3. Measurement of SVD spatial resolution

4. Search for $\tau \to \mu\mu\mu$ LFV decays

4.1 Untagged analysis strategy
4.2 Events reconstruction
4.3 Background suppression
4.4 Expected background yields
4.5 Study of the systematic uncertainties
4.6 Computation of the upper-limit on branching fraction

5. Conclusion and prospects

Analysis strategy

Aim: observe the $\tau^- \to \mu^-\mu^+\mu^-$ decay or set an upper limit on the branching fraction

$$\mathcal{B}_{\text{UL}}(\tau^- \to \mu^-\mu^+\mu^-) = \frac{s_{\text{UL}}}{\mathcal{L} \times 2\sigma_{\tau^-\tau^+} \times \varepsilon_{\text{sig}}}$$

Needed parameters:

- s_{UL}: the upper limit on the difference between observed and expected data yields $N_{obs} - N_{exp}$,
- $\mathcal{L}$: integrated luminosity of Belle II data (424 fb^{-1}),
- $\sigma_{\tau^-\tau^+} \simeq 0.919$ nb: tau-pair production cross-section,
- ε_{sig}: absolute signal efficiency $\to$ estimated in simulation.

Strategy

1. Untagged events reconstruction
2. Background rejection (optimised on the 362 fb^{-1} $\Upsilon(4S)$ sample)
   - Cut-based preselections
   - Boosted Decision Tree classifier

   Following the Punzi's figure of merit:

$$FOM = \frac{\varepsilon_{sig}}{3/2 + \sqrt{N_{bkg}}}$$

3. Estimation of surviving background in data (generalized on the 424 fb^{-1} full sample)
4. Estimation of systematic uncertainties
5. Computation of the expected upper-limit on branching fraction

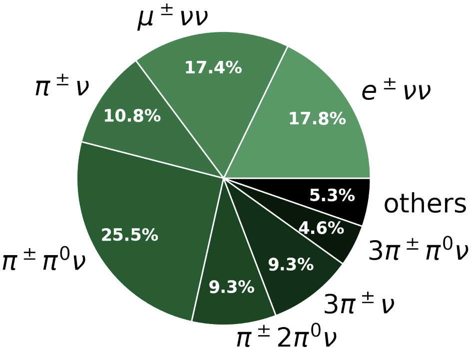
Tau $\longrightarrow$ 3 muons event reconstruction – *Tagged strategy*

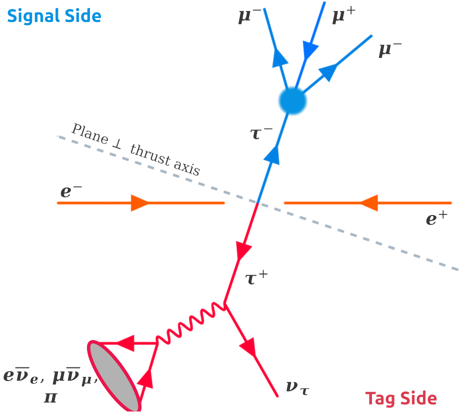
Belle & BaBar tagged strategy

Reconstruct τ pair events with an tagged strategy:

- **Signal side**: one τ decay into three muons
- **Opposite side**: oppositely charged τ (**tag**) constrained to 1-prong decays

$\Rightarrow$ only events with **exactly 4 tracks**





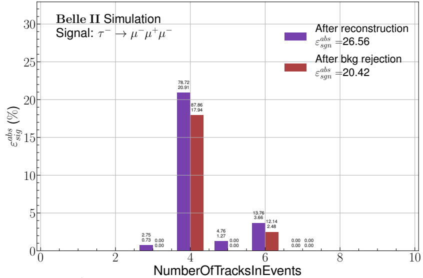
10.1016/j.physletb.2010.03.037. 2010; 10.1103/PhysRevD.81.111101. 2010

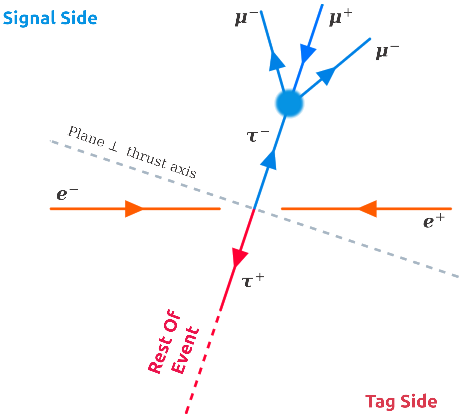
Tau $\to$ 3 muons event reconstruction – *Untagged strategy*

Belle II strategy

Reconstruct tau-pair events with an untagged strategy:

- **Signal side**: one τ decay into three muons
- **Untagged Opposite side**: **Rest of Event** (max 3 tracks) $\Rightarrow$ **signal efficiency** increase:
  - decays of τ_{tag} into 3-prong ($\mathcal{B} \sim 15\%$),
  - events with missing or additional track(s)...
- But **increase background contamination** mainly: $e^+e^- \to q\bar{q}$, $\tau^-\tau^+$, $\mu^-\mu^+$, $\ell^-\ell^+\mu^-\mu^+$.





Tau → 3 muons event reconstruction – *Event requirements*

Events reconstruction strategy

Reconstruct τ pair events with an untagged strategy:

- **Signal side**: one τ decay into three muons
- **Untagged Opposite side**: **Rest of Event** (max 3 tracks) ⇒ **signal efficiency** increase:
  - decays of τ_{tag} into 3-prong ($\mathcal{B} \sim 15\%$),
  - events with missing or additional track(s)...
- But **increase background contamination** mainly: $e^+e^- \rightarrow q\bar{q}$, $\tau^-\tau^+$, $\mu^-\mu^+$, $\ell^-\ell^+\mu^-\mu^+$.

Track requirements

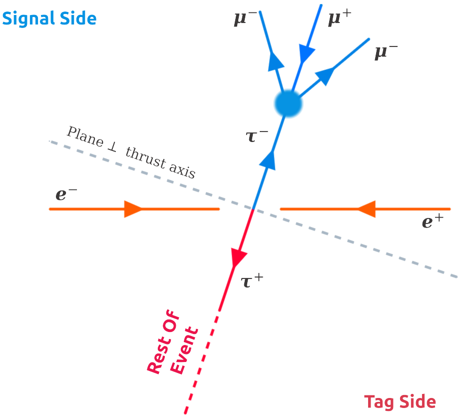
- Tracks coming from the interaction point
- Tracks identified as muons (likelihood-based)

Event requirements:

- Three muons in the same hemisphere, defined by the thrust

$$T = max_{n_T}\left(\frac{\sum_i |\mathbf{p_i} \cdot \mathbf{n_T}|}{\sum_i |\mathbf{p_i}|}\right)$$

- Pass trigger requirements



ε_{sig}^{abs} (%)	N_{bkg}	$N_{\tau\tau}$	$N_{q\bar{q}}$	$N_{\mu\mu}$	$N_{\ell\ell\mu\mu}$
34.35	1819.67	6.15	842.73	157.38	807.17

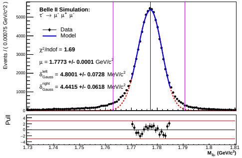
For 362 fb^{-1}

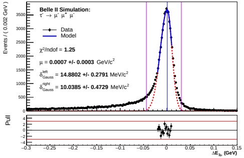
Signal Region in $M_{3\mu}$; $\Delta E_{3\mu}$ 2D plane

Measure resolutions to define **region** around **signal** centered at:
$M_{3\mu} \simeq 1.777\ \text{GeV}/c^2$
and $\Delta E_{3\mu} = E^{\text{CM}}_{3\mu} - \sqrt{s}/2 \simeq 0\ \text{GeV}$

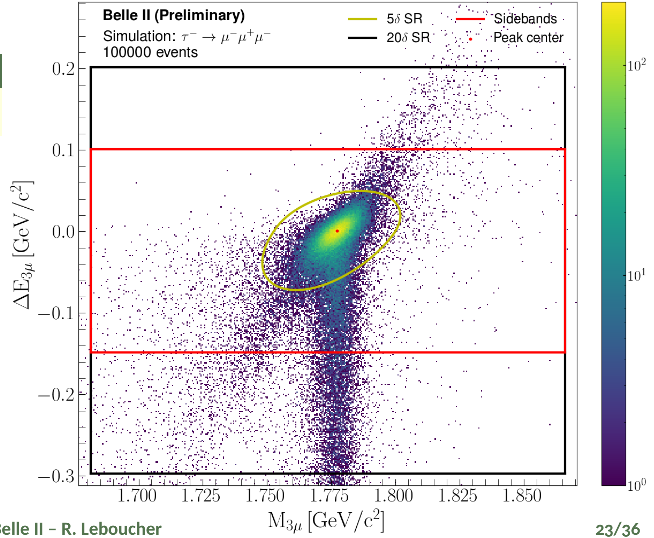
Fit $M_{3\mu}$ and $\Delta E_{3\mu}$

- **Asymmetric Gaussian:** on a limited range for the peak





- **20δSR**: To optimise background rejection and train the BDT
- **5δSR**: Final yields estimation
- **SB**: for data/Simulation comparisons with **5δSR** blinded

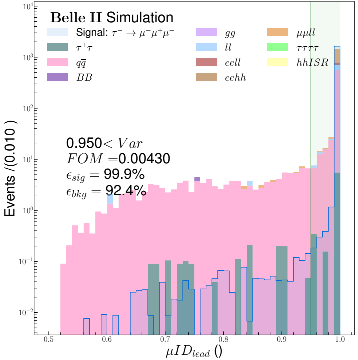


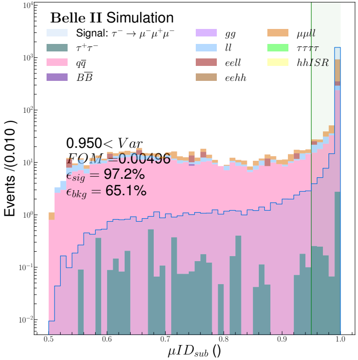
Cut-based preselections – *Muon identification selection*

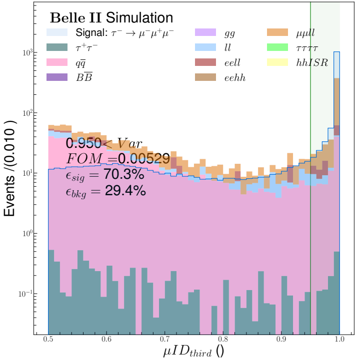
Muon identification selection:

- **Rank the three muons** according to muonID,
- Only cut on leading and subleading muonID.

$$\varepsilon_{sig} = 32.12\% \qquad N_{bkg} = 1043$$



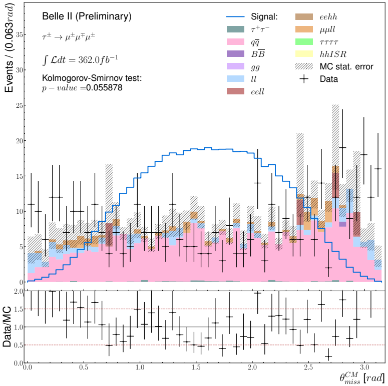


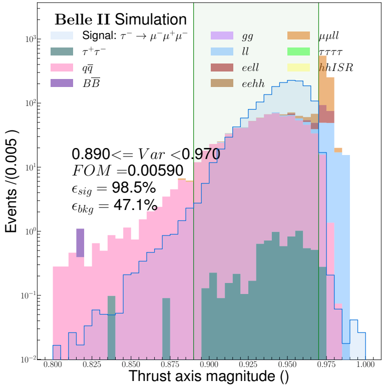


Cut-based preselections – *Additional selections*

Cut-based preselections to reject:

- missing di-photons events in the simulation
- $e^+e^- \to \mu^-\mu^+, \ell^-\ell^+\mu^-\mu^+$ backgrounds





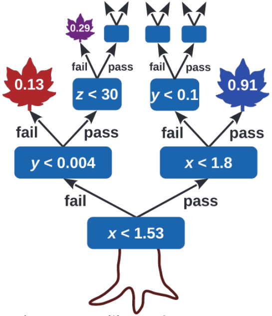
Preselection	ε^{abs}_{sig} (%)	N_{bkg}	$N_{\tau\tau}$	$N_{q\bar{q}}$	$N_{\mu\mu}$	$N_{\ell\ell\mu\mu}$
0.3 < theta$^{CM}_{miss}$ < 2.7 **0.89 < Thrust < 0.97**	24.32	253.74	2.05	220.12	2.06	29.51

For 362 fb^{-1}

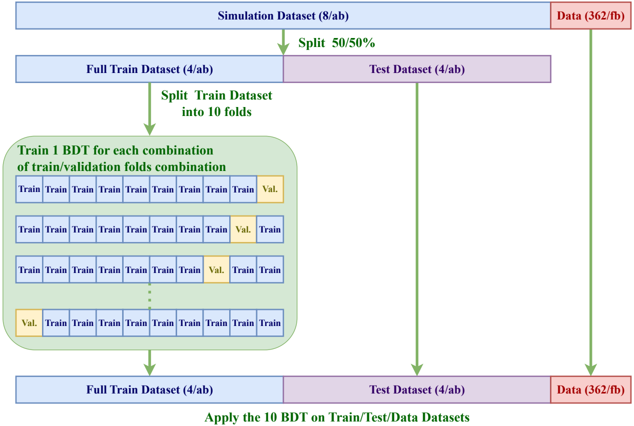
Boosted Decision Tree Classifier (BDT) – *Algorithm*

Remaining background: $q\bar{q}$ and tau pair $\Rightarrow$ Rejected by Boosted Decision Tree classifier

BDT classifiers are **machine learning algorithms** utilising **multiple decision trees** to create a highly accurate predictive model.



