LHCb

technische universität
dortmund

Development of the Topological Trigger for LHCb Run 3

Johannes Albrecht   Blaise Delaney   Greg Ciezarek
Niklas Nolte   **Nicole Schulte**   Mike Williams
On behalf of the LHCb Collaboration

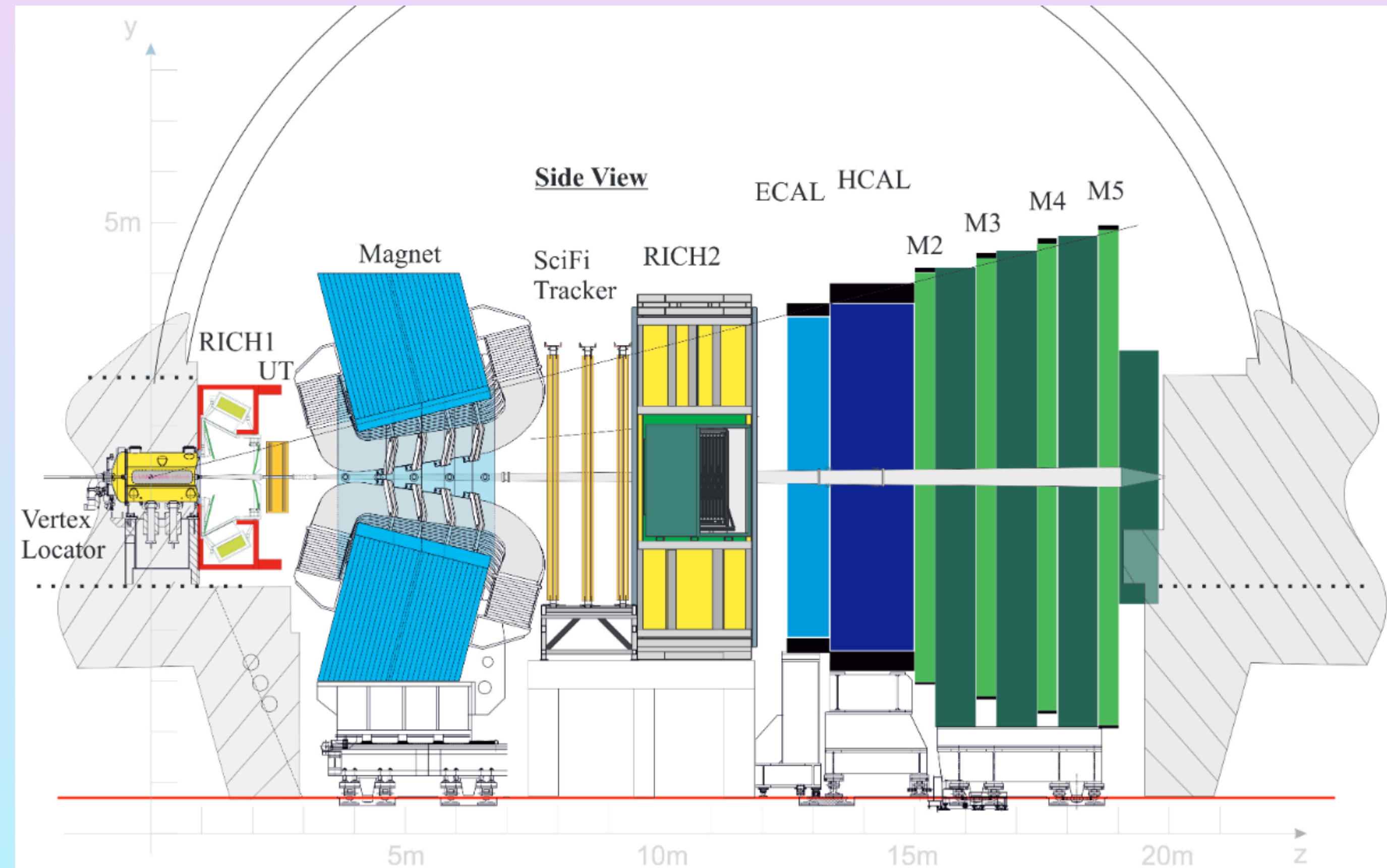ACAT 2022 - Track 1 Computing Technology for Physics Research
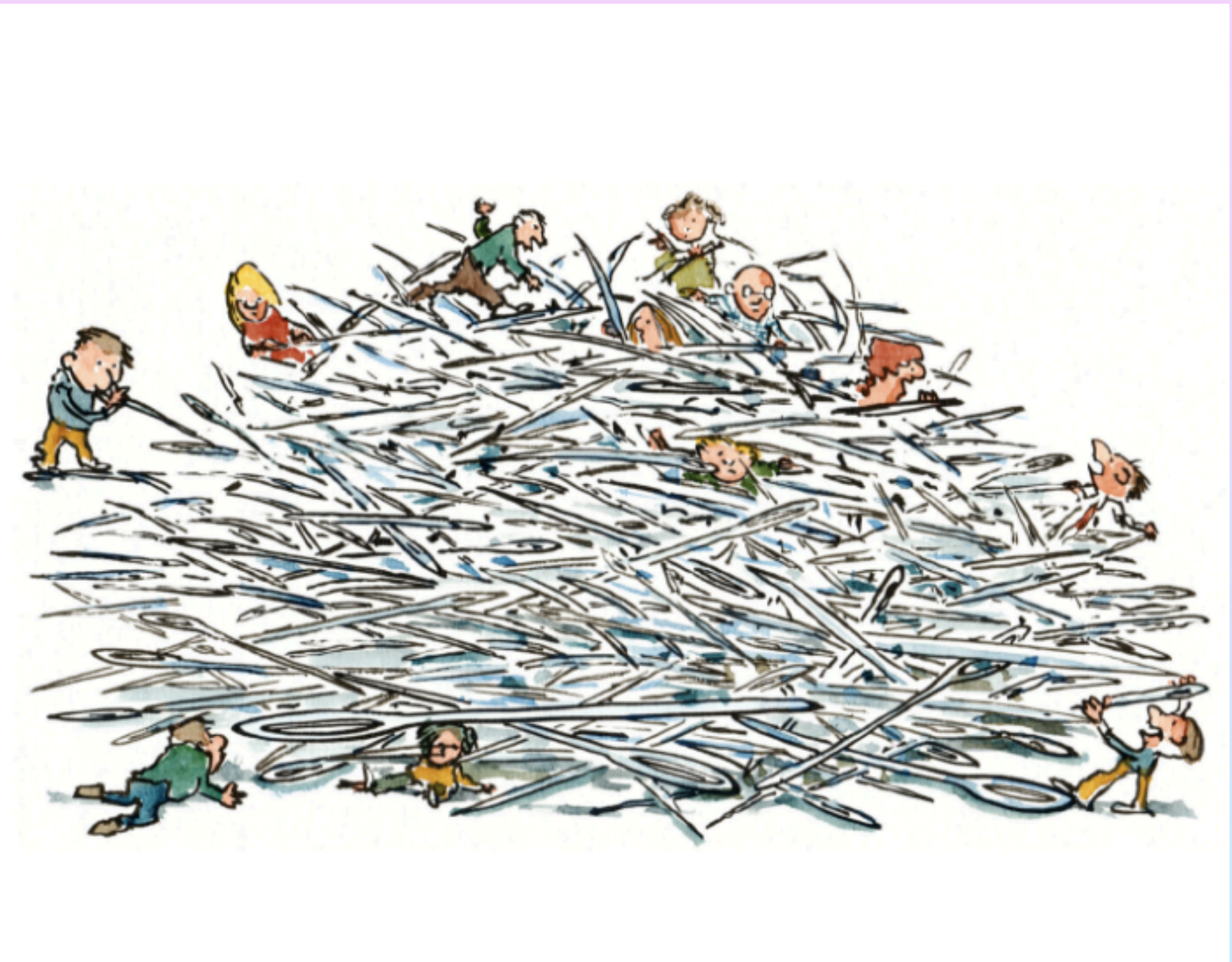
SFB1491

