SMARTHEP Annual Meeting

Università' di Milano-Bicocca, 30/09/2024 - 04/10/2024

ESR8 Micol Olocco, Prof. Johannes Albrecht

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, called H2020-MSCA-ITN-2020, under Grant Agreement n. 966086

SMARTHEP Annual Meeting

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, called H2020-MSCA-ITN-2020, under Grant Agreement n. 966086 [https://www.smarthep.org/]

1. Brief introduction
2. Project presentation: Flavour Tagging in Run 3 at LHCb
3. Conclusions

Introduction

technische universität dortmund

Who: Micol Olocco (ESR8), Prof. Johannes Albrecht

Where: TU Dortmund (Germany) - CERN

What: Real Time Analysis for global event triggering in LHCb

Particle Physics

"Study of the (anti-)deuteron production in pp collisions at 5 TeV" with ALICE (CERN)

Anomaly Detection

Anomaly Detection in large-radius jets, ATLAS

Natural Language Processing

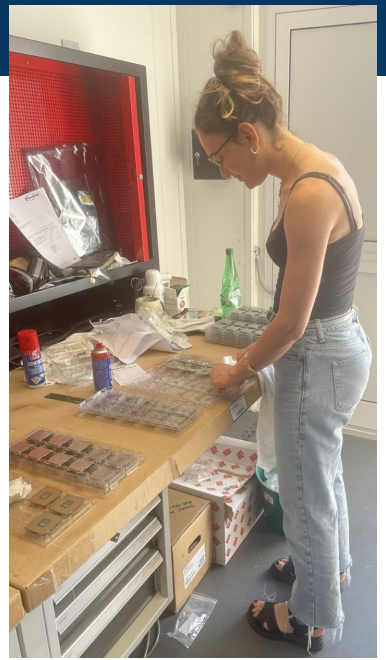
"Natural Language Processing techniques for error message analysis in WLCG data transfer" with Operational Intelligence (CERN)

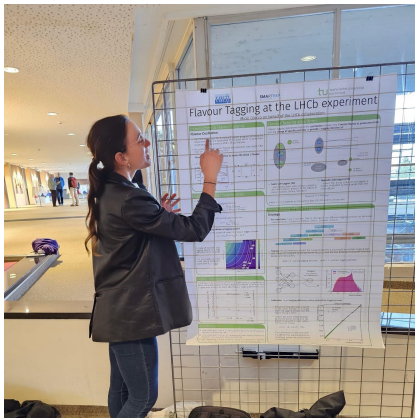
Consulting

Data Analyst in Accenture

Trainings & Talks

- **Data Manager shifts**
- **Trigger expert shifts**
- Presentation at the 112th LHCb week in Glasgow
- **Helping with the organization of the LHCb starterkit in November ‘24**
- “DPG Flavor meets Color School” in Bad Honnef
- SMARTHEP school on Edge Machine Learning
- **Organization of “SMARTHEP meets industries” event**
- Spontaneous work at the LHCb control room
  - GPUs installation
  - CPUs cleaning (which is not making ravioli)





Flavour Tagging algorithms

LHCb = LHC + b

Quark mixing allows flavour-changing currents in the weak interaction → Neutral mesons can oscillate between a particle and its antiparticle

How do we get the oscillation frequency?

→ **by measuring the time dependent oscillation asymmetry** $\mathcal{A}_{\text{mix}}^{\text{signal}}(t)$

$$\mathcal{A}_{\text{mix}}^{\text{signal}}(t) = \frac{N_{\text{unmixed}}(t) - N_{\text{mixed}}(t)}{N_{\text{unmixed}}(t) + N_{\text{mixed}}(t)} = \cos(\Delta m_d t)$$

$N\,(B^0 \to \textit{final state})$ $N(B^0 \to \overline{B}^0 \to \textit{final state})$ t = B^0 decay time

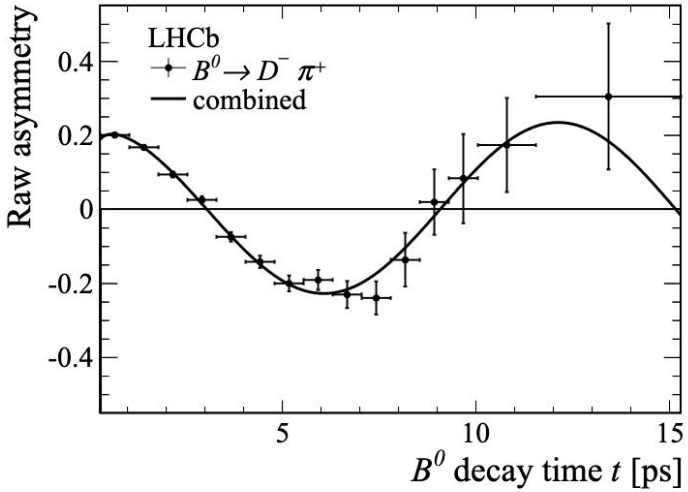


Fig: Raw mixing asymmetry A_{mix} (black points) for B0 → D−π+ [CERN-PH-EP-2012-315]