- Using python's XGBoost library for BDT
- Using python's Optuna library for optimization over 100 **hyper-parameters** setups



10.1142/12200. 2022

Boosted Decision Tree Classifier (BDT) – *Input variables*

Use 32 inputs variables containing:

Rest Of Event (ROE): combines all non-signal tracks and remaining ECL clusters,

Missing momentum: difference between the beam and the sum of all particles momenta,

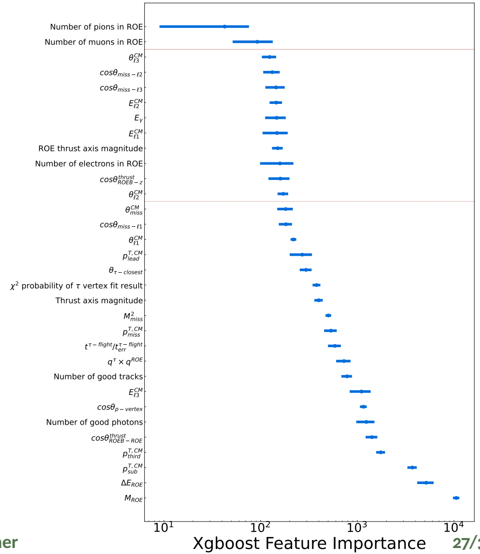
Topology: thrust and angles,

Ranked transverse momenta of the three muons,

Charged and neutral **particles multiplicities**.

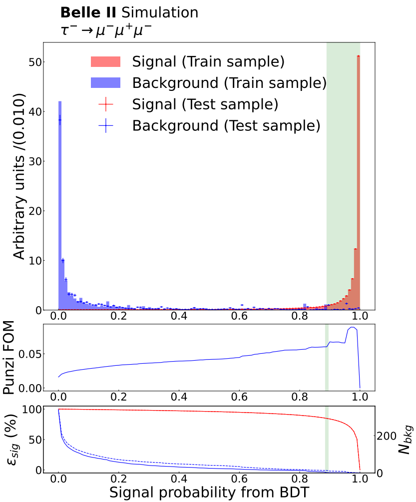
Variable selected following 3 criteria:

- importance in the BDT
- low correlation with $M_{3\mu}$ and $\Delta E_{3\mu}$
- data/simulation agreement



Boosted Decision Tree Classifier (BDT) – *BDT Output*

Averaged BDT's output scores in $\pm 20\delta$ SR box



Low background statistic in the SR $\Rightarrow$ Modify the BDT output and SR cut optimisation to minimize the sensitivity to fluctuations

We assume the background is uniformly distributed in the SR plane:

1. Try a cut on BDT output in the $\pm 20\delta$ SR box of area $\mathcal{A}_{20\delta}$
2. Number of surviving background events in $\pm n\delta$ SR ellipse of area $\mathcal{A}_{n\delta}$:

$$N_{bkg} = N_{bkg}^{20\delta} \times \mathcal{A}_{n\delta} / \mathcal{A}_{20\delta}$$

3. Keep the selection that maximizes Punzi's FOM

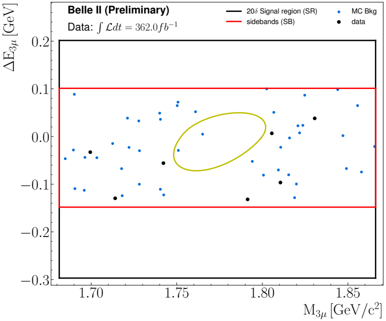
Final selection:

$\mu ID_{lead} > 0.95$
$\mu ID_{sub} > 0.95$
$\mu ID_{third} > 0.5$
$0.3 < \theta_{miss}^{CM} < 2.7$
$0.89 < thrust < 0.97$
$p^{BDT}(Signal) > 0.9$
5δ ellipse region

	Test
ε_{sig} (%)	21.76
N_{bkg} for 362 fb^{-1}	0.16
N_{bkg} for 4 ab^{-1}	2 ($q\bar{q}$)

Data/Simulation Comparison – *Sidebands*

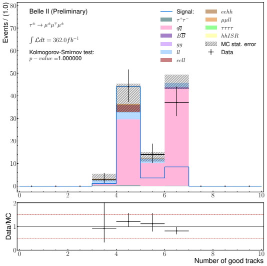
Check the Data/Simulation agreement in **blind sidebands region**: $\pm 20\delta M_{3\mu}$ and $\pm 10\delta \Delta E_{3\mu}$ while $\pm 5\delta$ box is blind.



Following the Review Committee suggestions, we applied a requirement on the total event charge

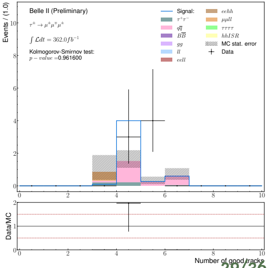
After preselection

Simulation	$121.13.0^{+9.8}_{-5.2}$
Data	$98.0^{+10.9}_{-9.9}$



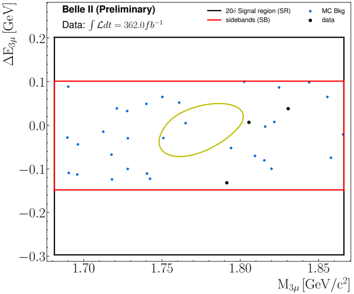
After preselection + BDT

Simulation	$3.29^{+1.24}_{-0.73}$
Data	$7.0^{+3.77}_{-2.58}$



Data/Simulation Comparison – *Sidebands*

Check the Data/Simulation agreement with total charge requirement.



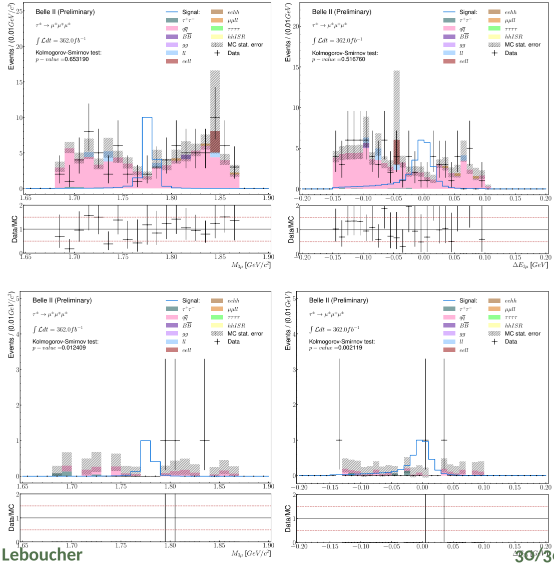
	Test
ε_{sig} (%)	20.42
N_{bkg} for 362 fb^{-1}	0.16

After preselection

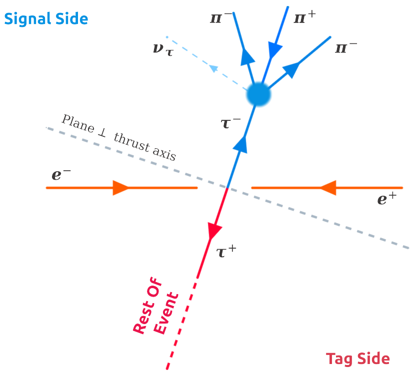
Simulation	$96.39^{+9.4}_{-4.5}$
Data	$80.0^{+9.9}_{-8.9}$

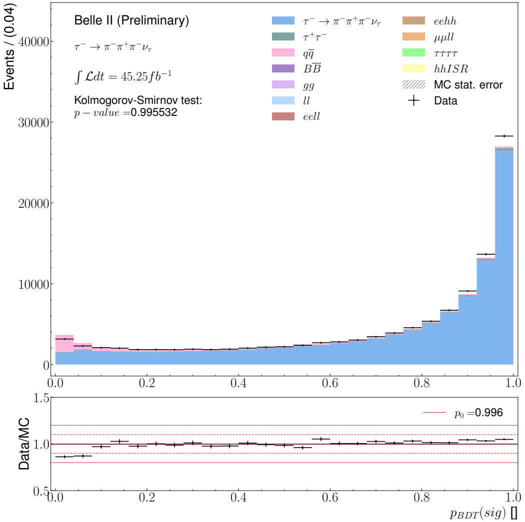
After preselection + BDT

Simulation	$2.05^{+0.71}_{-0.47}$
Data	$3.0^{+2.92}_{-1.63}$



Data/Simulation Comparison – $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$ Control Sample

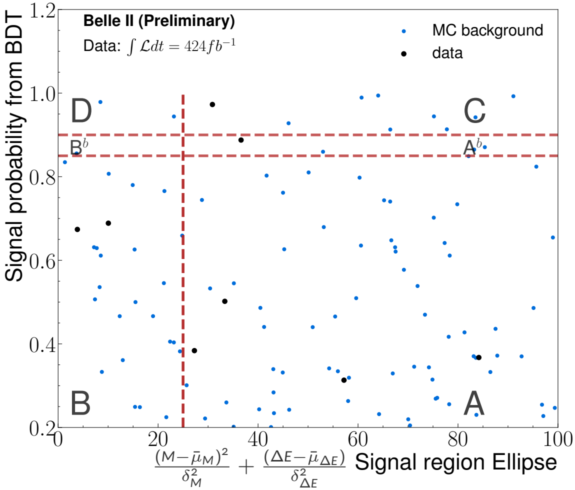




Similar reconstruction as $\tau^- \rightarrow \mu^-\mu^+\mu^-$

- **Avoid unblinding by reverting muon by pion identification**
- Cut-based selection to purify signal $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$

Background estimation in the signal region – *Data-driven ABCD method*



ABCD method

Fully **data-driven method**

1. Define a 2D plane: **Distance from the peaking signal** in SR plane VS **BDT output**
2. Define 4 regions ABCD
   D = $\pm 5\delta$ **SR** with $p(BDT) > 0.9$

$$N_D^{expected} = N_C \times R_{B/A}$$

$R_{B/A}$	$0.50^{+0.77}_{-0.40} stat$
N_C	$1.00^{+2.30}_{-0.83}$
$N_D^{expected}$	$0.50^{+1.38}_{-0.50}$

Validated with simulations

10.1103/PhysRevD.103.035021. 2021

Systematic uncertainties

Systematic uncertainties sources:

- **Particle identification efficiency:** (from Data/Simulation studies).
- **Tracking efficiency:** (from Data/Simulation studies) from tracking performances.
- **Trigger efficiency:** (from Data/Simulation studies) measured trigger efficiency relative to orthogonal lines.
- **BDT selection:** (from Data/Simulation studies) BDT efficiency discrepancy.
- **Luminosity and tau-pair production cross-section.**
- **Momentum scale:** bias due to imperfect magnetic field used in data reconstruction.

Provided by the **collaboration.**
Measured in $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ **control sample.**

Systematic uncertainties:

Quantity	Source	Value	Relative Systematic uncertainties (%) Low	Relative Systematic uncertainties (%) High
ε_{sig}	PID	20.42%	2.106	2.359
	Tracking	20.42%	1.018	1.018
	Trigger	20.42%	0.7	0.7
	BDT	20.42%	1.5	1.5
$\mathcal{L}$		424	0.6	0.6
$\sigma_{\tau\tau}$		0.919	0.326	0.326
N^{SB}_{data}	Momentum Scale	1.00	2.13	1.06

Computation of the upper-limit on branching fraction – *CLs method*

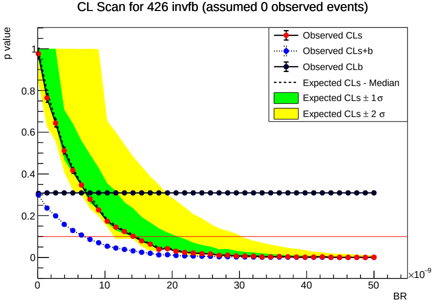
Branching fraction inputs:

$$\mathcal{B}_{\mathrm{UL}}(\tau^- \to \mu^-\mu^+\mu^-) = \frac{s_{\mathrm{UL}}}{\mathcal{L} \times 2\sigma_{\tau^-\tau^+} \times \varepsilon_{\mathrm{sig}}}$$

Quantity	Value	Relative Error (%)			
		Statistical		Systematics	
		Low	High	Low	High
ε_{sig}	20.42%	0.294	0.294	2.865	3.056
$\mathcal{L}$ (fb^{-1})	424	0.005	0.005	0.6	0.6
$\sigma_{\tau\tau}$ (nb)	0.919	-	-	0.326	0.326
N^{SB}_{data}	1.00	83.0	230.0	2.13	1.06
$R_{B/A}$	0.5	80.0	154.0	-	-

Compute upper-limit with CLs method

- Use `RooStat` **library** with generate 10k toys at each scanned points $\mathcal{B}$
- **Not yet unblinded $\Rightarrow$ assumes 0 observed events**



Expected upper-limit on branching fraction at 90% CL:
1.46×10^{-8}.