LHCb
RTA

nicole.schulte@cern.ch

The LHCb Detector in Run 3

- Forward arm spectrometer aiming to measure properties of beauty and charm properties

- Software-only trigger system

- Run 3 Luminosity:
$$\mathscr{L} = 2 * 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$$
  - ➡ 5x higher than in Run 2

- Brand new Tracking system and electronics

Why do we need a Trigger System?

Data acquisition: 30 MHz -> ~495 Ebit or 60 EB for one year*
➡Real-Time-Analysis (Fast selection of events before storage)

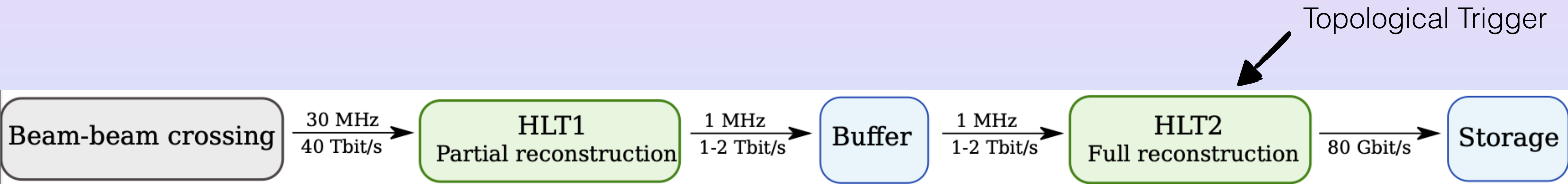*Assuming 143 days of pp-collisions in a year
and 40 Tbit/s data acquisition

