technische universität dortmund

Bundesministerium für Bildung und Forschung

FSP LHCb
Erforschung von Universum und Materie

Flavour Tagging at the LHCb experiment

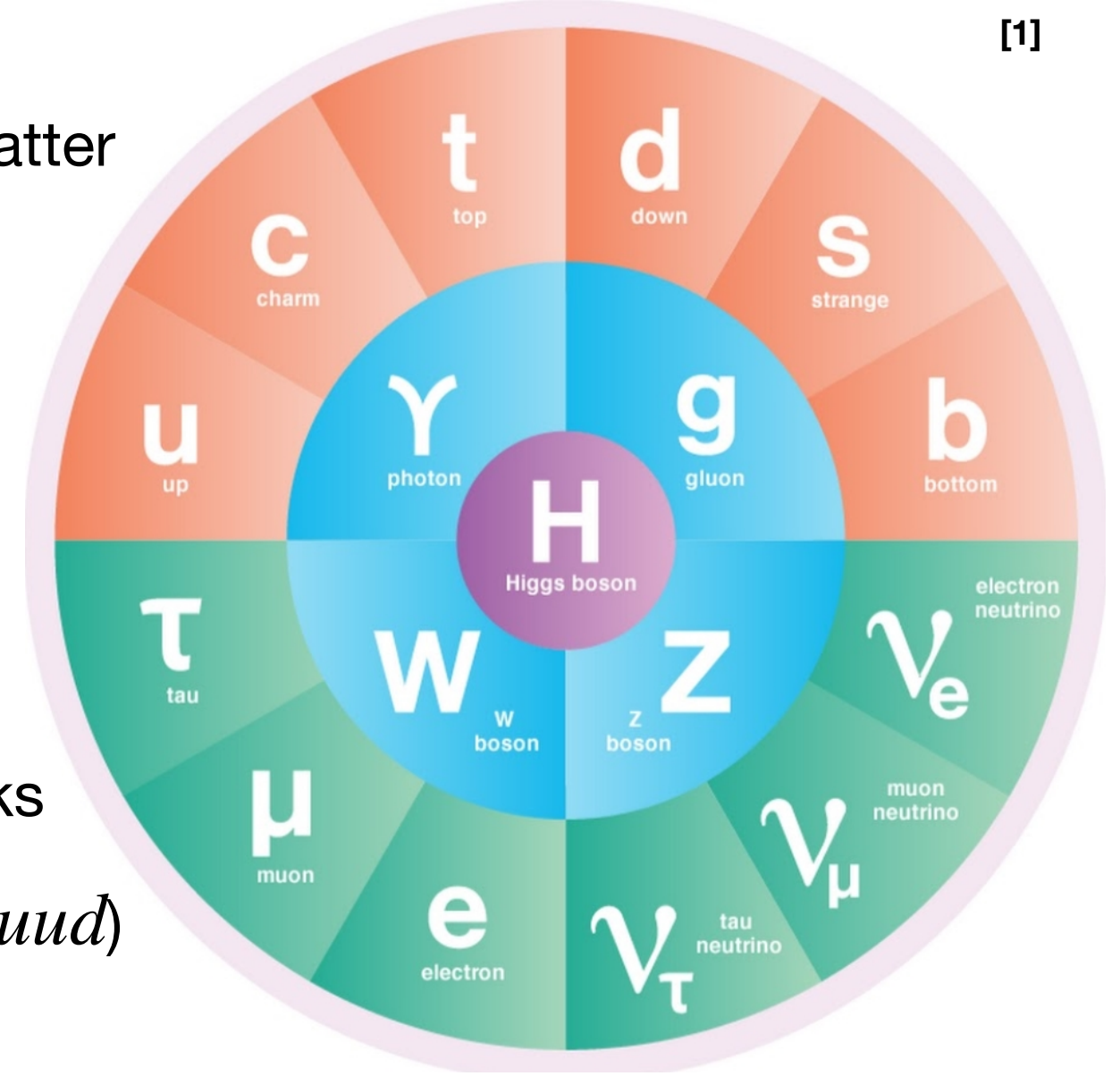
Jonah Blank (Technische Universität Dortmund)