Belle with 782 fb^{-1}

$\mathcal{B}_{UL}$	ε_{sig} (%)	N_{bkg}	N_{obs}
2.1×10^{-8}	7.6	0.13	0

10.1140/epjc/s10052-011-1554-0. 2011

1. The Lepton Flavour Violation

2. The Belle II experiment

3. Measurement of SVD spatial resolution

4. Search for $\tau \to \mu\mu\mu$ LFV decays

4.1 Untagged analysis strategy

4.2 Events reconstruction

4.3 Background suppression

4.4 Expected background yields

4.5 Study of the systematic uncertainties

4.6 Computation of the upper-limit on branching fraction

5. Conclusion and prospects

Conclusion and prospects

CPPM

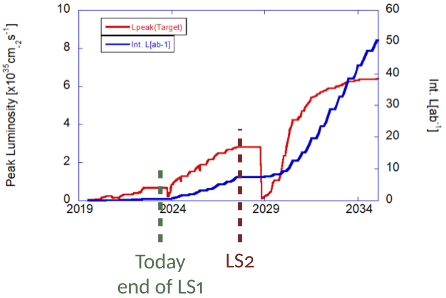
Conclusion:

- First search for $\tau^- \to \mu^-\mu^+\mu^-$ with an untagged reconstruction at Belle II
- Expect to be competitive with Belle (2.1×10^{-8} for $782\,\text{fb}^{-1}$) even with 1 observed event ($2.\times10^{-8}$)
- Independent crosscheck in Belle II with tagged and cut-based strategy (15% signal efficiency for UL 2×10^{-8})
- Waiting review committee green light for unblinding

Prospects:

- Belle II is going to resume data taking after a first long shutdown $\Rightarrow$ expect more luminosity
- Near future Belle II will lead the search for τ LFV decays with at least $5\,\text{ab}^{-1}$ before the second shutdown $\Rightarrow$ by extrapolation 2.9×10^{-9} for $5\,\text{ab}^{-1}$ and 2.1×10^{-9} for $25\,\text{ab}^{-1}$
- Longer-term Belle II efforts will possibly be joined by Super Charm and Tau factories, fixed target experiments or even the FCCee

Expected upper-limit on branching fraction at 90% CL: 1.46×10^{-8}.



Thank you!

Backups

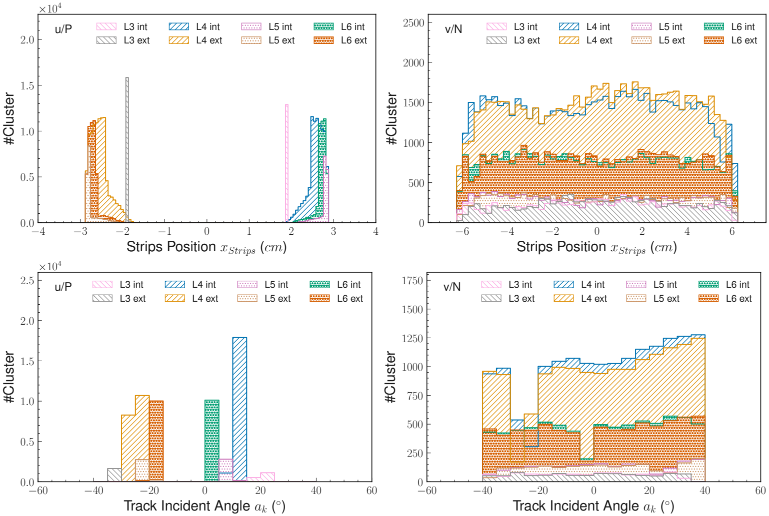
6. Backups: Measurement of the SVD spatial resolution

6.1 Geometrical correction
6.2 Detailled Overlapped method
6.3 Other methods
6.4 Methods comparison
6.5 Overlapped method discrepancies checks

7. Backups: Search for $\tau \to \mu\mu\mu$ LFV decays

7.1 Detailled reconstruction
7.2 Signal Region
7.3 Cut-based preselections
7.4 BDT checks
7.5 Old optimisation method
7.6 Checks efficiencies at all energies
7.7 Background composition
7.8 Number of good tracks descripency
7.9 Background estimation
7.10 Uncertainties
7.11 Trigger efficiency
7.12 Studies overlapping
7.13 Signal/Background comparison
7.14 Data/Simulation Sidebands
7.15 Data/Simulation Control Sample

Overlapping region



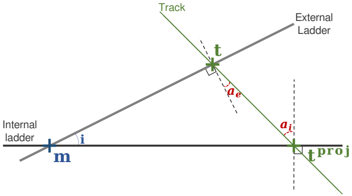
Geometrical correction

Project parallel to the tracks, the external residual on the internal ladder:

u/P Side:

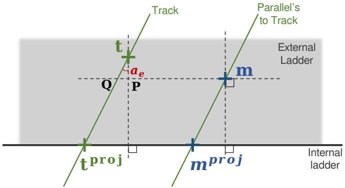
$$\Delta R = \frac{R_{\text{int}} - R_{\text{ext}} \times C}{\sqrt{1+C^2}}$$

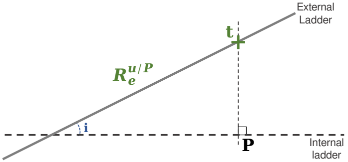
$$\text{with } C = \frac{\cos a_e}{\cos a_i}$$



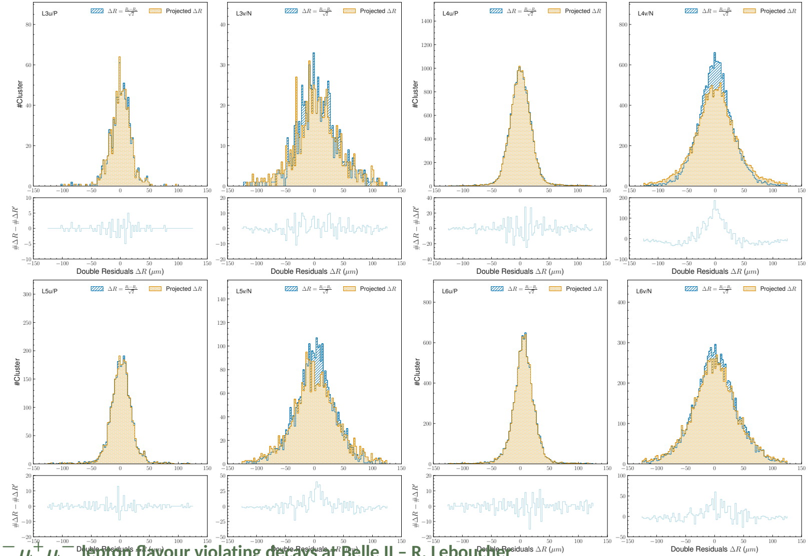
v/N Side:

$$\Delta R = \frac{R_{\text{int}} - \left(R_{\text{ext}} + \sin(i)\tan\left(a_e\right) R_{\text{ext}}^{u/P}\right)}{\sqrt{2}}$$





Geometrical correction



Spatial resolution – *Overlapped sensor method*

Method for estimate resolution with overlapping:

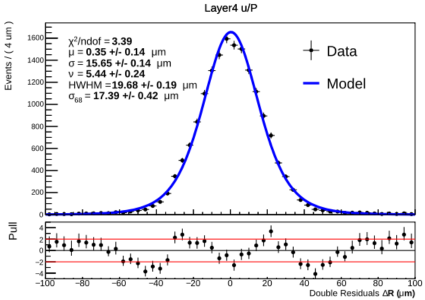
1. Apply geometrical correction factor on double residuals
2. Fit DeltaRes with a student's T-distribution:

$$T(X;\nu,\mu,\sigma) = \frac{\exp\left(\Gamma\left(\frac{\nu+1}{2}\right) - \Gamma\left(\frac{\nu}{2}\right)\right)}{\sigma\sqrt{\Pi\nu}} \left(1 + \frac{(X-\mu)^2}{\sigma^2\nu}\right)^{-\frac{\nu+1}{2}}$$

- Normalisation parameter N
- Number of degree of freedom ν
- Mean μ
- Vairance σ^2

3. The resolution is the `sigma-68` of the fitted student's T-model:

$$\sigma_{CL} = \sigma_{68}(T(X,\nu,\mu,\sigma))$$
$$= \frac{\chi_{84}(T(X,\nu,\mu,\sigma)) - \chi_{16}(T(X,\nu,\mu,\sigma))}{2}$$

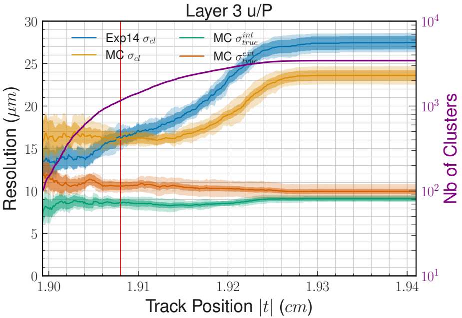


Method for estimate resolution uncertainties:

1. Vary fitted parameters (N, $\mu\nu\sigma$) within the fit uncertainties
2. Compute student's T-model with new parameters
3. Taking `sigma-68` resolution of this new model
4. Take as resolution uncertainty for each layer half the maximal variation of the recomputed `sigma-68`

Frame Title

- At the sensor edges some strips are masked
- Masked strips are not simulated and can introduce a bias in the track position
- Important to select a fiducial area
- The optimal cut value: compares the resolutions obtained from data and simulations at different t cut values



Cut Values:

Layer 3	$\lvert t^{u/P}\rvert < 1.908$ cm & $\lvert t^{v/N}\rvert < 5.9$ cm
Layer 4	$\lvert t^{u/P}\rvert < 2.82$ cm & $\lvert t^{v/N}\rvert < 5.9$ cm
Layer 5	$\lvert t^{u/P}\rvert < 2.84$ cm & $\lvert t^{v/N}\rvert < 5.9$ cm
Layer 6	$\lvert t^{u/P}\rvert < 2.82$ cm & $\lvert t^{v/N}\rvert < 5.9$ cm

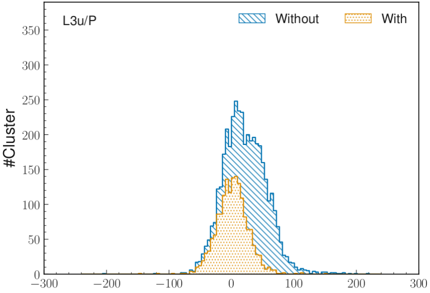


Figure 3: Double Residuals ΔR μm

Event by Event Method

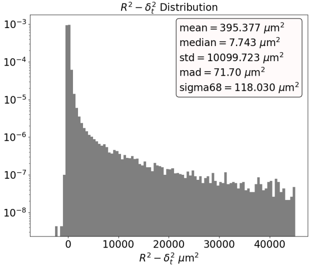
Event by event method:

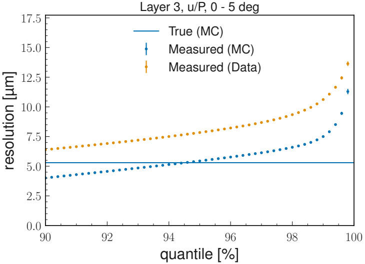
Subtract in quadrature the effect of the error on track extrapolation on residual:

$$\sigma_{cl}^2 = \langle R^2 - \sigma_t^2 \rangle_{trunc} \quad (1)$$

- Discrepancy between true and measured resolution on simulation
- Solved by optimising the quantile truncation on $R^2 - \sigma_t^2$ following:

$$\mathrm{FOM} = \frac{(\sigma_{true} - \sigma_{cl})^2}{(\Delta\sigma_{cl})^2} \quad (2)$$





Global Method

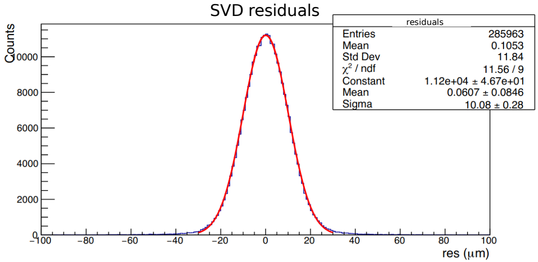
Global Method:

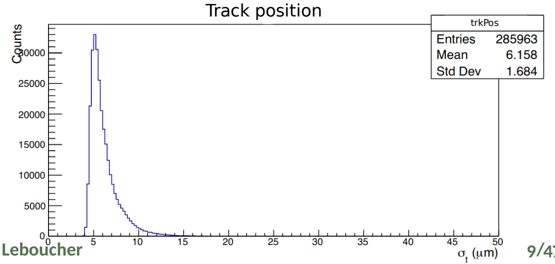
Based on Mean Absolute Deviation (MAD) as the best estimator:

- no optimization on MC is needed.

$$\sigma_{cl}^2 = \langle R^2 - \sigma_t^2 \rangle \simeq mad(R)^2 - median(\sigma_t)^2 - mad(\sigma_t)^2$$

- $mad(y) = 1.4826 \times median(|\ y - median(y)\ |)$
- mad and median more robust against outliers than standard deviation and mean.