LHCb's Trigger System in Run 3

➡ Remove Hardware Trigger L0
➡ Conversion to a software-based Trigger System

Topological Trigger

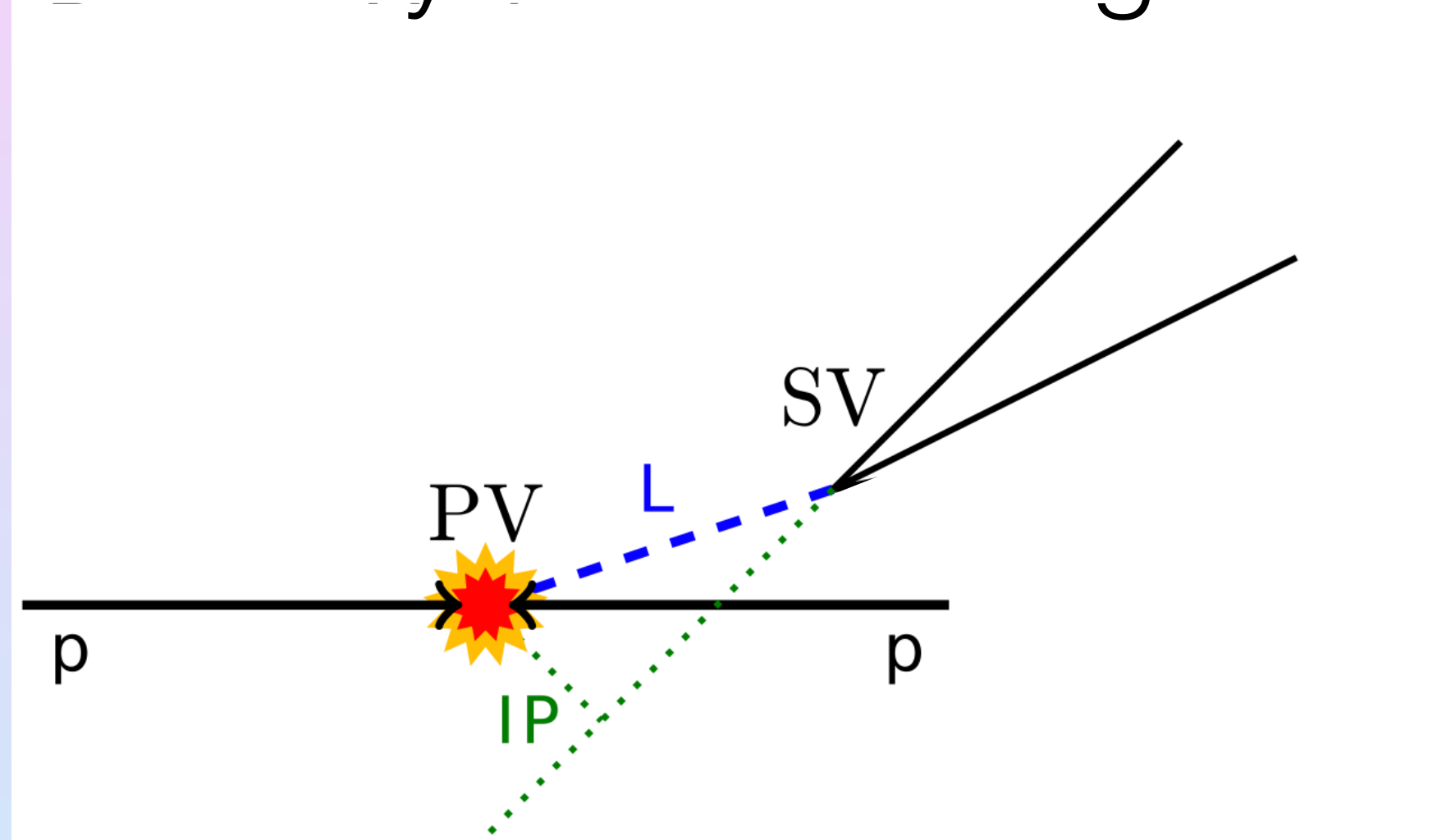| Beam-beam crossing | → 30 MHz / 40 Tbit/s → | HLT1 Partial reconstruction | → 1 MHz / 1-2 Tbit/s → | Buffer | → 1 MHz / 1-2 Tbit/s → | HLT2 Full reconstruction | → 80 Gbit/s → | Storage |