Johannes Albrecht, Quentin Führing and Vukan Jevtic

Quarks & Hadrons

Quarks:

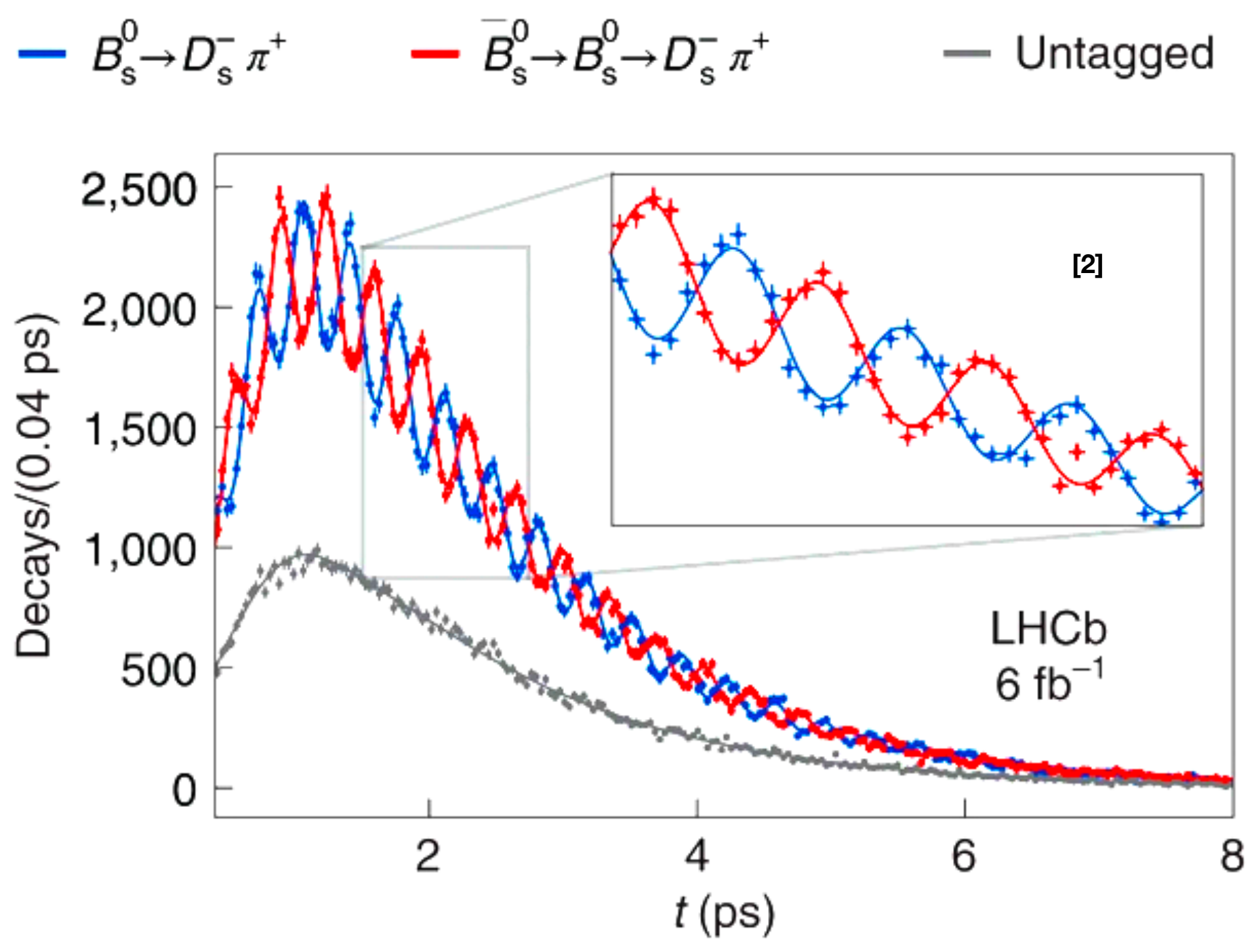
- Fundamental building blocks of visible matter in the universe
- 3 generations of particles
- Carry colour charge (RGB)
  → not observable as free states