Comparison of the different methods

- Different method developed to estimate the spatial resolution :

Event By Event	Global	Overlapped
Good estimation of spatial resolution Data/MC agreement Optimization on MC needed	Good estimation of spatial resolution No optimization needed Small Data/MC discrepancies	Marginally sensitive to Coulomb scattering Estimated resolution higher than expected

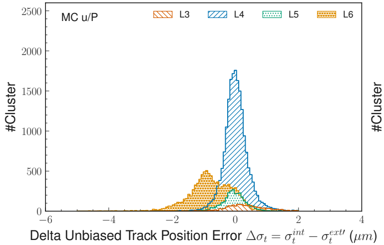
	Digital	EBE	Global	Overlapped
Layer $3u/P(\mu\mathrm{m})$	7	7	9	15
Layer $456u/P(\mu\mathrm{m})$	11	10	11	16 – 17
Layer $3v/N(\mu\mathrm{m})$	23	24	23	33
Layer $456v/N(\mu\mathrm{m})$	35	32	35	29 – 36

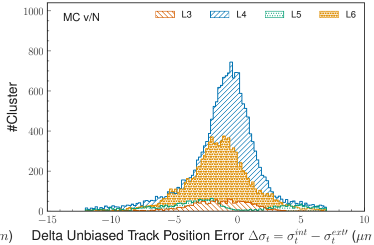
Overlapped method discrepancies checks

Check the assumption of errors on track position cancellation

Layer	u/P Side Median (μm)	u/P Side σ_{68} (μm)	v/N Side Median (μm)	v/N Side σ_{68} (μm)
3	$0.07 \pm 0.00 \pm 0.05$	$0.56 \pm 0.03 \pm 0.00$	$-1.19 \pm 0.04 \pm 0.20$	$1.37 \pm 0.07 \pm 0.10$
4	$-0.06 \pm 0.00 \pm 0.00$	$0.38 \pm 0.00 \pm 0.02$	$-0.92 \pm 0.01 \pm 0.21$	$2.71 \pm 0.03 \pm 0.35$
5	$-0.10 \pm 0.00 \pm 0.00$	$0.45 \pm 0.01 \pm 0.03$	$-2.66 \pm 0.05 \pm 1.82$	$10.23 \pm 0.27 \pm 2.16$
6	$-0.79 \pm 0.01 \pm 0.05$	$0.83 \pm 0.01 \pm 0.01$	$-1.83 \pm 0.022 \pm 0.14$	$3.18 \pm 0.05 \pm 0.32$

Non-null cancellation has a too small impact (1 μm) to explain the observed discrepancy





Overlapped method discrepancies checks

Check the assumption that $\sigma_{cl}^2 = Var(\Delta\varepsilon_m) = Var(\Delta R) - Var(\Delta\varepsilon_t)$ for the Layer 4 u/P side.

	Median (μm)	σ_{68} (μm)
$\Delta\varepsilon_m$	$0.032 \pm 0.00 \pm 0.00$	$9.01 \pm 0.09 \pm 0.29$
$\Delta\varepsilon_t$	$0.14 \pm 0.00 \pm 0.01$	$3.52 \pm 0.04 \pm 0.08$
ΔR	$-0.44 \pm 0.00 \pm 0.05$	$12.52 \pm 0.13 \pm 0.40$

We get:

$Var(\Delta\varepsilon_m) = Var(\Delta R) - Var(\Delta\varepsilon_t)$

$9.01 = \sqrt{12.52^2 - 3.52^2}$

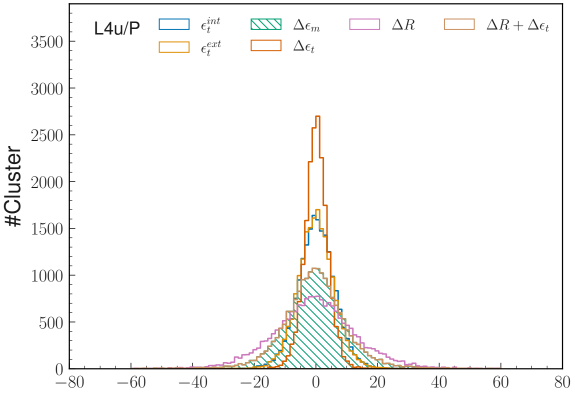
$9.01 \neq 12.01\ \mu\mathrm{m}$

While:

$Var(\Delta\varepsilon_m) = Var(\Delta R + \Delta\varepsilon_t)$

$9.01 = 9.01\ \mu\mathrm{m}$

ΔR and $\Delta\varepsilon_t$ appears to be correlated and so $\sigma_{cl}^2 \neq Var(\Delta R)$



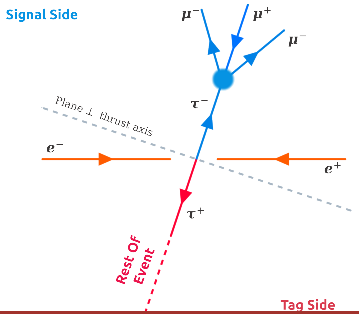
6. Backups: Measurement of the SVD spatial resolution
6.1 Geometrical correction
6.2 Detailled Overlapped method
6.3 Other methods
6.4 Methods comparison
6.5 Overlapped method discrepancies checks
7. Backups: Search for $\tau \to \mu\mu\mu$ LFV decays
7.1 Detailled reconstruction
7.2 Signal Region
7.3 Cut-based preselections
7.4 BDT checks
7.5 Old optimisation method
7.6 Checks efficiencies at all energies
7.7 Background composition
7.8 Number of good tracks descripency
7.9 Background estimation
7.10 Uncertainties
7.11 Trigger efficiency
7.12 Studies overlapping
7.13 Signal/Background comparison
7.14 Data/Simulation Sidebands
7.15 Data/Simulation Control Sample

Tau $\to$ 3 muons event reconstruction – *Untagged strategy*

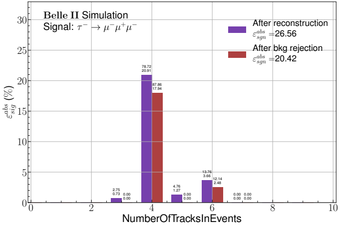
Belle II strategy

Reconstruct tau-pair events with an untagged strategy:

- **Signal side**: one τ decay into three muons
- **Untagged Opposite side**: **Rest of Event** (max 3 tracks) $\Rightarrow$ **signal efficiency** increase:
  - decays of τ_{tag} into 3-prong ($\mathcal{B} \sim 15\%$),
  - events with missing or additional track(s)...
- But **increase background contamination** mainly: $e^+e^- \to q\bar{q},\ \tau^-\tau^+,\ \mu^-\mu^+,\ \ell^-\ell^+\mu^-\mu^+$.



- 8 ab^{-1} of simulated backgrounds: $e^+e^- \to \tau^-\tau^+, q\bar{q}, , \ell\ell, \ell\ell XX, hh$ ($q = u, d, s, c, \ell = e, \mu;\ X = e, \mu, \tau, \pi, K, p, h = \pi, K$)
- 362 fb^{-1} of $\Upsilon(4S)$ data



Event requirements

Track requirements:

- Good tracks:
  - $|dz| < 3$ cm and $|dr| < 1$ cm
- Muon identification:
  - *muonID* > 0.5

Event requirements:

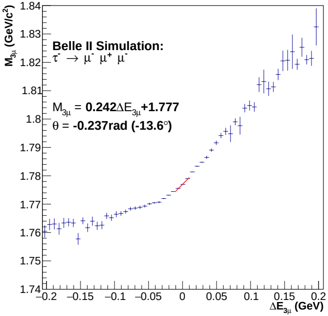
- Require the **3 leptons** in the same hemisphere wrt thrust axis
- Pass Level1 (L1) ECL and CDC **triggers**
- Pass **tauLFV skim**
- Correction on γ efficiency and energy
- Applying PID weights to correct data/Simulation efficiency differences

Full reconstruction table for 362 fb^{-1}

		weighted	ε_{sig}^{rel}	ε_{sig}^{abs}	$N_{bkg}^{weighted}$	$N_{\tau\tau}^{weighted}$	$N_{q\bar{q}}^{weighted}$	$N_{B\bar{B}}^{weighted}$	$N_{lowm}^{weighted}$
Reconstruction	train	231966.00	46.39	46.39	30157.59	1439.31	20569.29	422.09	7726.89
	test	232412.00	46.48	46.48	30483.66	1452.89	20864.05	440.92	7725.80
TruthMatch	train	231487.00	99.79	46.30	30157.59	1439.31	20569.29	422.09	7726.89
	test	231900.	99.78	46.38	30483.66	1452.89	20864.05	440.92	7725.80
21delta SR	train	204804.00	88.47	40.96	2220.60	14.03	953.69	0.72	1252.16
	test	205204.00	88.49	41.04	2225.49	13.39	965.73	1.45	1244.92
LML+HIE+CDC Trigger	train	193683.00	94.57	38.74	1994.35	9.86	922.01	0.72	1061.75
	test	194059.00	94.57	38.81	2013.99	8.51	936.49	1.45	1067.54
TauLFV Skim	train	190891.00	98.56	38.18	1994.35	9.86	922.01	0.72	1061.75
	test	191246.00	98.55	38.25	2013.99	8.51	936.49	1.45	1067.54
LID correction 0.5	train	171512.27	89.85	34.30	1803.70	7.98	830.10	0.70	964.92
	test	171752.08	89.81	34.35	1819.67	6.15	842.73	1.23	969.56

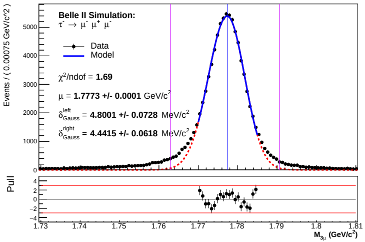
Signal Region in $M_{3\mu}$; $\Delta E_{3\mu}$ 2D plane

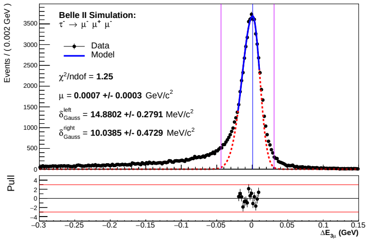
Measure resolutions to define **region** around **signal** centered at:
$M_{3\mu} \simeq 1.777\ \mathrm{GeV}/c^2$
and $\Delta E_{3\mu} = E_{3\mu}^{\mathrm{CM}} - \sqrt{s}/2 \simeq 0\ \mathrm{GeV}$



Fit $M_{3\mu}$ and $\Delta E_{3\mu}$

- **Asymmetric Gaussian**: on a limited range for the peak





	$M_{3\mu}$ MeV/c^2	$\Delta E_{3\mu}$ MeV
mean	1777.35±0.07	0.67±0.27
delta left	4.80±0.07	14.88±0.28
delta right	4.44±0.06	10.04±0.47
±20δ SR	[1.7533, 1.7996]	[−0.0737, 0.0509]
±20δ SR	[1.7456, 1.8031]	[−0.0708, 0.0491]

Cut-based preselections – *Muon identification selections*

Lepton MuonID Selection

The selection $\mu_{ID} > 0.95$ has an efficiency below 90% for each muon.

