Search for the multilepton final state @LHCB
\[ B^+ \rightarrow K^+ J/\psi \gamma^*(\rightarrow \mu^+\mu^-) \]
Construction of knowledge on elementary particles and their interaction
= constant interconnection between theory buildings and experimental observations

Observing new phenomena is a nice goal... but one must first make them happen!

LHC
27 km circumference
13 TeV proton-proton collisions
40 MHz collision rate

Search for multilepton final state $B^+ \rightarrow J/\psi K^+ \gamma^* (\rightarrow \mu^+ \mu^-)$ @LHCb
A B-HADRON SPECIALISED DETECTOR

- Single-arm, forward spectrometer, $2 < \eta < 5$
- $b\bar{b}$ production peaks (back-)forward
- Aiming for $\sqrt{s} \gtrsim 13$ TeV for 3rd data taking run (planned 2022-2024)

MAJOR UPGRADE FOR RUN 3 (2022)

- Event readout rate $\times 40$
- Upgrade of ~all detectors
- Full software trigger, with increased computing power

LARGELY DEVELOPED @ EPFL!

Search for multilepton final state $B^+ \rightarrow J/\psi K^+ \gamma^*(\rightarrow \mu^+ \mu^-)$ @LHCb
Ok... but why?

- FLAVOUR ANOMALIES IN SEMILEPTONIC B-DECAYS

\[ b \to c\ell \nu_\ell \] shows anomaly @ 3\( \sigma \)
using yields of \( \tau \) vs. light leptons \((\mu, e)\) [HFLAV summary, 2019]

\[ b \to u\ell \nu_\ell \] can be probed in \( B^+ \to \ell \nu_\ell \) decays
Only measurement [Belle Collab, 2004]
\[
\frac{BR(B^+\to\tau\nu_\tau)_{\text{exp}}}{BR(B^+\to\tau\nu_\tau)_{\text{SM}}} = 1.31\pm0.27 @1.5\sigma
\]

- HOW CAN LHCb CONTRIBUTE? \( B^+ \to \ell \nu_\ell \) is difficult or impossible at LHCb due to
  single-track signature and helicity suppression ☹️
However, adding an Final State Radiation (FSR) tagging with \( \gamma^* \to \mu^+\mu^- \) allows to:
  - increase BF to available dataset \((-10^{-8})\) [Bharucha et al., 2021]
  - reconstruct a decay vertex of \( B^+ \to \mu^+\nu_\mu \gamma^* (\to \mu^+\mu^-) \) and \( B^+ \to e^+\nu_e \gamma^* (\to \mu^+\mu^-) \)
@LHCb and probe Lepton Flavour Universality by comparing their yields

Experimental challenge \️→ use of very soft muons

Sonia Bouchiba
Search for multilepton final state \( B^+ \to J/\psi K^+ \gamma^* (\to \mu^+\mu^-) \) @LHCb
Experimental challenge 💪 → use of very soft muons

**Goal:** Develop a dedicated treatment of soft muon selection and identification with a reference channel $B^+ \rightarrow K^+ J/\psi \gamma^*(\rightarrow \mu^+\mu^-)$

**Benefits:**

- visible with available dataset (though have never been observed so far)

- usage of $J/\psi \rightarrow \mu^+\mu^-$ data stream allows to get an unbiased sample $\gamma^* \rightarrow \mu^+\mu^-$

- imposes minimum $P$ of 3 GeV on soft muons → make sure they reach the muon stations and can be identified

![Momentum $P$ vs number of entries](image1)

![Transverse momentum $PT$ vs number of entries](image2)
The challenge: dealing with different backgrounds!

**MISIDENTIFIED MUONS** → 1 or 2 hadrons ($K$ or $\pi$) decaying to muons, tricking the muon stations, are misID as muons.

**RESONANT DECAYS** → $\omega, \phi, \rho \rightarrow \mu^+\mu^-$ with $B \sim 10^{-8}$ and $\chi_{c1}, \chi_{c2} \rightarrow J/\psi \mu^+\mu^-$ contributions;

**COMBINATORIAL BACKGROUND** → Random tracks put together by mistake by the selection algos.

**PARTIALLY RECONSTRUCTED DECAYS** → Non-signal decays with an extra particle not entering the selection (e.g. $B^+ \rightarrow J/\psi K^+\eta(\rightarrow \mu^+\mu^+\gamma)$, with missed $\gamma$).

**Dominant background**
To be studied with a dedicated Boosted Decision Tree (BDT) only with Particle Identification (PID) vars.