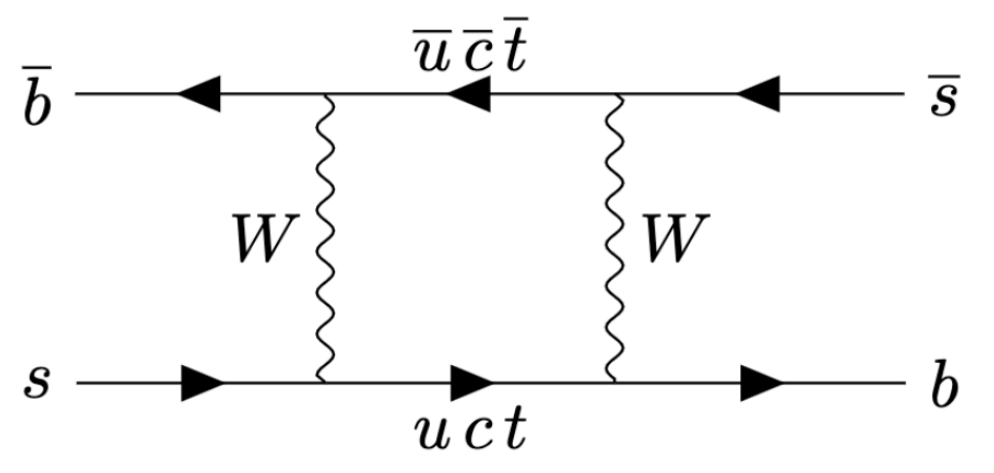
Hadrons:

- Colour-neutral bound states of (anti)quarks
- Baryons made of qqq ($\bar{q}\bar{q}\bar{q}$), e.g. proton (uud)
- Mesons made of $q\bar{q}$, e.g. $B_s^0/\bar{B}_s^0$ ($\bar{b}s/b\bar{s}$)

Neutral B meson mixing & CPV



- Matter & antimatter connected via **C**harge & **P**arity (CP) transformation
- Baryons & charged mesons are fixed to (anti)particle state
- Neutral mesons constantly oscillate between states ("flavour")