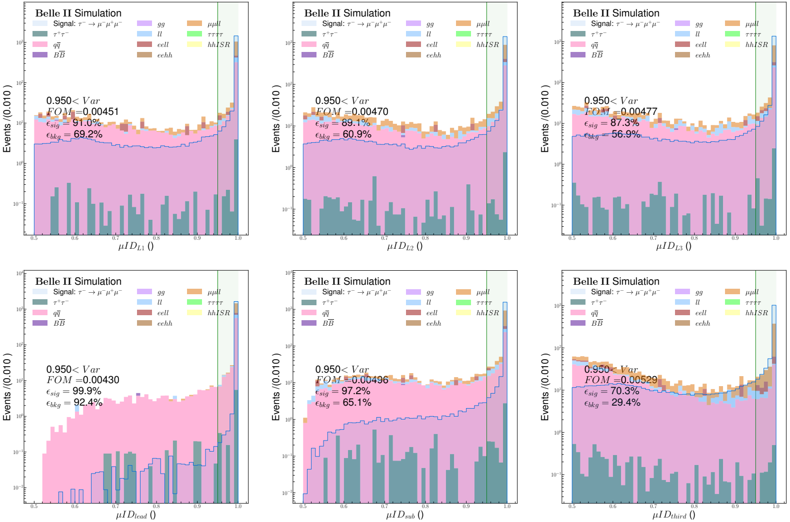
ε_{sig}^{rel}	ε_{sig}^{abs}	ε_{bkg}	N_{bkg} $362\ \text{fb}^{-1}$
68.08%	22.32%	24.28%	452.

Ranked MuonID Selection

- Rank the three Lepton μ_{ID}
- Cut on Third muonID is degrading $\varepsilon_{sig} < 70\%$
- Only cut on lead and sub muonID: $\varepsilon_{sig} > 95\%$

The muonID selection becomes less effective in rejecting background but also allows more background to train BDT.

ε_{sig}^{rel}	ε_{sig}^{abs}	ε_{bkg}	N_{bkg} $362\ \text{fb}^{-1}$
97.07%	32.12%	61.90%	1043.



Cut-based preselections – Additional selections

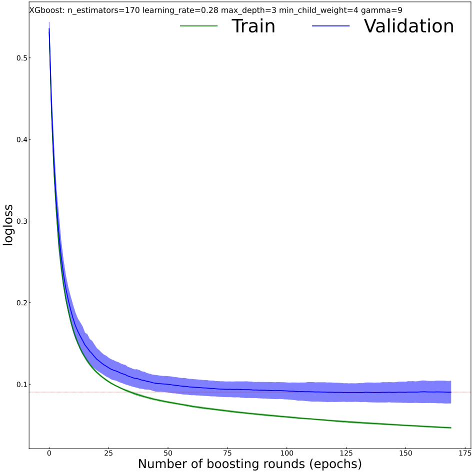
Numbers obtained with 4 ab^{-1} simulation train sample scaled to 362 fb^{-1} with additional requirements:

- 20δ Box Signal region
- LID selection: "$\mu_{ID}^{lkh,lead} > 0.95$ & $\mu_{ID}^{lkh,sub} > 0.95$ & $\mu_{ID}^{lkh,third} > 0.5$"

Name	Preselection	ε_{sig}^{rel} (%)	ε_{sig}^{abs} (%)	ε_{bkg}^{rel} (%)	N_{bkg}	$N_{\tau-pair}$	$N_{q\bar{q}}$	$N_{B\bar{B}}$	N_{lowm}	$\gamma\gamma$	$\ell\ell$	$ee\ell\ell$	$eehh$	$\mu\mu\ell\ell$	$\tau\tau\tau\tau$	hhISR
Reference	$0.3 < \theta_{miss}^{CM} < 2.7$	96.88	31.11	89.99	938.82	3.08	287.52	0.	648.22	0.	49.29	44.32	0.	554.61	0.	0.
Set 1	$0.3 < \theta_{miss}^{CM} < 2.7$ $0.89 < Thrust < 0.97$	95.48	30.67	30.83	321.64	2.96	270.87	0.	47.82	0.	8.56	9.63	0.	29.63	0.	0.
Set 2	$0.3 < \theta_{miss}^{CM} < 2.7$ $0.935 < ROE\tau$ thrust<0.95	96.35	30.94	61.78	644.50	2.58	244.94	0.	396.97	0.	13.25	29.66	0.	354.07	0.	0.
Set 3	$0.3 < \theta_{miss}^{CM} < 2.7$ $E_{vis}^{CM} < 10.$	90.54	29.08	14.89	155.30	2.98	127.81	0.	24.52	0.	6.91	12.62	0.	4.98	0.	0.
Set 4	$0.3 < \theta_{miss}^{CM} < 2.7$ $E_{miss}^{CM} > 0.6$	90.22	28.98	14.69	153.29	2.91	125.85	0.	24.52	0.	6.91	12.62	0.	4.98	0.	0.
Set 5	$0.3 < \theta_{miss}^{CM} < 2.7$ $p_{miss}^{T,CM} > 0.4$	91.12	29.26	15.89	165.74	2.77	135.08	0.	27.90	0.	7.11	15.50	0.	5.28	0.	0.
Set 6	$0.3 < \theta_{miss}^{CM} < 2.7$ $M_{ROE(masked)} < 2.2$ $-5. < \Delta E_{ROE(masked)} < -0.2$	90.76	29.15	16.49	172.08	2.62	106.08	0.	63.39	0.	14.34	22.93	0.	26.11	0.	0.

BDT overtraining health

Train & Validation Logloss in the function of the boosting rounds to visualize the training behaviour:



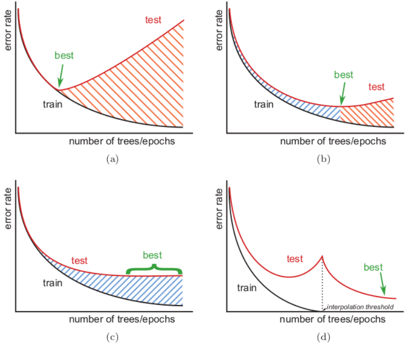
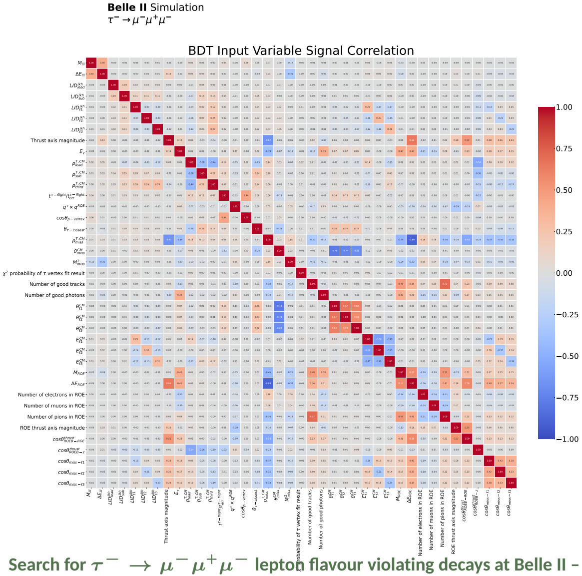
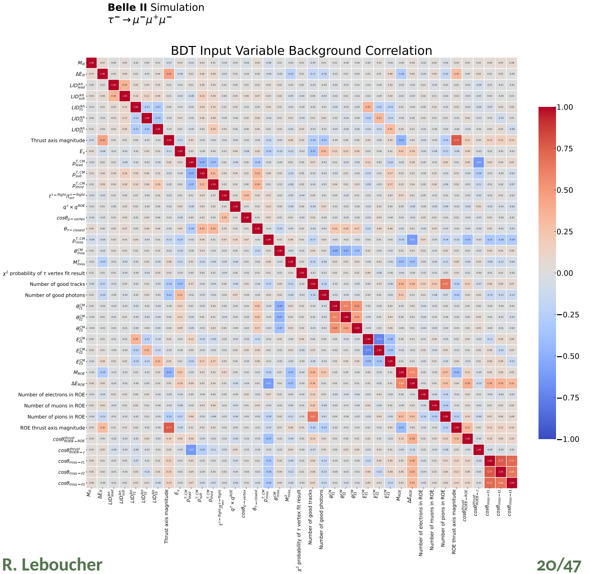


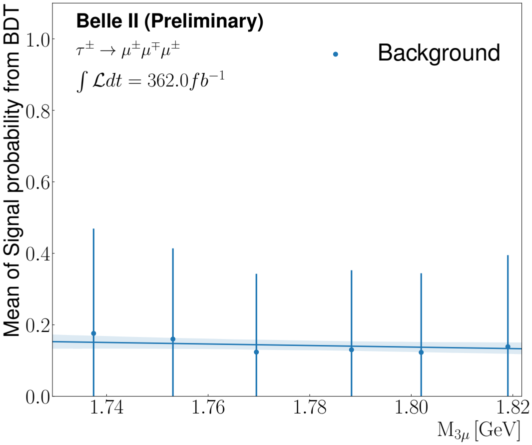
Fig. 3. Overtraining estimation using the error rate as a function of the number of trees (for boosted decision trees) or epochs (for neural networks). Black curves are measured on the training sample and red curves on the validation sample. The optimal classifier corresponds to the "best" label. The hatched areas represent overtraining: beneficial in blue (but underfitting), detrimental in orange (overfitting). (a) Typical curves, with the best model at the minimum of the testing curve, and overfitting beyond with decrease of performance. (b) The best model is overtrained but still improves performance. (c) Typical curves for boosted decision trees with flattening testing error rate: all models in the flat area perform equally well despite increasing overtraining. (d) Interpolation regime: the best classifier is obtained after the training error has reached zero.

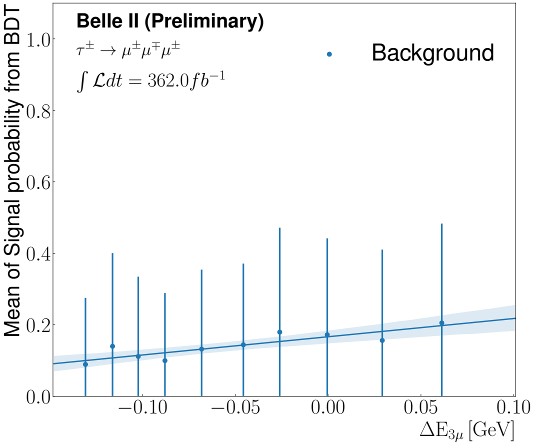
Variables Correlations





Signal shape sculpting check





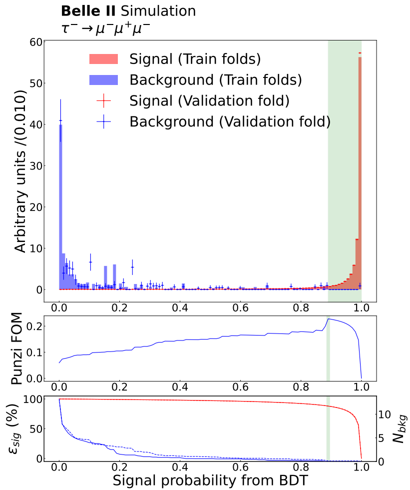
Signal region optimisation

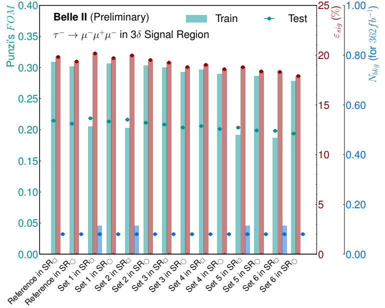
Previous optimisation of the BDT output and SR:

1. Fit resolution δ of SR variables and define asymmetric Ellipse and Rectangle
2. Find selection on BDT output maximising Punzi's FOM in validation folds (from 10-folding)
3. Find the best SR form (rectangle or ellipse) at 3δ and the best preselection in 4 ab^{-1} train

Final selection
$\mu ID_{lead} > 0.95$
$\mu ID_{sub} > 0.95$
$0.3 < \theta^{CM}_{miss} < 2.7$
$0.89 < thrust < 0.97$
$p^{BDT}(Signal) > 0.89$
3δ ellipse region

	Train	Test
N_{bkg}	0.00	0.08
ε_{sig} (%)	19.61	19.70
Punzi's FOM	0.131	0.131

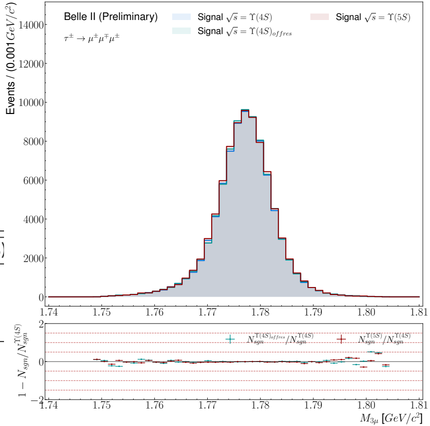


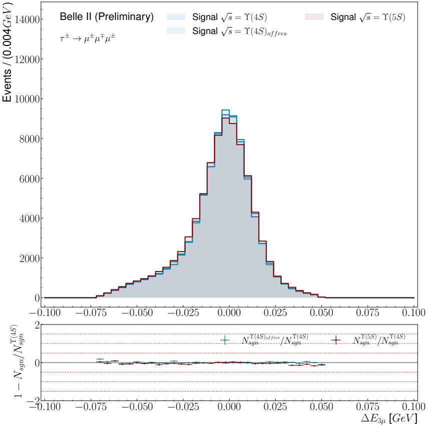


Signal efficiency at different energies

Apply the same $\tau^- \to \mu^-\mu^+\mu^-$ reconstruction and background rejection on signal samples produced at different energies:

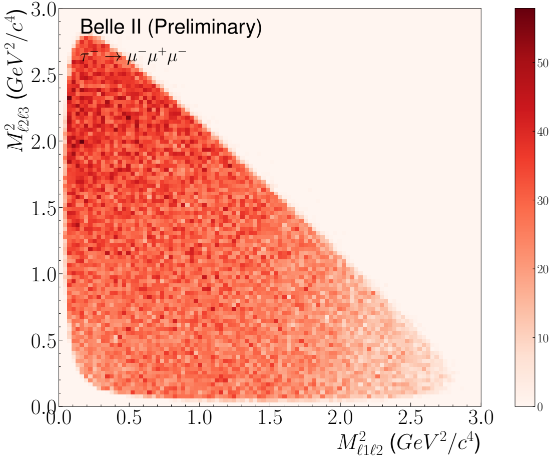
Signal Produced	ε_{sig}^{abs} (%)
$\Upsilon(4S) = 10.58$ GeV	20.42
Off-resonance -60 MeV	20.54
10.81 GeV	20.72



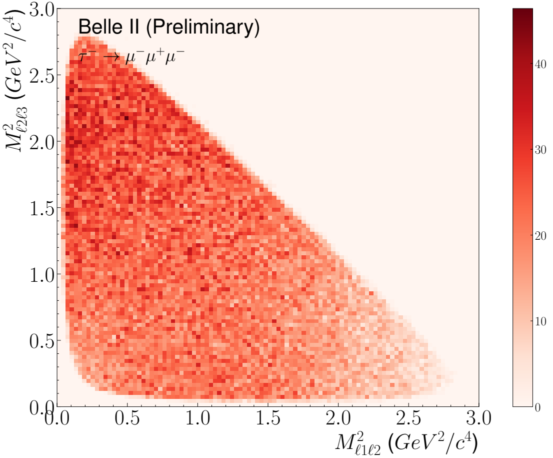


Phase space distribution

After reconstruction:



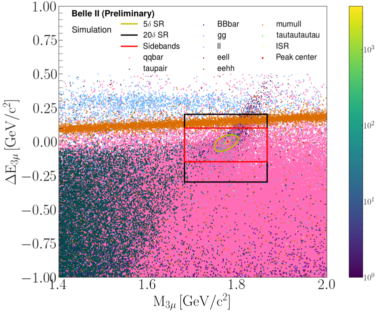
After background rejection:



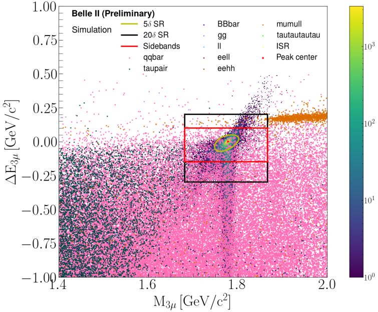
Muons with same charge as the τ : $\ell_{1,3}$; opposite charge ℓ_2

Background composition

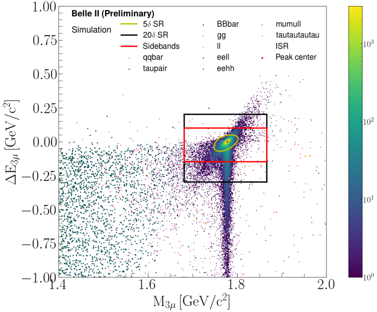
After reconstruction:



After preselection:



After background rejection:



Background composition

After preselection cut:

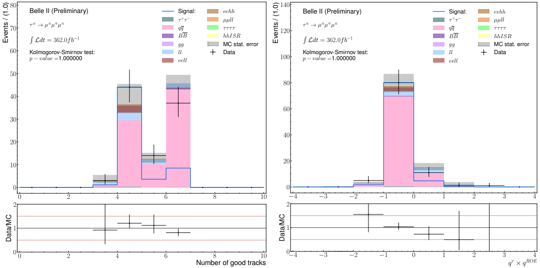
$mcPDG_{\ell 1}$	$mcPDG_{\ell 2}$	$mcPDG_{\ell 3}$	type	sample	occurrence
π	π	π	qqbar	uubar	19
π	π	π	qqbar	ddbar	7
K	π	π	qqbar	ssbar	3
K	π	π	qqbar	uubar	2
π	π	π	qqbar	ssbar	2
π	K	π	qqbar	ccbar	2
μ	π	π	qqbar	uubar	1
π	K	K	qqbar	uubar	1
K	π	π	qqbar	ddbar	1
π	NaN	π	qqbar	ddbar	1
π	π	NaN	qqbar	uubar	1
π	K	π	qqbar	ssbar	1
μ	NaN	π	qqbar	ccbar	1
π	π	K	qqbar	uubar	1
π	π	μ	qqbar	uubar	1
π	μ	π	qqbar	ddbar	1
K	K	K	qqbar	ssbar	1

After BDT Cut:

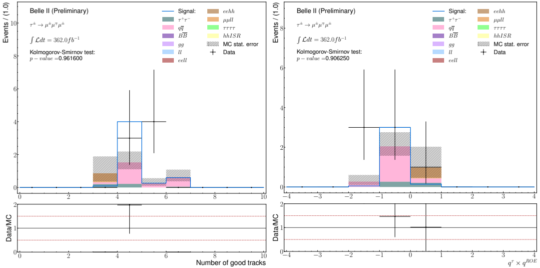
$mcPDG_{\ell 1}$	$mcPDG_{\ell 2}$	$mcPDG_{\ell 3}$	type	sample	occurrence
K	π	π	qqbar	ssbar	1

Number of good tracks descripency – *Extended sidebands*

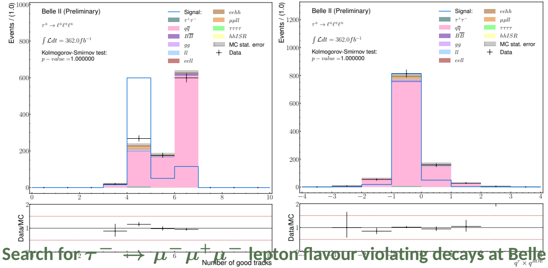
Plots in sidebands after preselection:



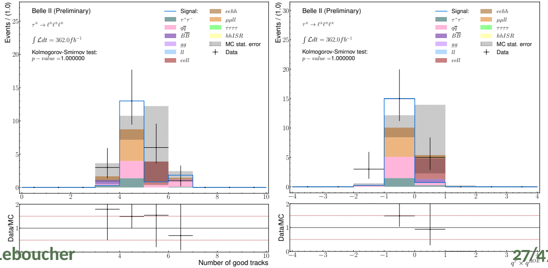
Plots in sidebands after preselection and BDT cut:



Plots in extended-sidebands after preselection:

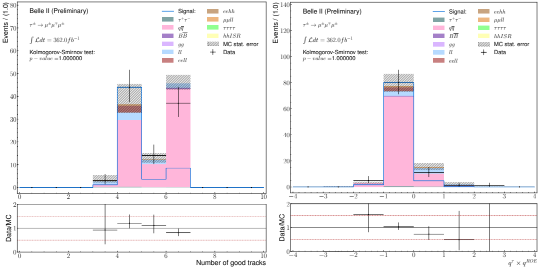


Plots in extended-sidebands after preselection and BDT cut:

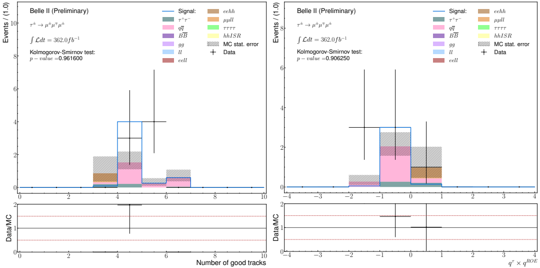


Number of good tracks descripency – *Revert LID*

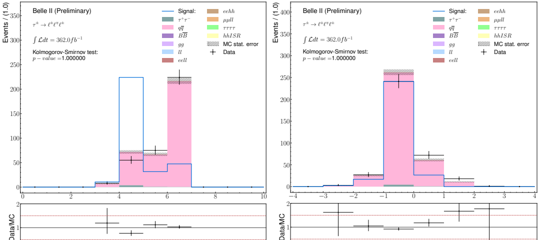
Plots in sidebands after preselection:



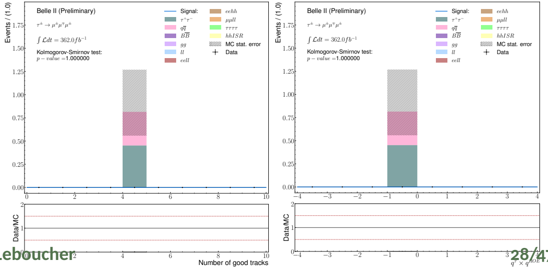
Plots in sidebands after preselection and BDT cut:



Plots in extended-sidebands after preselection revert muonID:

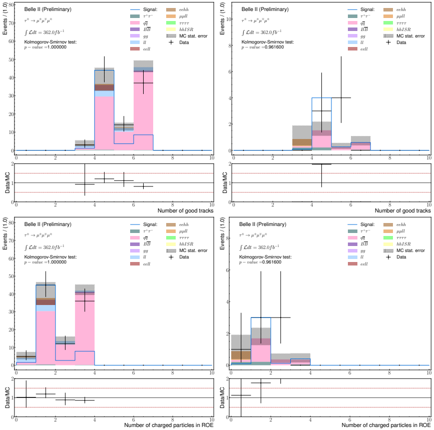


Plots in extended-sidebands after preselection and BDT cut and revert muonID:

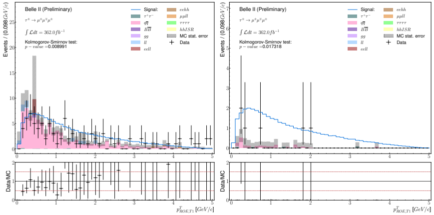


Number of good tracks descripency

Plots in sidebands after preselection (left) and BDT cut (right):



Plots in sidebands after preselection (left) and BDT cut (right):



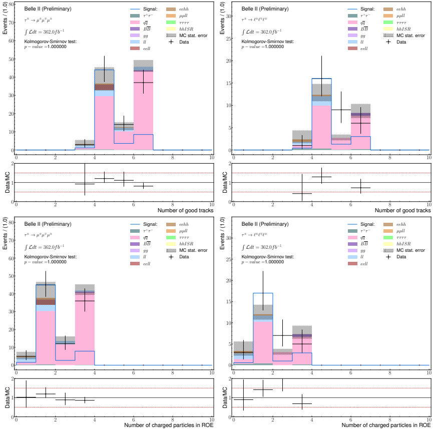
Number of good tracks descripency

Data events in sidebands with exactly 5 tracks after applying preselections and BDT cut at 0.9

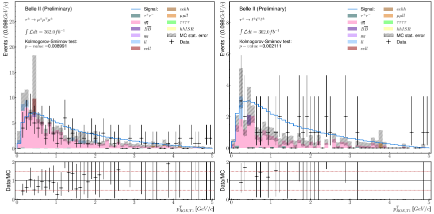
Colum	nGoodTracks	SigBDTProbability	T1_roe_PDG	T2_roe_PDG	T1_roe_pt	T2_roe_pt	T1_roe_electronID	T2_roe_electronID
2859	5	0.951149002	-11	211	0.232124849	0.444896184	0.580009369	0.124428372
4959	5	0.968264543	13	-211	0.769164055	1.327150574	3.12E-23	0.000176022
14023	5	0.900836594						
16559	5	0.968536004	11	11	0.267997908	0.646096216	0.999728114	0.989322256

Number of good tracks descripency

Plots in sidebands after preselection (left) and BDT cut 0.2 (right):



Plots in sidebands after preselection (left) and BDT cut 0.2 (right):

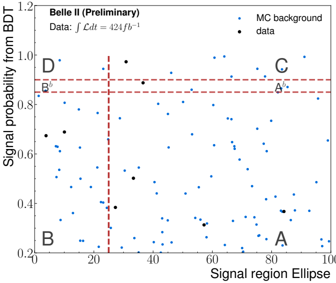


Number of good tracks descripency

Data events in sidebands with exactly 5 tracks after applying preselections and BDT cut at 0.2