**Peaking backgrounds**
(dangerous!) Add in fit model / veto.

**Exponential distribution**
Supress with dedicated BDT.

**Neglect / add in fit model**
The challenge: dealing with different backgrounds!

First look at a small subset of data with the preselection applied:

- Misidentified muons contributed by 
  - Misidentifications 
  - Peaky backgrounds 
  - Resonance decays

- Double misidentification peak on top of signal $\odot$
  - $B^+ \rightarrow K^+ J/\psi \pi^+\pi^-$
  - Where both pions are misidentified as muons

- Double MisID peak on top of signal $\odot$
  - $B^+ \rightarrow K^+ J/\psi \pi^+\pi^-$
  - Where both pions are misidentified as muons

- MisID + Partially reconstructed decays
  - $B^+ \rightarrow K^+ J/\psi \pi^+\pi^- X$

- First look at a small subset of data with the preselection applied:
  - $M(K^+ J/\psi \mu^+\mu^-) [\text{MeV}/c^2]$
Combinatorial background Boosted Decision Tree (BDT)

**Goal:** have a very efficient selection with ~100% signal efficiency

**TRAINING VARIABLES**
B-meson's (mother)
- impact parameter
- direction angle of flight
- quality of the constrained track fit
- position of decay (secondary vertex)

J/ψ flight distance
Minimum transverse momentum of J/ψ muons

**USED SAMPLES**
- **Background:** inclusive data sample 2018 (Magnet Down), high mass sideband (M(B)>5.4 GeV)
- **Signal:** MC simulation 2018 (Magnet Down), \(B \rightarrow (\chi_{c1} \rightarrow J/\psi \mu^+ \mu^-)K\) with normal \(\chi_{c1}\) mass distribution

Cut efficiencies
Rejecting misID soft dimuon background

The situation:

Learn with simulation: $D^0 \rightarrow K^- \pi^+ (\rightarrow \mu^+ \nu_\mu)$, with reconstructed $K^-$ and $\mu^+$, and designed to access to generator level truth!

2 "types" of features
- muon identification parameters
- track quality (kink finder) to discard $\pi \rightarrow \mu \nu_\mu$ decays in flight
**MisID soft muon BDT**

*Let's mix all this!*

**TRAINING VARIABLES** (single soft muon only)

**PID variables**
- Global NN-computed muon/kaon probability
- Global pion likelihood:
  
  \[ \text{PID}(X) = \log(\text{Likelihood}(X)) - \log(\text{Likelihood}(\pi)) \]
- RICH pion likelihood
- Kink-tracker
- Kinematics to add dependency of the PID in the soft muon momentum

**SAMPLES**

- **Background**: \( D^0 \rightarrow K\pi(\rightarrow \mu\nu) \) sample weighted in momentum and in transverse momentum sample wrt simulation
- **Signal**: MC simulation 2018 MD, \( B \rightarrow \chi_{c1} \rightarrow J/\psi \mu^+\mu^- )K \) with normal \( \chi_{c1} \) mass distribution

---

Sonia Bouchiba

Search for multilepton final state \( B^+ \rightarrow J/\psi K^+\gamma^*(\rightarrow \mu^+\mu^-) \) @LHCb
Is our method more powerful than a standard PID cut?

↓ Signal efficiency ↓

↓ MisID background efficiency ↓

![Graphs](image-url)
Cuts applied on data sample


Signal: $27 \pm 7$ evts
Misidentified background: $17 \pm 17$ evts
Combinatorial background: $386 \pm 25$ evts
Take-home messages

• The decay channel $B^+ \rightarrow K^+ J/\psi \gamma^* (\rightarrow \mu^+\mu^-)$ is currently searched for at LHCb

• Particular attention is paid to soft muons background rejection.

• A dedicated momentum-dependent BDT is trained with specific variables.

  ☞ It shows a better rejection of decays in flight in background compared to the standard muon identification algorithm 😊

• This study will be extended to FSR muons used in other channels, in order to probe Leptonic Flavour Universality in $B^+ \rightarrow \ell\nu\ell$ decays without suffering of helicity suppression

• Results will be used to improve data-taking efficiency for soft muons
Thank you!
Backup
How to simulate $\gamma^* \rightarrow \mu^+ \mu^-$ signal?