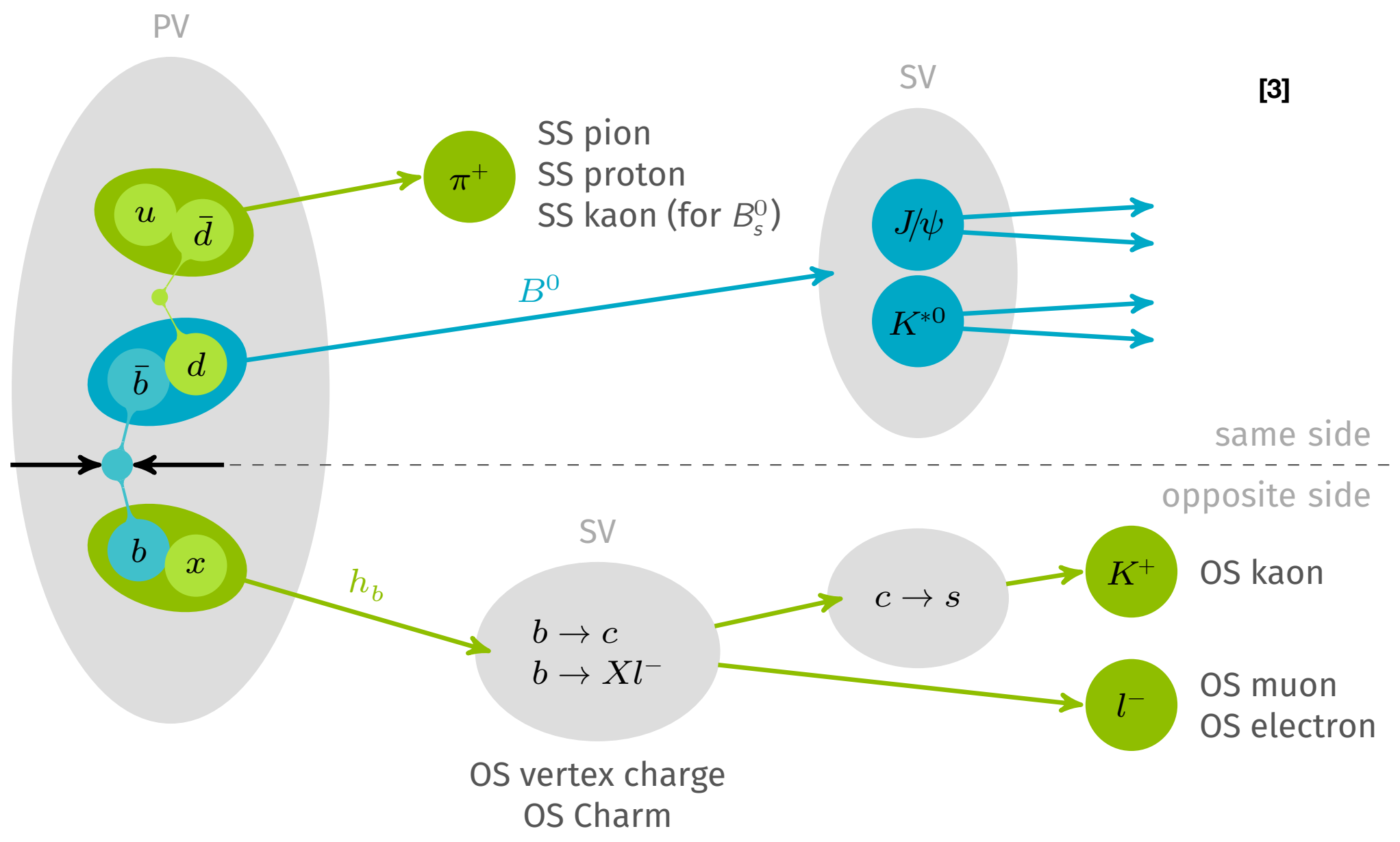
- CP symmetry violated (CPV) → difference between matter and antimatter
- Neutral (B) mesons: time-dependent CPV measurement possible due to oscillation
- Correct knowledge of flavour at decay needed to determine asymmetry in mixing and decay

A method is needed to identify state of particle at production.

Current Flavour Taggers at LHCb

Same Side (SS) taggers:

- Use charged particles produced alongside the signal $B^0_{(s)}$ meson (**Frag**mentation) to directly identify the signal flavour
  - SS Kaon for B_s^0 identification
  - SS Pion/Proton for B^0 identification



Opposite Side (OS) taggers:

- Use properties of non-signal b hadron decay coming from the initial $b\bar{b}$ pair → independent of signal meson type (B^0/B_s^0)
  - OS Vertex Charge: measures average charge of OS tracks
  - OS Charm: uses $\Lambda_c^\pm/D^\pm$ from $b \to c$ transition
  - OS Kaon: uses charged kaon in the final state
  - OS Muon/Electron: use lepton from semi-leptonic B decay

Tagging outputs:

- Tag decision d: prediction of initial signal B flavour
- Mistag estimate η: per event prediction of how likely the wrong d was assigned

Flavour Tagging characteristics:

- Tagging efficiency: fraction of successfully tagged events

$$\varepsilon_{tag} = \frac{N_{right} + N_{wrong}}{N_{right} + N_{wrong} + N_{untagged}}$$

- Average mistag probability: fraction of events with wrong tag decision
- Mistag prediction does not necessarily match the true mistag: calibrate η on flavour specific decays of real mesons using a linear function

$$\omega(\eta) = p_0 + p_1(\eta - \langle\eta\rangle)$$

- Effective tagging power: reduced efficiency due to dilution $D = 1 - 2\omega$ of mistaged events

$$\varepsilon_{eff} = \varepsilon_{tag} D^2$$

All tagging algorithms use a selection based approach as well as multivariate analysis.
The selection process depends on good particle identification and specific domain knowledge.
Good performance therefore requires a trade-off between tag quality and tagging efficiency.

The Inclusive Flavour Tagger (IFT)

The goal is to find a more general approach for LHCb to determine the production flavour of a decaying meson. Using advanced Machine Learning techniques the IFT aims to increase the performance and potentially replace the current ensemble of tagging algorithms.

The idea:

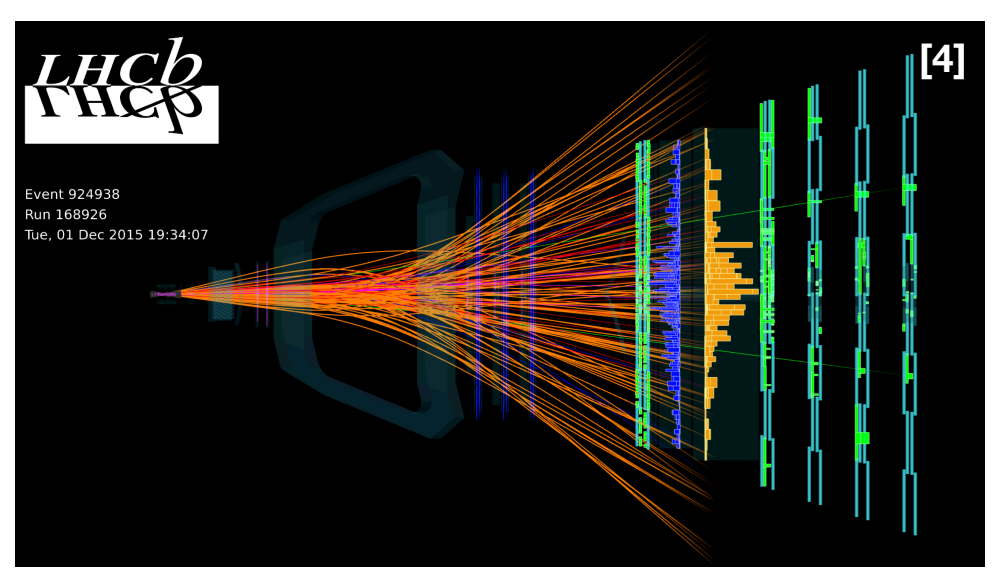
- Use all reconstructed tracks & vertices in an event not related to signal decay chain
  → not rely on specific segments and physics assumptions
- Apply only a loose selection to achieve high $\varepsilon_{tag} \lesssim 100\,\%$
  → let the neural network (NN) learn to recognise important information