Colum	nGoodTrac	SigBDTProba	T1_roe_PDG	T2_roe_PDG	T1_roe_pt	T2_roe_pt	T1_roe_electronID	T2_roe_electronID
1368	5	0.313182198	-13		4.532335049		1.57E-20	
2401	5	0.485624406	211	-211	1.229043729	0.27551261	0.001468953	0.060522662
2507	5	0.726254858	11	-211	2.452555429	0.18460046	0.99948458	0.293043059
2859	5	0.951149002	-11	211	0.232124849	0.44489618	0.580009369	0.124428372
3894	5	0.584622681	211	-321	0.623276238	1.00199294	0.0039538	3.93E-06
4959	5	0.968264543	13	-211	0.769164055	1.32715057	3.12E-23	0.000176022
7531	5	0.4313305	211	2212	2.109725136	1.29809413	0.325301955	0.002274573
14023	5	0.900836594						
16559	5	0.968536004	11	11	0.267997908	0.64609622	0.999728114	0.989322256

Background estimation in the signal region – *Data-driven ABCD method*

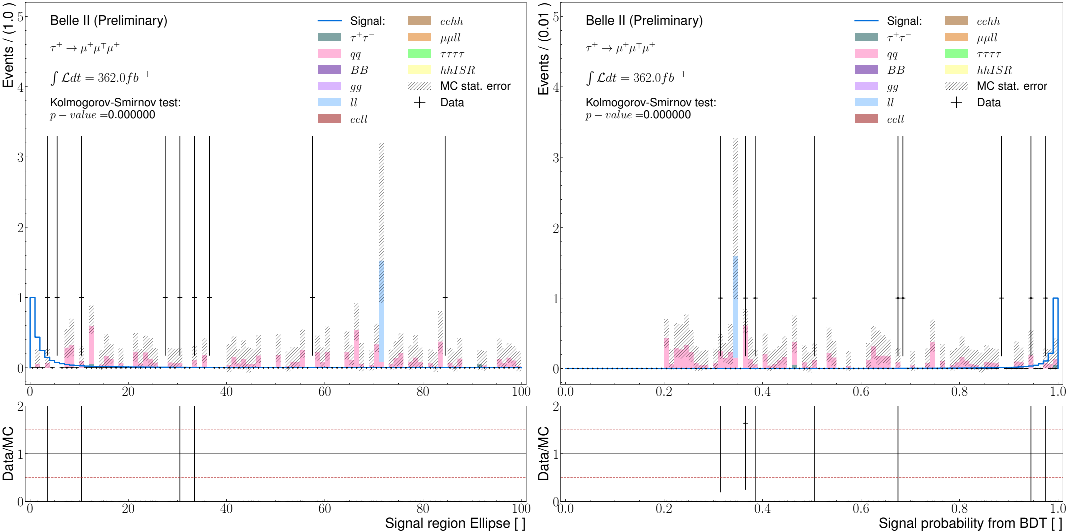


Data	
N_A	$4.00^{+3.16}_{-1.91}$
N_B	$2.00^{+2.64}_{-1.29}$
$R_{B/A}$	$0.50^{+0.77}_{-0.40}$
N_C	$1.00^{+2.30}_{-0.83}$
$N_D^{expected}$	$0.50^{+1.38}_{-0.58}$

Simulation	
N_A	$6.47^{+1.91}_{-0.95}$
N_B	$2.23^{+0.65}_{-0.42}$
$R_{B/A}$	$0.34^{+0.14}_{-0.08}$
$R_{D/C}$	$0.32^{+0.58}_{-0.28}$
N_C	$0.49^{+0.51}_{-0.23}$
$N_D^{expected}$	$0.17^{+0.19}_{-0.09}$
N_D	$0.16^{+0.24}_{-0.12}$

	N_{sgn}	ε_{sig}	$N_{sgn}^{expBelleBR}$	N_{bkg}
Zone A	$1913.20^{+44.75}_{-43.74}$	$0.38\%^{+0.01}_{-0.00}$	$0.06^{+0.00}_{-0.00}$	$6.47^{+1.91}_{-0.95}$
Zone B	$8964.92^{+95.69}_{-94.68}$	$1.79\%^{+0.02}_{-0.02}$	$0.29^{+0.00}_{-0.00}$	$2.23^{+0.65}_{-0.42}$
Zone C	$10177.07^{+101.88}_{-100.88}$	$2.04\%^{+0.02}_{-0.02}$	$0.33^{+0.00}_{-0.00}$	$0.49^{+0.51}_{-0.23}$
Zone D	$102103.70^{+320.54}_{-319.54}$	$20.42\%^{+0.06}_{-0.06}$	$3.34^{+0.01}_{-0.01}$	$0.16^{+0.24}_{-0.12}$

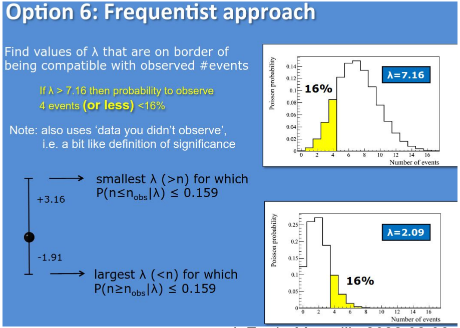
Background estimation in the signal region – *Data-driven ABCD method*

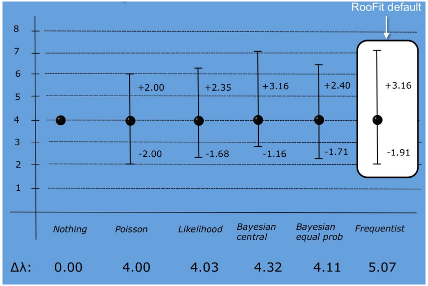


Statistical uncertainties

Asymmetric error bars on data yields ("vanilla case")

- after discussion at past tau meeting, we assign asymmetric uncertainties to yields in **data and MC**
  - before computed as symmetrical Poisson uncertainties sqrt(N), for N entries in bin_i
- adopt frequentist approach and find iteratively λ1, λ2 so that $P(n<=N_{bin}|\lambda1) <=0.16$ and $P(n>=N_{bin}|\lambda2) <= 0.16$





- in each bin error bars are defined as:
  - $err\_stat\_up = \lambda1 - N_{bin}$,
  - $err\_stat\_low = N_{bin} - \lambda2$

7

L.Zani - Marseille 2023.03.06 - Tau to lepton phi unboxing

Systematics uncertainties – *Trigger*

Strategy

- Measure trigger efficiency on $\tau \to \pi\pi\pi\nu$ Control sample
- Relative efficiencies w.r.t. orthogonal lines (ECL Vs CDC)

$$\varepsilon_{ECL} = \frac{ECL\ AND\ CDC}{CDC} \qquad \varepsilon_{CDC} = \frac{ECL\ AND\ CDC}{ECL}$$

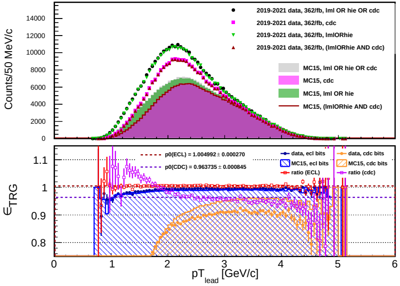
- Assign systematic uncertainty as the p^T discrepancy weighted to trigger efficiency measured with TSIM on signal simulation as $N_{TRG} / N_{Reconstruction}$

Result:

$$\sigma_{TRG} = \varepsilon^{TSIM}_{ECL} \cdot \delta_{ECL} + \left(\varepsilon^{TSIM}_{tot} - \varepsilon^{TSIM}_{ECL}\right) \cdot \delta_{CDC}$$

$$= 0.80 \cdot \frac{0.005 + 0.004}{2} + (0.95 - 0.8) \cdot \frac{0.04 + 0.05}{2}$$

$$= 1\%$$

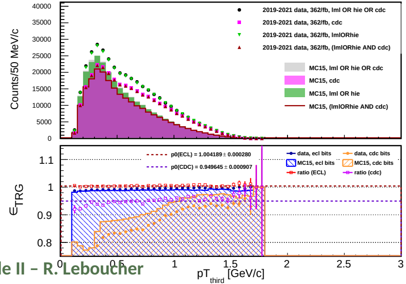


p^T_{lead}:

$$\delta_{TRG} = \left|1 - \frac{\varepsilon^{Data}_{TRG}}{\varepsilon^{MC}_{TRG}}\right|$$

$$\delta_{CDC} = 0.0363$$

$$\delta_{ECL} = 0.005$$



p^T_{third}:

$$\delta_{TRG} = \left|1 - \frac{\varepsilon^{Data}_{TRG}}{\varepsilon^{MC}_{TRG}}\right|$$

$$\delta_{CDC} = 0.0504$$

$$\delta_{ECL} = 0.0042$$

Systematics uncertainties – *LID & Momentum scale*

Systematics from LID on signal efficiency:

- Track LID weights w_{LID} and variation $w_{LID}^{stat,sys}$ are given to correct simulations. A global LID weight is defined as:

$$w_{LID}^{stat,sys} = \prod_{i=1}^{3} w_{LID,\ell i} \times w_{LID,\ell i}^{stat,sys}$$

- Compute the number of signals after background rejection for each weight:

$$\sigma_{LID}^{stat,sys} = \sqrt{\left(\frac{n_{LID} - n_{LID}^{stat,sys}}{n_{produced}}\right)^2} \times \frac{1}{\varepsilon_{sig}^{abs}}$$

Results:

	σ_{LID}^{abs}	σ_{LID}^{rel} (%)
stat ↓	0.0042	2.08
stat ↑	0.0047	2.29
sys ↓	0.0007	0.36
sys ↑	0.0011	0.55

Total error: $\sqrt{{\sigma_{LID,stat}^{rel}}^2 + {\sigma_{LID,sys}^{rel}}^2} = 2.36\%$

Systematics from momentum scale on expected data:

- Corrections and up/down variations are provided to fix the bias caused by the magnetic field map
- Reconstruction run on data for each momentum scale working point
- Compute the expected data after preselection with the ABCD method

Results:

	$N^{extected}$	σ_{LID}^{abs}	σ_{LID}^{rel} (%)
True momentum scale	80	0.0	0.00
Low momentum scale	78	2.0	2.50
Hig momentum scale	81	1.0	1.25

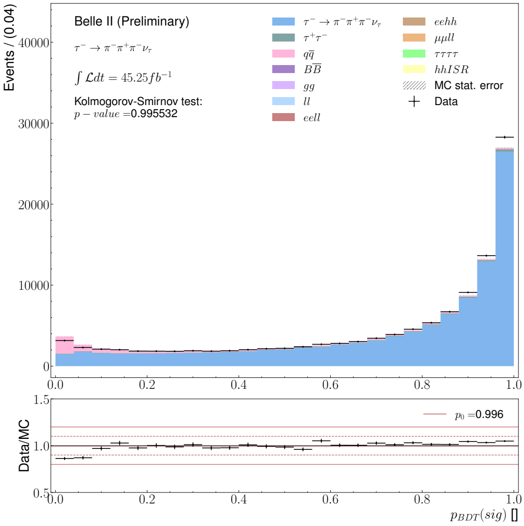
Systematics uncertainties – *BDT*

Compute control sample in the control sample:

- ε^{CS}_{data}: relative data efficiency on the BDT output subtracting non 3π events estimated in MC
- $\varepsilon^{CS}_{3\pi}$ relative data efficiency on the BDT output

At relative, a BDT output threshold with relative efficiency equals the relative efficiency (81.5%) on $\tau^- \rightarrow \mu^-\mu^+\mu^-$ at 0.9.
Look at the difference between control sample data ε^{CS}_{data} and simulation $\varepsilon^{CS}_{3\pi}$ efficiencies relative to unity is computed as:

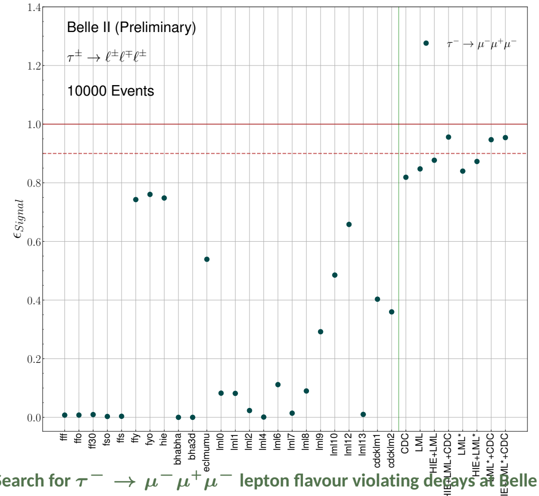
$$\delta = \left|1 - \frac{\varepsilon^{CS}_{data}}{\varepsilon^{CS}_{3\pi}}\right| = 1.5\%$$