Flavour Tagging algorithms

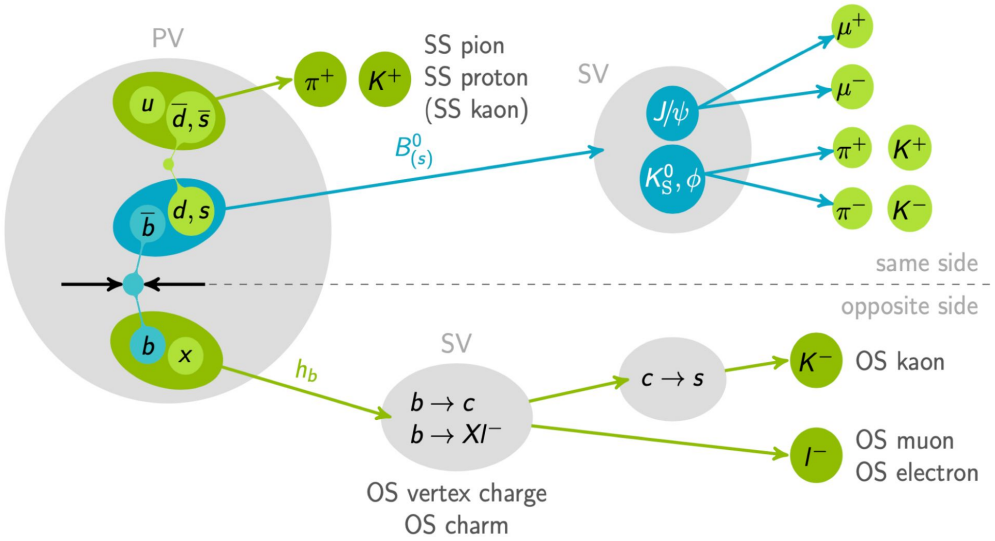
Exploit **correlation** between the **B meson flavour at production time** and the **charge of specific particles** to provide a **tagging decision (d)** .

$d_{predicted} = (\pm 1) \cdot Q$

- if $d_{predicted} = 1 \rightarrow \bar{b}$
- if $d_{predicted} = -1 \rightarrow b$

According to the ***tagging particle:***

- **Opposite side taggers (OS)**
- **Same side taggers (SS)**



Flavour Tagging algorithms

Flavour Tagging algorithms performance is evaluated by:

Effective Tagging Efficiency (Tagging Power)

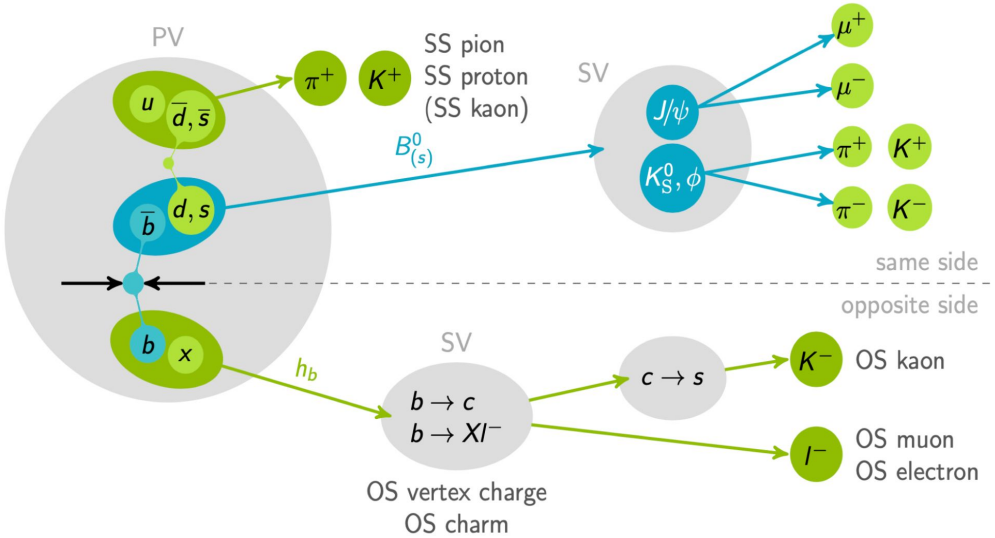
$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} \cdot (1 - 2\omega)^2$$