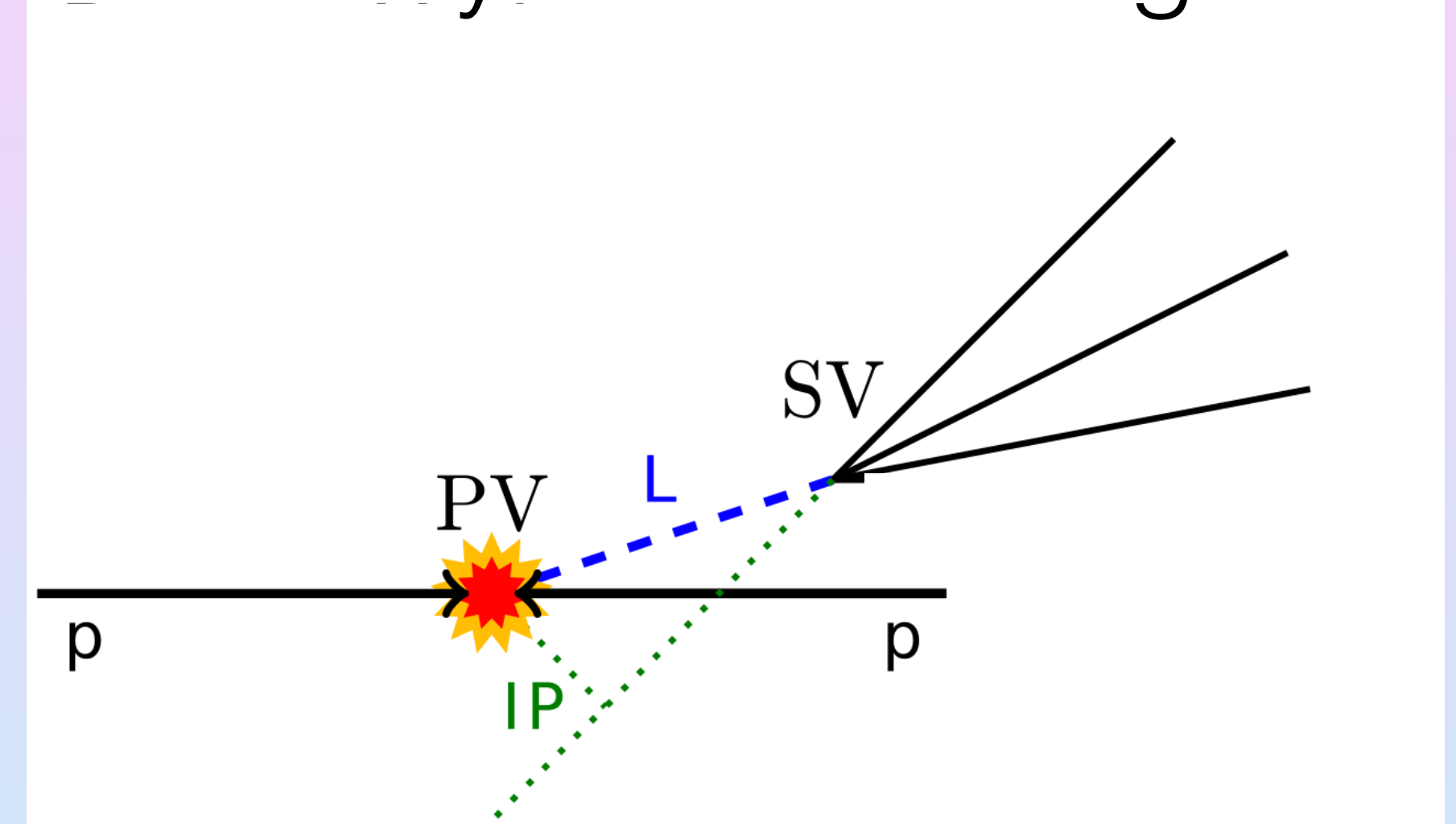## Reduce the incoming data by a factor of 500!

Assuming again 143 days of pp-collisions
➡ ~1Ebit or 125PB for storage

HLT2 Reconstruction and Selection

TwoBody Selection Algorithm

ThreeBody Selection Algorithm

1. HLT2 particle reconstruction and identification
2. Filter particles using kinematic and spatial properties
3. Combine Particles to decay candidates
4. Filter decay candidates
5. ThreeBody: Repeat Steps 3 and 4
6. Trigger Decision

What is the Topological Trigger?

- Inclusive selection targeting **any** beauty decay

- Discriminate beauty events from charm contributions and soft QCD

- Develop inclusive signature by training on various exclusive beauty samples
  - Each sample provides equal amount of signal to avoid biases