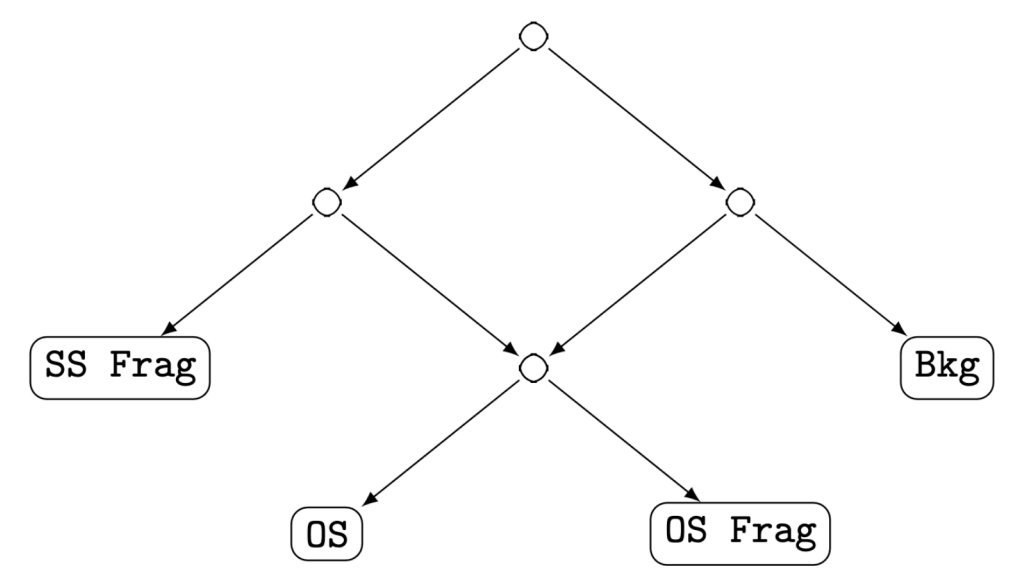
First implementations:

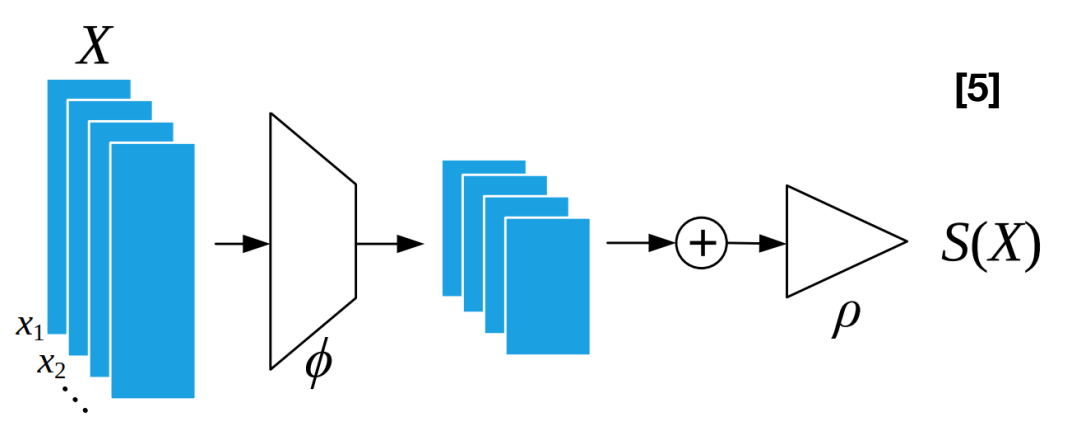
- Categorise tracks based on topology (SS Frag, OS, OS Frag)
  → multiclass Boosted Decision Tree (BDT) output as extra input for NN
- Number of tracks/vertices varies from event to event
  → DeepSet for fast training & to adapt for different input sizes