Tagging Efficiency: $\varepsilon_{\text{tag}} = \dfrac{N_{\text{tagged}}}{N_{\text{tagged}} + N_{\text{untagged}}}$

Mistag: $\omega = \dfrac{N_{\text{false}}}{N_{\text{tagged}}}$

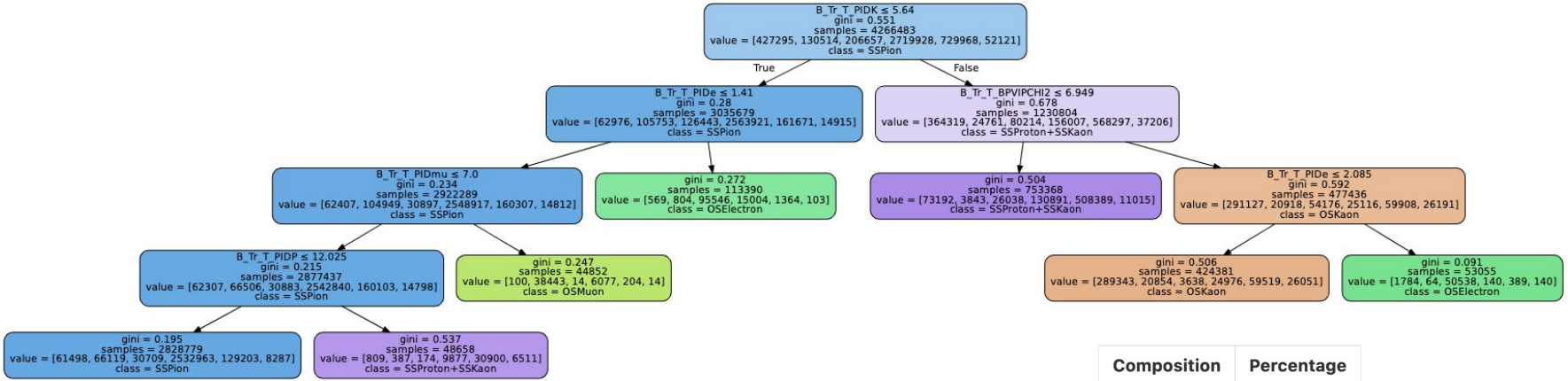
Where is ML?

1) Given a tagger type, pre-select the corresponding tagging particles
   - Correctly identify the tagging particle type
   - Pick up the particle with the correct charge-B flavour relation



1) Decision Tree Training

Merged sample of all decays used $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$, $B_s^0 \to D_s^- \pi^+$, $B^0 \to D^- \pi^+$



Composition	Percentage
SSPion	64%
OSKaon	10%
SSKaon	10%
SSProton	7%
OSElectron	5%
OSMuon	3%
OSProton	1%

Among 41 input features → PIDs + χ^2 of the best PV impact parameter identified as the most powerful cutting features

Where is ML?

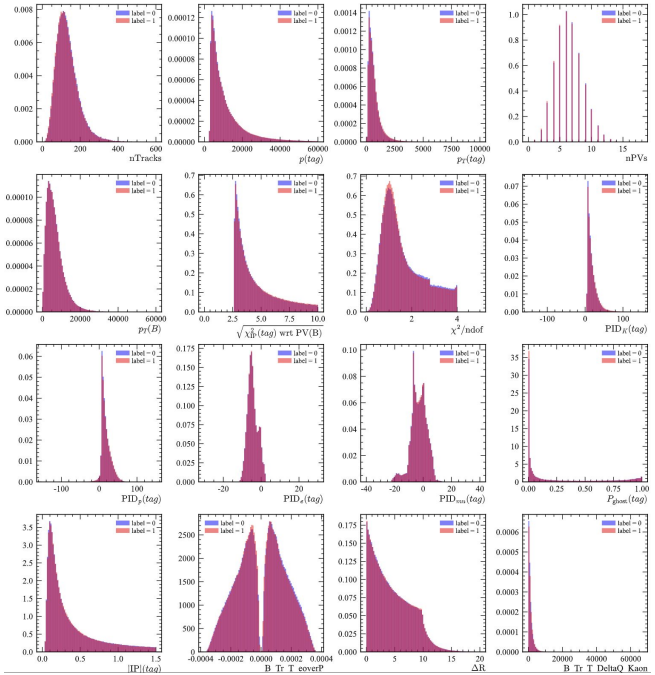
1) Given a tagger type, pre-select the corresponding tagging particles
   - Correctly identify the tagging particle type
   - Pick up the particle with the correct charge-B flavour relation
2) Train a Neural Network on wrong/correct tagging decision to predict the mistag η