- Based on the topology of beauty decays:
  - Relatively high mass ~5GeV
  - Displaced decay vertex due to lifetime $\sim\mathcal{O}(1\,\text{ps})$
  - Boosted in the forward direction ~travels $\mathcal{O}(1\,\text{cm})$ before decay
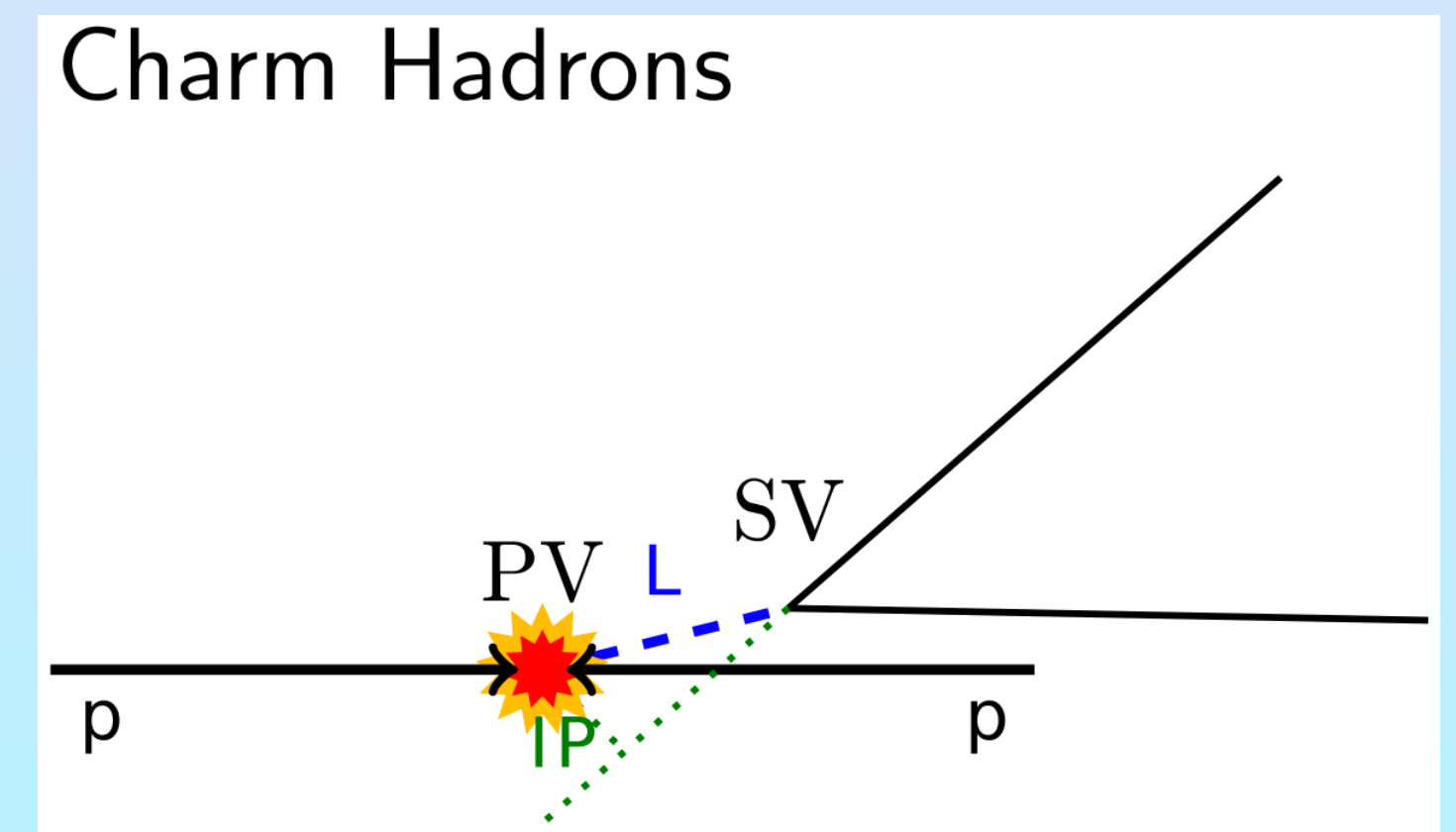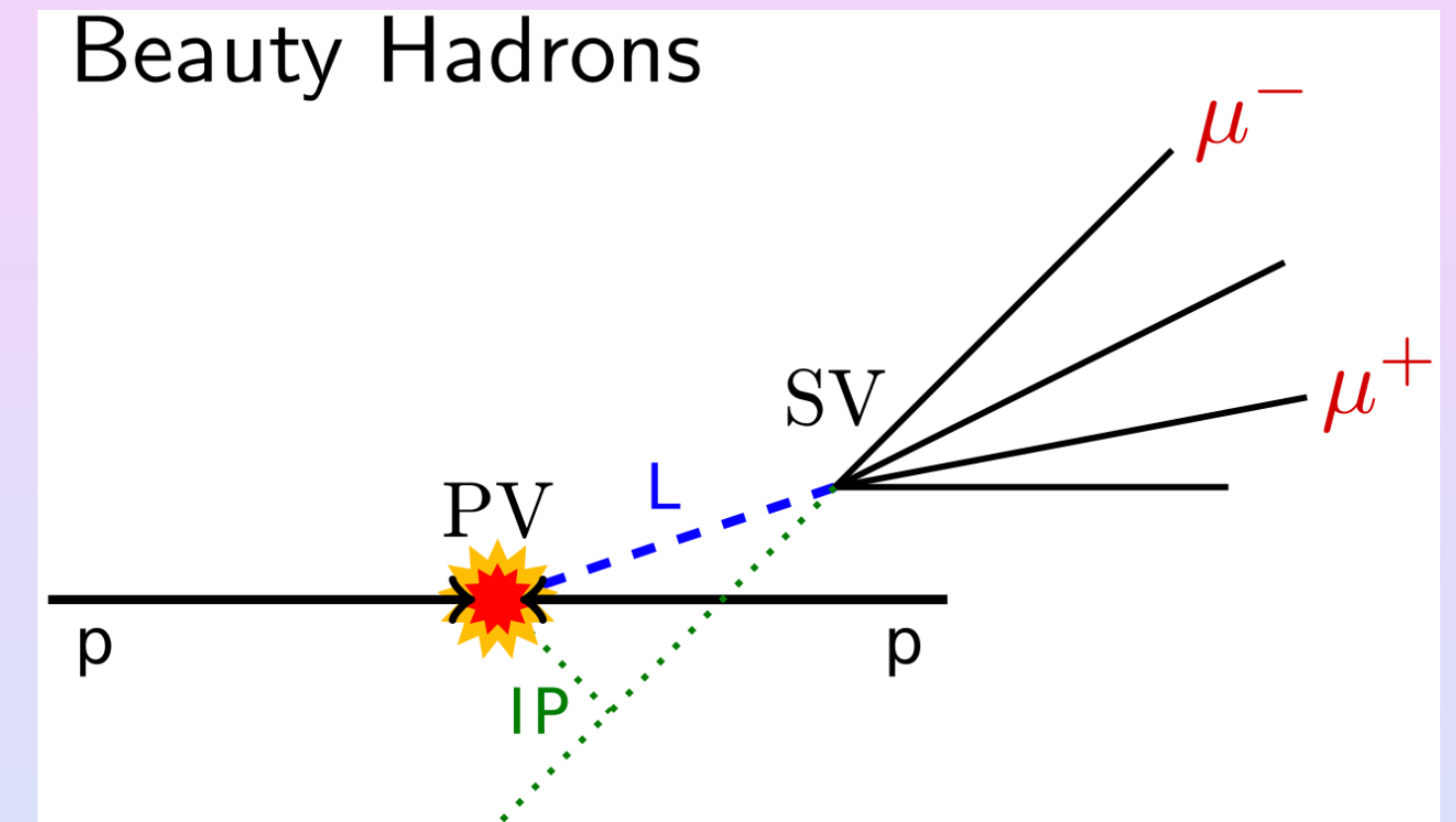
Why is the Topological Trigger important?