Trigger efficiency

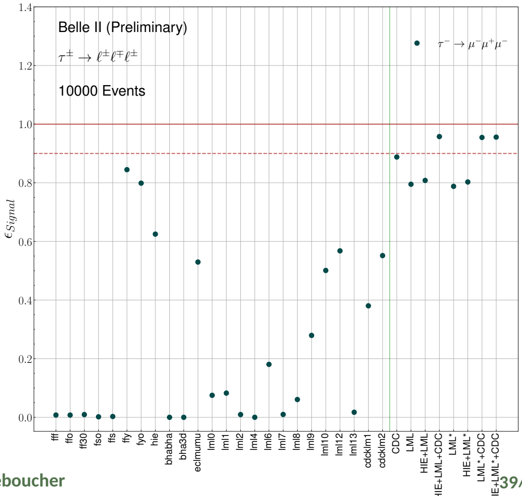
Trigger efficiency after the reconstruction:

$$\varepsilon_{TRG}^{sgn} = \frac{N_{Pass\ trigger}}{N_{Reconstructed}} = 95.26\%$$



Trigger efficiency after the background rejection cuts:

$$\varepsilon_{TRG}^{sgn} = \frac{N_{Pass\ trigger}}{N_{PassingSelection}} = 95.57\%$$



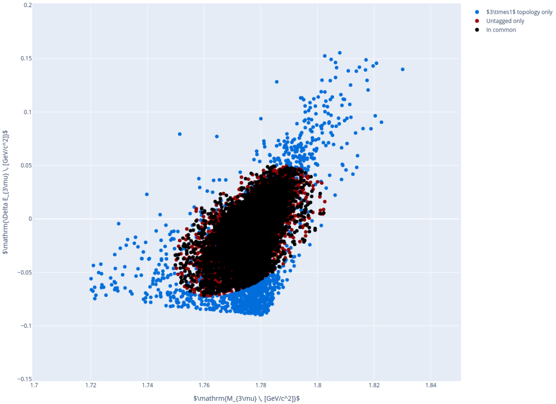
Signal overlapping between the two methods

Signal events retained after the full selection for each strategy:

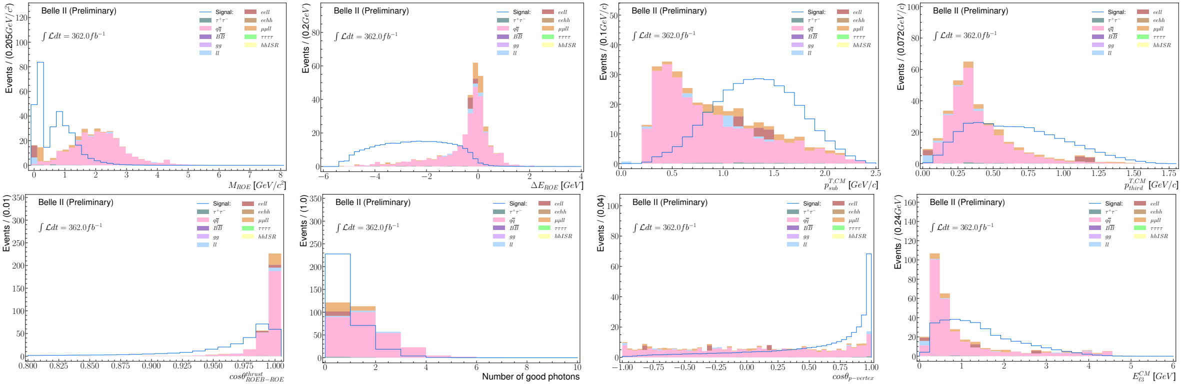
	Produced	Passing Selection	In common	Ratio
3X1 topology	100000	17489	14876	85.06%
Untagged	100000	23209	14876	64.10%

Signal events retained after the full selection of each strategy (only 1 tracks in the ROE):

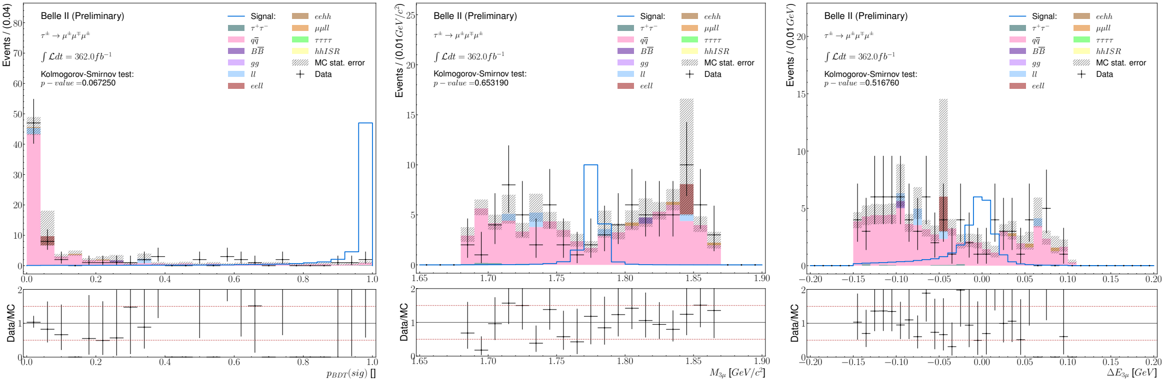
	Produced	Passing Selection	In common	Ratio
3X1 topology	100000	17489	14876	85.06%
Untagged	100000	20422	14876	72.84%



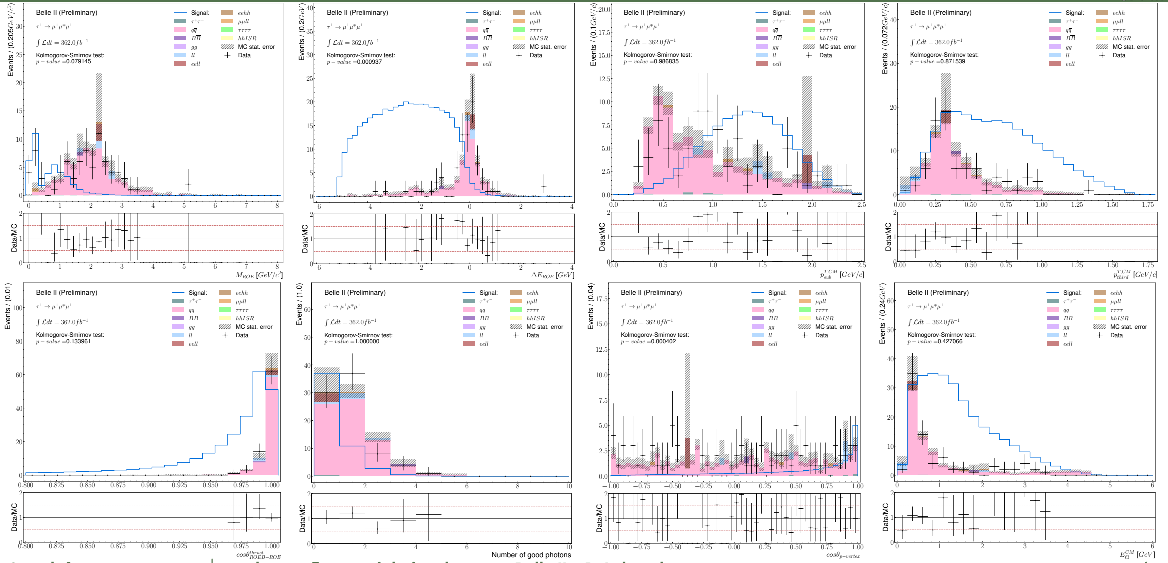
Signal/Background comparison – *BDT Variables*



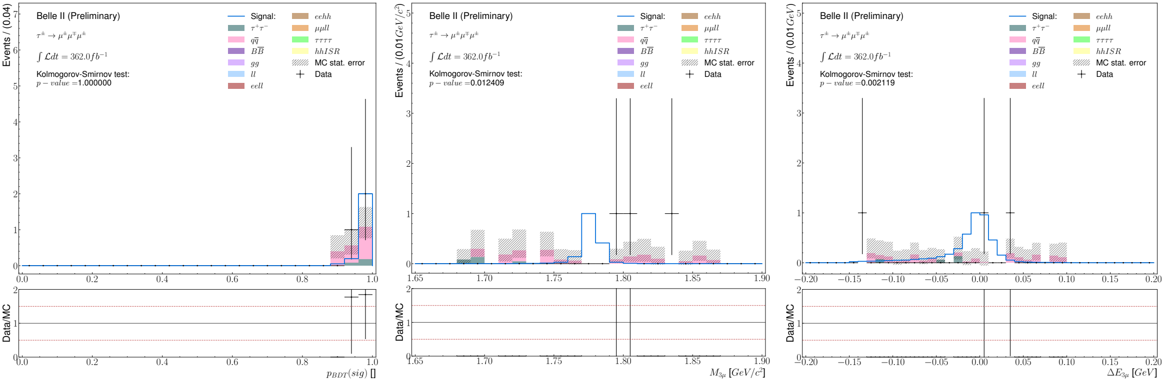
Data/Simulation Sidebands – *BDT Variables*



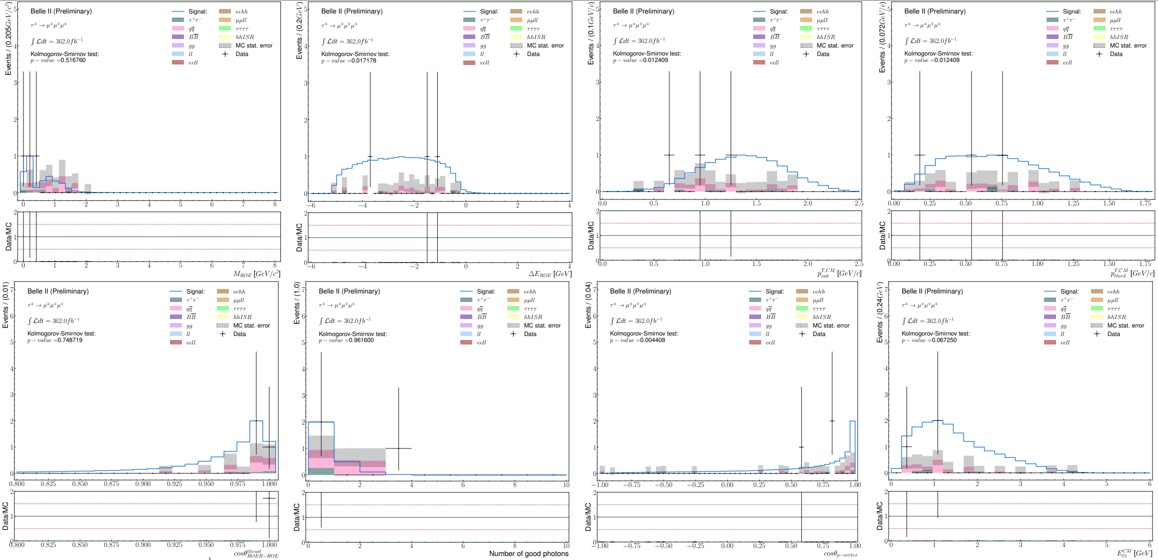
Data/Simulation Sidebands – *BDT Variables*



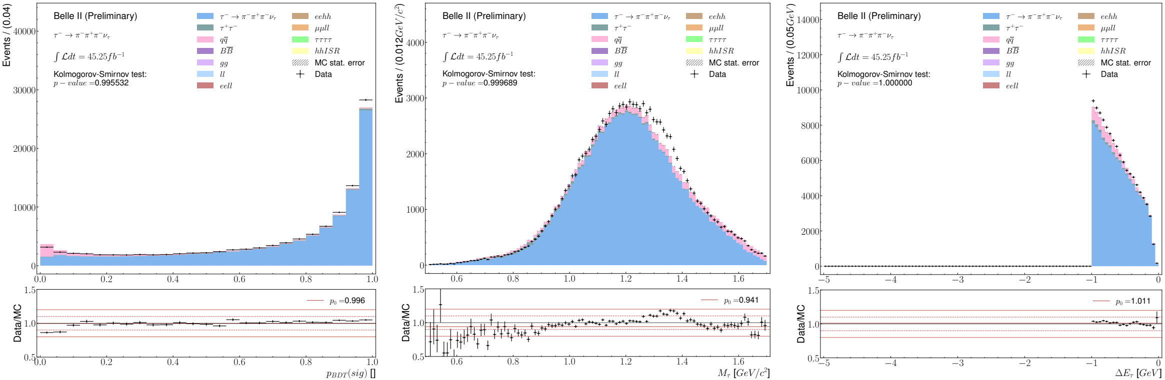
Data/Simulation Sidebands – *BDT Variables*



Data/Simulation Sidebands – *BDT Variables*



Data/Simulation Control Sample – *BDT Variables*



Data/Simulation Control Sample – *BDT Variables*