- No decay model in EvtGen for such $\gamma^*$
- **Assumption**: dimuon coming from Initial State Radiation (ISR) (choice among other possible processes)
- Identify $\gamma^*$ as $\omega$ and use model “vector to 2 leptons” (VLL)

- Fast simulation tool RapidSim (global LHCb response)
- Requested full-sim sample (~0.5M) for flat mass distribution, can be reshaped later
Data fitting

Simulation sample

\[ B \to (\chi_{c1} \to J/\psi \mu^+ \mu^-)K \]

Fitted with a Crystal Ball shape
Tail parameters fixed

MidID muons sample

\[ B^+ \to K^+ J/\psi (\pi^+ \to \mu^+ \nu_\mu)(\pi^- \to \mu^- \bar{\nu}_\mu) \]

Fitted with a Crystal Ball shape
Tail parameters fixed
Preselection details

• Preselection:

<table>
<thead>
<tr>
<th>B</th>
<th>Kaon</th>
<th>Jpsi (→ mu1 mu2)</th>
<th>Soft muons (mu3, mu4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOS L0 Muon OR DiMuon decision</td>
<td>PIDK &gt; -2</td>
<td>50 MeV window around Jpsi mass</td>
<td>Soft muons Tracks’ GhostProb &lt; 0.3</td>
</tr>
<tr>
<td>B_DIRA_OWNPV &gt; 0.9999</td>
<td>Track’s GhostProb &lt; 0.3</td>
<td>mu1,2 ProbNNmu &gt; 0.05</td>
<td></td>
</tr>
</tbody>
</table>

• Overall: require soft muons to both pass "isMuon" - cut @P=3GeV
• (Correlated background cuts)

<table>
<thead>
<tr>
<th>p [GeV/c]</th>
<th>IsMuon</th>
<th>Required stations</th>
<th>IsMuonLoose</th>
<th>Required stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>p &lt; 3</td>
<td>Always false</td>
<td></td>
<td>Always false</td>
<td></td>
</tr>
<tr>
<td>p &lt; 6</td>
<td>M2 &amp; M3</td>
<td></td>
<td>At least two of M2–M4</td>
<td></td>
</tr>
<tr>
<td>6 &lt; p &lt; 10</td>
<td>M2 &amp; M3 &amp; (M4</td>
<td></td>
<td>M5)</td>
<td></td>
</tr>
<tr>
<td>p &gt; 10</td>
<td>M2 &amp; M3 &amp; M4 &amp; M5</td>
<td></td>
<td>At least three of M2–M5</td>
<td></td>
</tr>
</tbody>
</table>
isMuon ALGORITHM
- tracks are extrapolated to the muon stations
- Requires 2 or 3 hits in the muon stations depending on particle momentum (backup)

↓ Misidentification probabilities for kaons and pions → more likely to miss muons at low momenta

\[
\Delta LL(X-Y) = \log(\text{Likelihood}(X)) - \log(\text{Likelihood}(\pi))
\]
gives the likelihood for the probed particle X to be an Y

Taken from [LHCb Collab., 2019]
BDTs cut efficiencies

**Benchmark Combinatorial BDT**

For 3542 signal and 4224 background events, the maximum $S/S+B$ is 54.3064 when cutting at 0.0517.

**Benchmark soft muon BDT**

For 3542 signal and 4136 background events, the maximum $S/S+B$ is 53.9123 when cutting at -0.0153.
Comparison of stripping lines

inclusive line

B_LOKI_MASS_JpsiConstr [B_LOKI_MASS_JpsiConstr>0]

htemp

Entries 16941
Mean 5378
Std Dev 178.7

B_LOKI_MASS_JpsiConstr [B_LOKI_MASS_JpsiConstr>0]

htemp

Entries 3305
Mean 5290
Std Dev 109.4

min(mu3_PT,mu4_PT)

h1

Entries 3305
Mean 349.7
Std Dev 172.6

tight 3/4 line

inclusive
Category BDT2 with different conditions on mu3_P

Category in TMVA

• separate the training data (and accordingly the application data) into disjoint sub-populations exhibiting significantly different properties.

• The separation into phase space regions is done by applying requirements on the input and/or spectator variables (in this case mu3_P is spectator)

• In each of these disjoint regions (each event must belong to one and only one region), an independent training is performed using the most appropriate MVA method, training options and set of training variables in that zone.