1. One of the main beauty Triggers for LHC**b**
2. Selection writes directly to storage
3. Pre-Filter for other selections

Electrons	Muons
HLT1 in Run I	
TrackAllL0	TrackAllL0
HLT1 in Run II	
TrackMVA	TrackMVA
HLT2 in Run I	
Topo(2,3,4)BodyBBDT	Topo(2,3,4)BodyBBDT
TopoE(2,3,4)BodyBBDT	TopoMu(2,3,4)BodyBBDT
HLT2 in Run II	
Topo(2,3,4)Body	Topo(2,3,4)Body
TopoE(2,3,4)Body	TopoMu(2,3,4)Body

Table from: Vitalii Lisovskyi, Study of rare b-baryon decays and test of lepton universality at LHCb

The Topological Trigger: Then vs. Now

Major upgrades in Run 3 for detector and software require innovative improvements for trigger algorithms
Machine Learning Algorithms have improved majorly over the past few years

Run 2: Boosted Decision Tree Algorithm

- Followed same selection strategy as in Run 1 with updated preselection and training of the BDT

Run 3: Neural Network (Monotonic Net)

- Reinvent the selection strategy by moving to a different machine learning algorithm that provides robustness against detector effects and sensitivity to long-lived beauty objects

Monotonic Net: ArXiv: 2112.00038, M. Williams, N. Nolte, O. Kitouni

LHCb-PUB-2014-031, S. Benson et. al. Trigger selections for the LHCb upgrade

Monotonic Net - Robustness
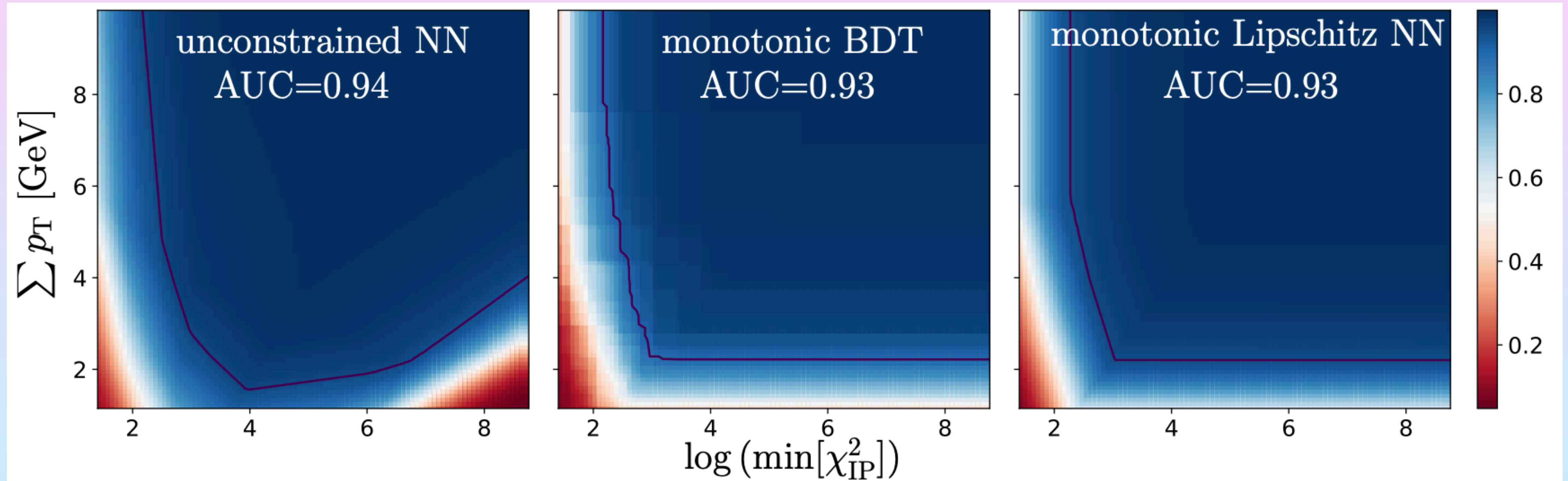
Neural Network architecture provides two main features:

- Robustness: Introduce Lipschitz constant λ:

    - Having inputs x and y and corresponding Prediction f(x) and f(y):
    |f(x) - f(y)| ≦ λ |x - y|

    - Means: Change of the output is constrained by the change in the input -> Detector effects are reduced -> Robustness ensured

ArXiv: 2112.00038, M. Williams, N. Nolte, O. Kitouni

M. Williams, N.Nolte, O.Kitouni: Provably Robust and Monotonic Networks for Heavy Flavor Selections at LHCb, ACAT 2021

Monotonic Net - Monotonicity

- Monotonicity: Given feature i and inputs x and y:

    - If $x_i < y_i$ ensure that the Prediction follows $f(x_i) < f(y_i)$

    - Chose features where monotonic behaviour could be beneficial

    - Chose direction of the monotonicity: positive or negative

For more details see Ouail Kitouni's talk from Monday

M. Williams, N.Nolte, O.Kitouni: Provably Robust and Monotonic Networks for Heavy Flavor Selections at LHCb, ACAT 2021

Monotonic Net - An Example

Blue regions: Classified as signal; Red regions: Classified as background; Black Line: Decision Boundary

➡Monotonic Net is more sensitive to **long-lived signal** candidates than unconstrained Neural Networks

➡Monotonic Net provides a smoother decision boundary than other monotonic architectures (**Robustness**)

ArXiv: 2112.00038, M. Williams, N. Nolte, O. Kitouni

M. Williams, N.Nolte, O.Kitouni: Provably Robust and Monotonic Networks for Heavy Flavor Selections at LHCb, ACAT 2021

(Monotonic) Features

- Kinematic properties: (transverse momentum p_T)

- Spatial properties: (Flight Distance FD, Impact Parameter IP, Distance of closest approach DOCA)

- Conceptually the **positive monotonic** features are:
    - Transverse Momentum
    - Flight Distance
    - χ^2 of the Impact Parameter (IP)

➡ Highly displaced tracks with a high momentum

➡ Long-Lived beauty objects
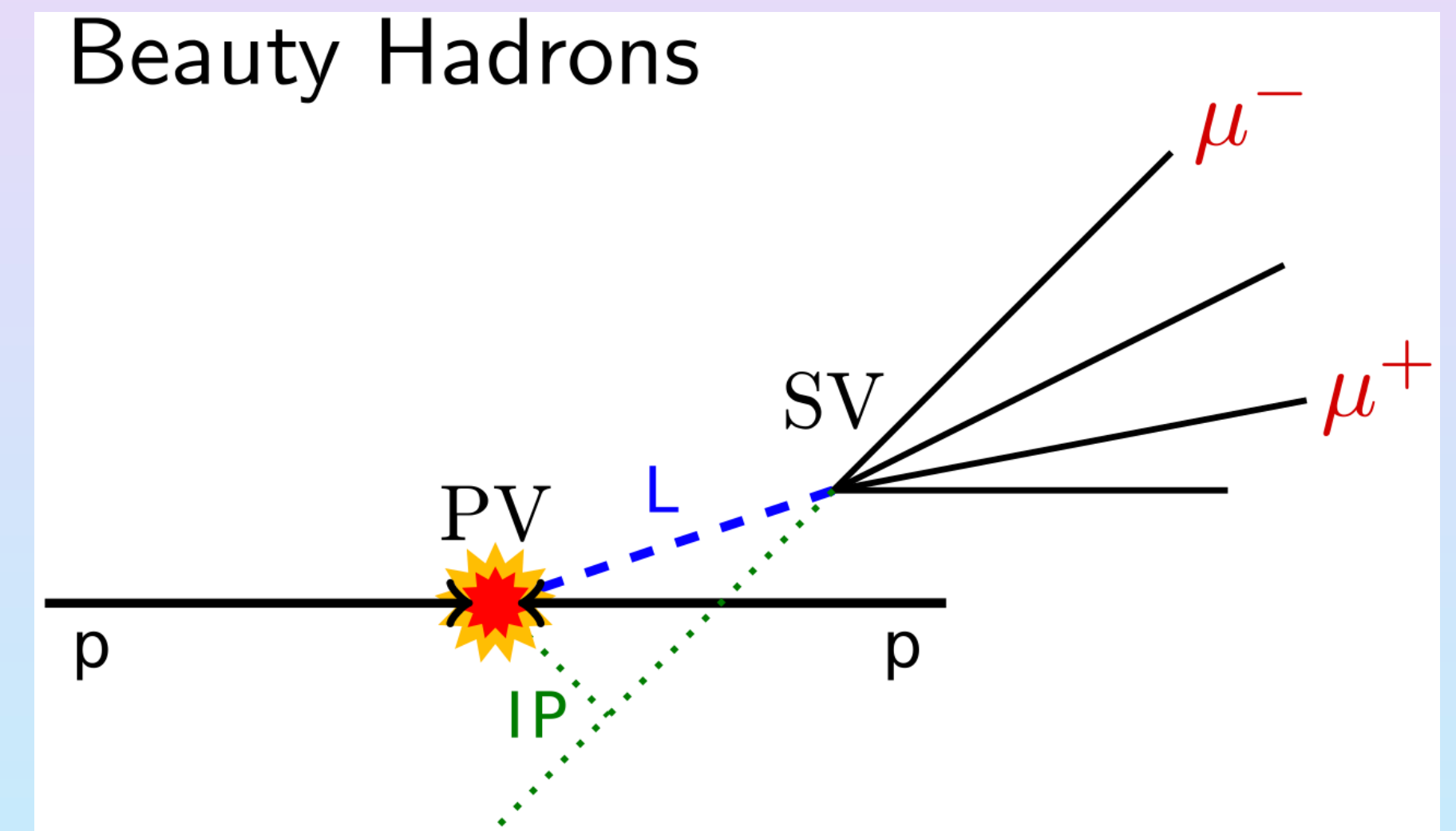
Beauty Hadrons