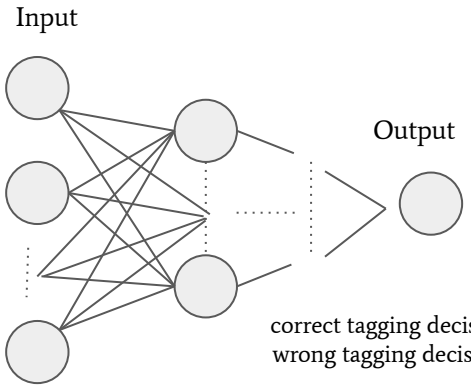
$$d_{predicted} = (\pm 1) \cdot Q$$

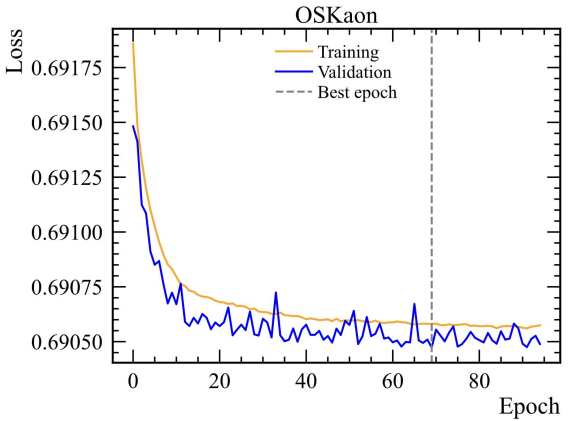
- η = probability of a wrong tagging decision $\Leftrightarrow$ p_{wrong} probability of belonging to the class 'wrong'

2) NN training

An example: OSKaon







AUC-ROC ~ 0.52!

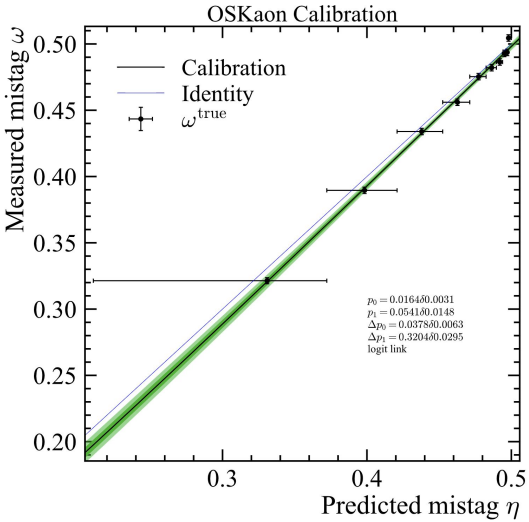
FT is a difficult task!

NN training

An example: OSKaon FT is a difficult task! However...

Tagger calibration:

- fit the predicted mistag η and the observed mistag ω
- to disentangle η from the possible dependence on the decay and sample used