• According to TMVA devs, the division into categories in presence of distinct sub-populations reduces the correlations between the training variables, improves the modelling, and hence increases the classification and regression performance.
Category BDT2 with different conditions on mu3_P

Background: weighted sample of $D^0 \rightarrow K\pi(\rightarrow \mu\nu)$

Signal: presel. inclusive MC 2018 MD, Evn 121450


**Categories:**
Cat 1: $\mu_3_P < 6000$
Cat 2: $\mu_3_P > 6000 \land \mu_3_P < 12000$
Cat 3: $\mu_3_P > 12000$

**Training Variables** (first attempt, will need the $var(p, p_T)$ performance to target more systematically what is best)
Cat 1: $\mu_3\_TRACK\_CHI2:log10(\text{abs}(\mu_3\_ProbNNmu)):log10(\text{abs}(\mu_3\_ProbNNk)):\mu_3\_PIDmu:\mu_3\_PIDK$
Cat 2 = vars of Cat 1
Cat 3 = vars of Cat 1 + $\mu_3\_RichDLLmu:\mu_3\_RichDLLk$

**Results:**

![TMVA overtraining check for classifier: Category](image)

![Background rejection versus Signal efficiency](image)
**Motivation**: "combinatorial" background,

- with 5 tracks random (not likely), IPChi2 already
gives an information about compatibility (already cut
on that)

**Aim**: check quality of [Kmu3mu4] or [Jpsimu3mu4] constrained vertex to discard the candidates
with double misID

=> take the sum of the IP chi2 or the **max**
Correlated background suppression

$K - \mu_3\mu_4$ correlations [234]
Correlated background suppression
\( \psi - K \) correlations [014]

Temporary
Correlated background suppression

$J/\psi - \mu_3 \mu_4$ correlations [014]
Soft dimuon misID background rejection
Kink calculator tool: principle

Case A: "Fake signal muons" are muons, but coming from decays-in-flight of hadrons in the tracking stations.

- 2 body decay, so change of direction of the muon track. This change of slope is the "kink".

- Interesting parameters

  track reconstructed
  Daughter muon associated with the track

  TRACK_kinkChi2: for each point on the track, measure the track chi2 difference between forward and backward tracks. Return the maximum difference along one track.

  TRACK_kinkChi2z: returns the z-coord of the maximum track chi2 difference.

  MATCHChi2: same as TRACK_kinkChi2 but only for kinks in the magnet
Soft dimuon misID background rejection
Gen-level matching

Motivation:

• MC truth gives the TRUEID but gives no info about the nature of the decay in flight. In particular, where did it happen?

• Crosscheck Kink algorithm. Can we trust it? Can we use it in a MVA?
Soft dimuon misID background rejection

**Kink calculator tool: does it tell the (MC)truth?**

**Where are the kinks happening?**

- **kinkZ** is the z-coord. of the greatest change in $\chi^2$ along a track

- Compare (MC TRUTH z) to [z-coord of max kinkChi2]
  - Signal muon: **prompt**, created in VELO + multiple scattering in T1-3
  - misID samples: muons created in-flight mostly

---

**Signal $\gamma^* \to \mu \mu$**

**$D^0 \to (\pi \to \nu \mu)K$**

---

Mu soft TRUEORIGINVERTEX_Z **(MC TRUTH !)**

Mu soft TRACK_kinkChi2z
Soft dimuon misID background rejection
Kink calculator tool: sounds good, doesn't work

Plot the scatter plot

If kink algo works: kinkchi2z = TRUEORIGINVERTEX
But: we observe random distribution 😞

Probably due to multiple scattering that "steals" the max chi2 difference

Expectation

Reality

=> Not adding kink tool in the MVA since not trusting them
More sophisticated: **kink** compter of the track, could spot on-flight creation of muons

\[
\begin{align*}
\mathcal{B}(K \rightarrow \mu \nu_\mu) &= 63.56 \% [\text{PDG 2019}], & \tau_K &= 1.2 \cdot 10^{-8} \text{s} \\
\mathcal{B}(\pi \rightarrow \mu \nu_\mu) &= 99.99 \% [\text{PDG 2019}], & \tau_\pi &= 2.6 \cdot 10^{-8} \text{s}
\end{align*}
\]

**Algorithm:**

For each track, collect the batch of mes. **nodes** (collection of track **states**)

For each node (with step \( s \)):

\[
\text{chi2kink} = \left( \text{chi2tot} - \text{chi2forward}[i] - \text{chi2backward}[i + 1] \right) / \text{ndof};
\]

Return the **maximum** one along the track.

S, ndof are hardcoded (4 and 5)

And \( z_{\text{kink}} = 0.5 \times (\text{measnodes}[\text{kink}] \rightarrow z() + \text{measnodes}[\text{kink} + 1] \rightarrow z()) \);
• Muons are created on-flight from hadrons
  • Though, the reconstructed SV (ORIVX) are pretty similar
  • … but the TRUE z-coord of muon creation is very different
  • → much longer flight of K and pi -> muons tends to be created until T1-3 stations.. (then harsh cut)