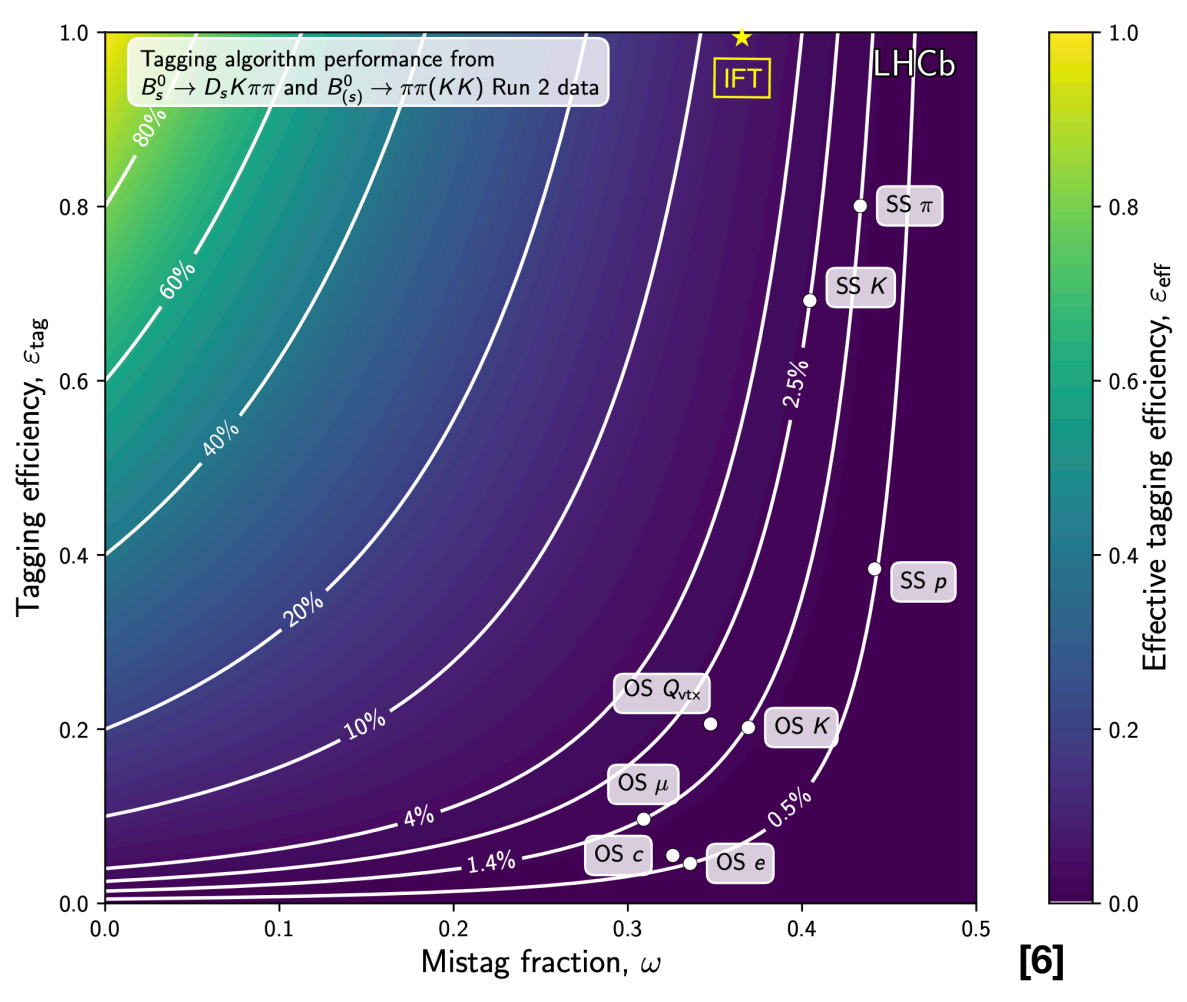
Training data:

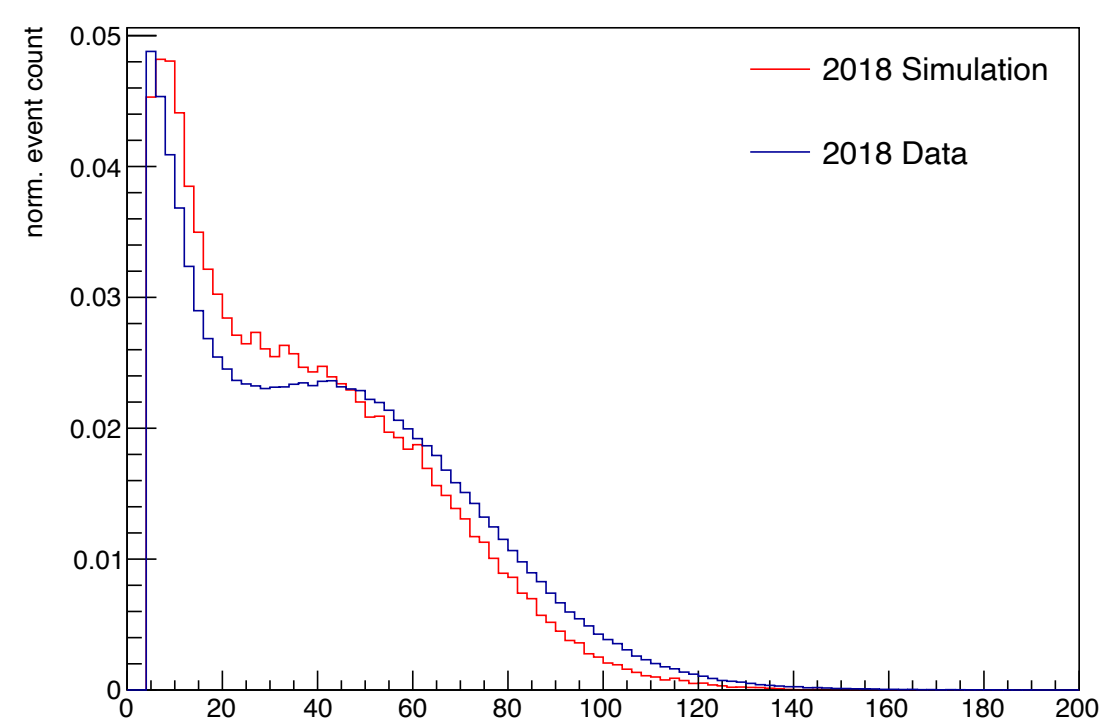
- Real data:
  - Background contamination
  - Limited statistics for specific decay channels
  - Difficult labelling in neutral meson decays
- Simulation:
  - Not a perfect description of reality
  - Pure signal samples
  - Truth information available
- Idea: train two separate taggers & combine their predicted tag decisions
  - SS tracks dependent on signal decay type: train SS-IFT on individual simulated samples for different decays
  - Train OS-IFT using real "self-tagging" $B^\pm \to J/\psi K^\pm$ decays
    → can be cross-applied on other channels











The IFT as the future of flavour tagging:

- Benefits:
  - Increased tagging power & high tagging efficiency
  - Simple framework: only one tagger needs to be trained the same way for different channels
- Challenges:
  - Complete event information used → sensitive to differences between simulation and real data
  - Background contribution have to be accounted for during training with real self-tagging decays
  - Real data training SS tagger not feasible for neutral meson decays due to oscillation
  - NN structure is a "black box" → hard to interpret compared to classical taggers
- Plans:
  - Test using detector output directly as input features instead of/in addition to reconstructed high level variables
  - Optimisation of training procedure & network architecture
  - Validate performance on different channels in real data

[1] https://www.energy.gov/science/doe-explainsthe-standard-model-particle-physics
[2] LHCb collaboration. Precise determination of the B_s^0–$\bar{B}_s^0$ oscillation frequency. *Nat. Phys.* 18, 1–5 (2022)
[3] https://twiki.cern.ch/twiki/bin/view/LHCb/FlavourTaggingConferencePlots
[4] https://twiki.cern.ch/twiki/bin/view/LHCb/LHCbPlots2015
[5] Manzil Zaheer et al. *Deep Sets*. 2017 https://arxiv.org/abs/1703. 06114
[6] Comparison of Flavour Tagging performances displayed in the ω-$\varepsilon_{\mathrm{tag}}$-plane LHCB-FIGURE-2020-02

CS & Physics Meet-Up by Lamarr & B3D, 29.11. - 01.12.23