Calibrated tagging power = 1.95+-0.04

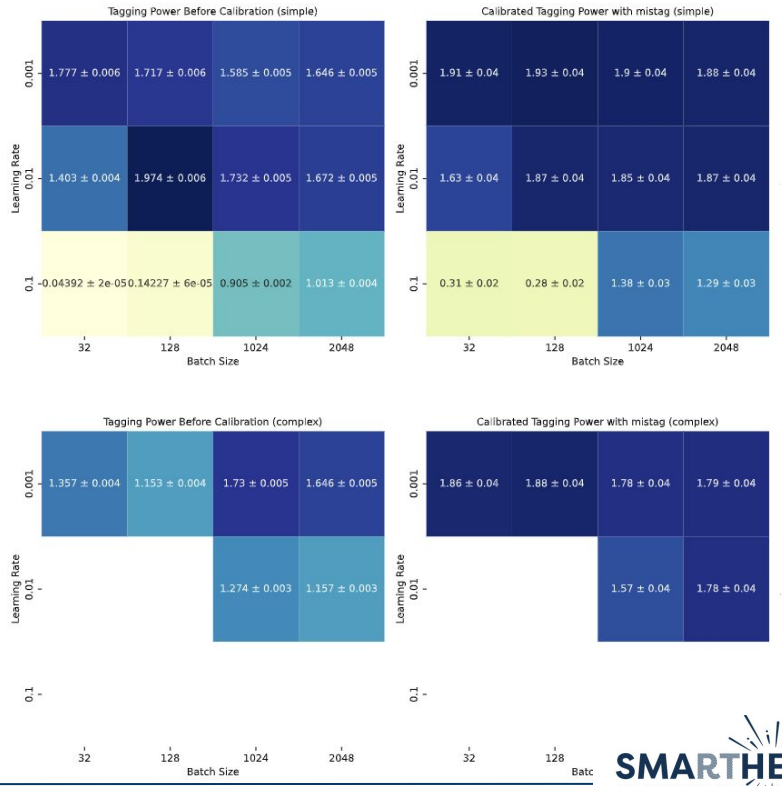
VS Run2 OSKaon ~ 1.4

Optimization: grid search

Explore different hyperparameter combinations and NN architectures to maximise the calibrated tagging power:

- NN architecture:
  - simple: one hidden layer with 3 nodes
  - complex: two hidden layers with 32 and 64 nodes
- Learning rate: 0.001, 0.01, 0.1
- Training batch size: 2024, 1024, 128, 32

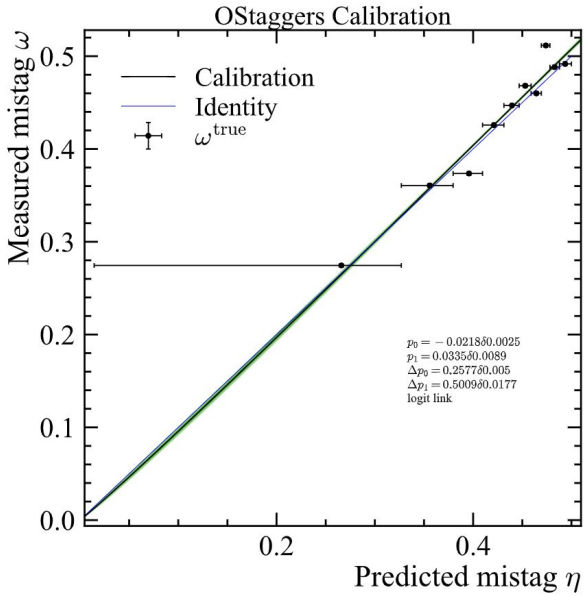
OSKaon, random seed=45



Performance check

Check performances of the tagger combination on Bs2JpsiPhi and Bd2DPhi:

- OSKaon, OSMuon, OSElectron, SSKaon on Bs2JpsiPhi



Tagging power: 3.99+/-0.04

→ close to the performance of the same combination of Run 2 taggers applied on Run 3 MC data.

Flavor Tagging: next steps

Definitely not the end of the story...

- Retrain pre-selection cuts including significant features not available before (ex. PROBNNe, PROBNNp etc)
- Retrain all the taggers with the new pre-selections
- Optimization
- ....
- Final step: port the taggers into the Run3 LHCb software to perform Physics Analysis
    - participate to a physics analysis (example: measurement of the oscillation frequency)

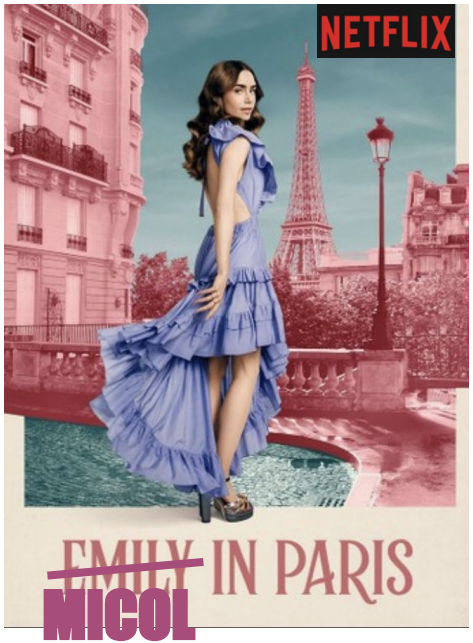
Secondment

Where: IBM in Paris

Who: with Pierre Feillet (IBM) and ESR Laura Boggia

What: Anomaly Detection techniques applied to fraud management

When: to be agreed in this week



Looking forward to it!

Conclusions

Plans for the upcoming year:

- Train and port taggers into LHCb software to make CP time dependent measurement possibles with Run 3 taggers
- Contribute to a physics analysis
- Keep on with the commissioning work for LHCb
- Secondment in Paris

Thank you for your attention

Performance check

Check performances of the tagger combination on Bs2JpsiPhi and Bd2DPhi:

- SSProton, SSPion on Bd2DPi

- ❑ Many things are going wrong... (still under study)
- ❑ Calibrated tagging power ~ 0 for both SSProton and SSPion