μ^-

μ^+

SV

L

PV

IP

p   p

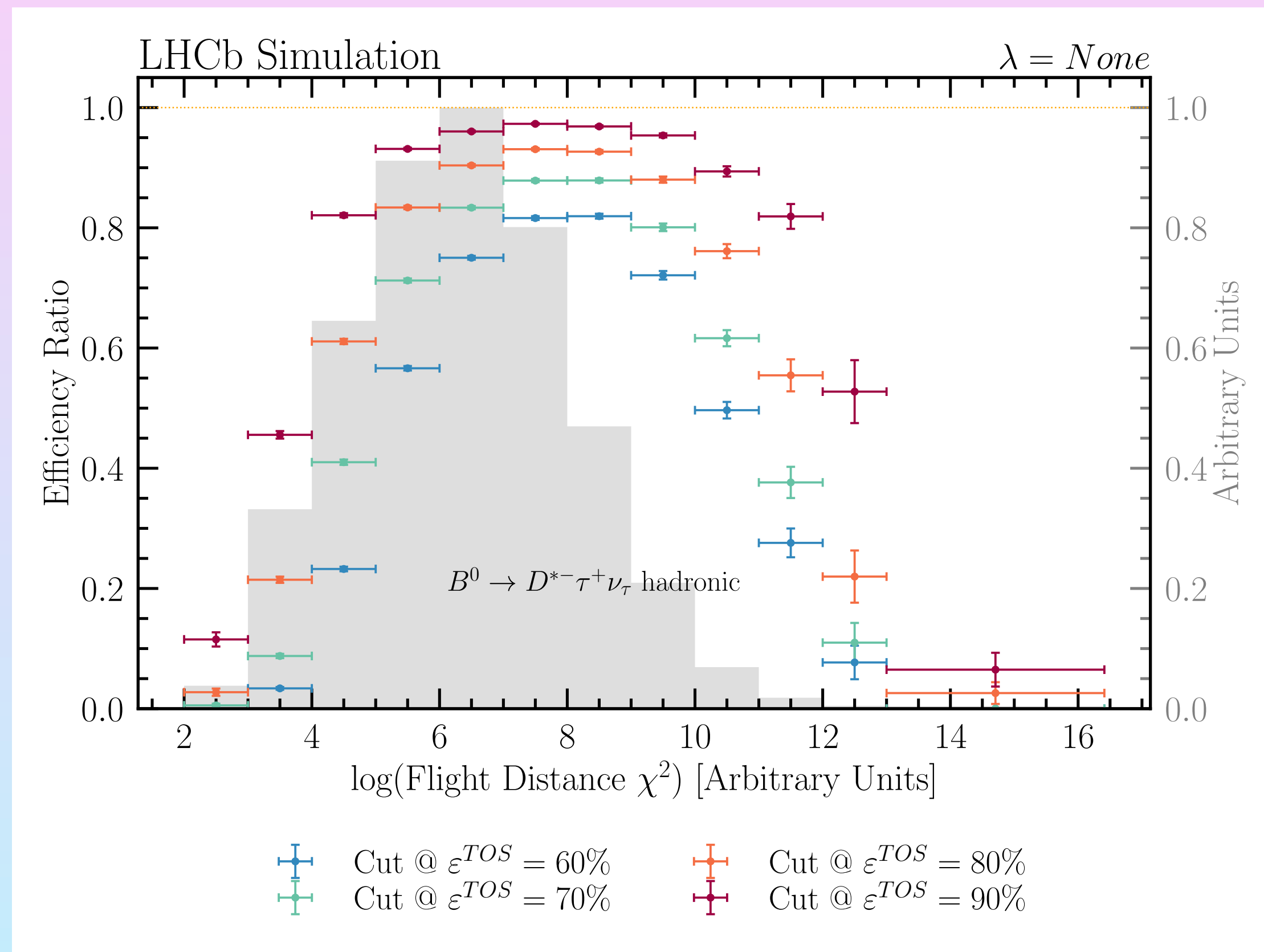# The Monotonic Net in the Topological Trigger
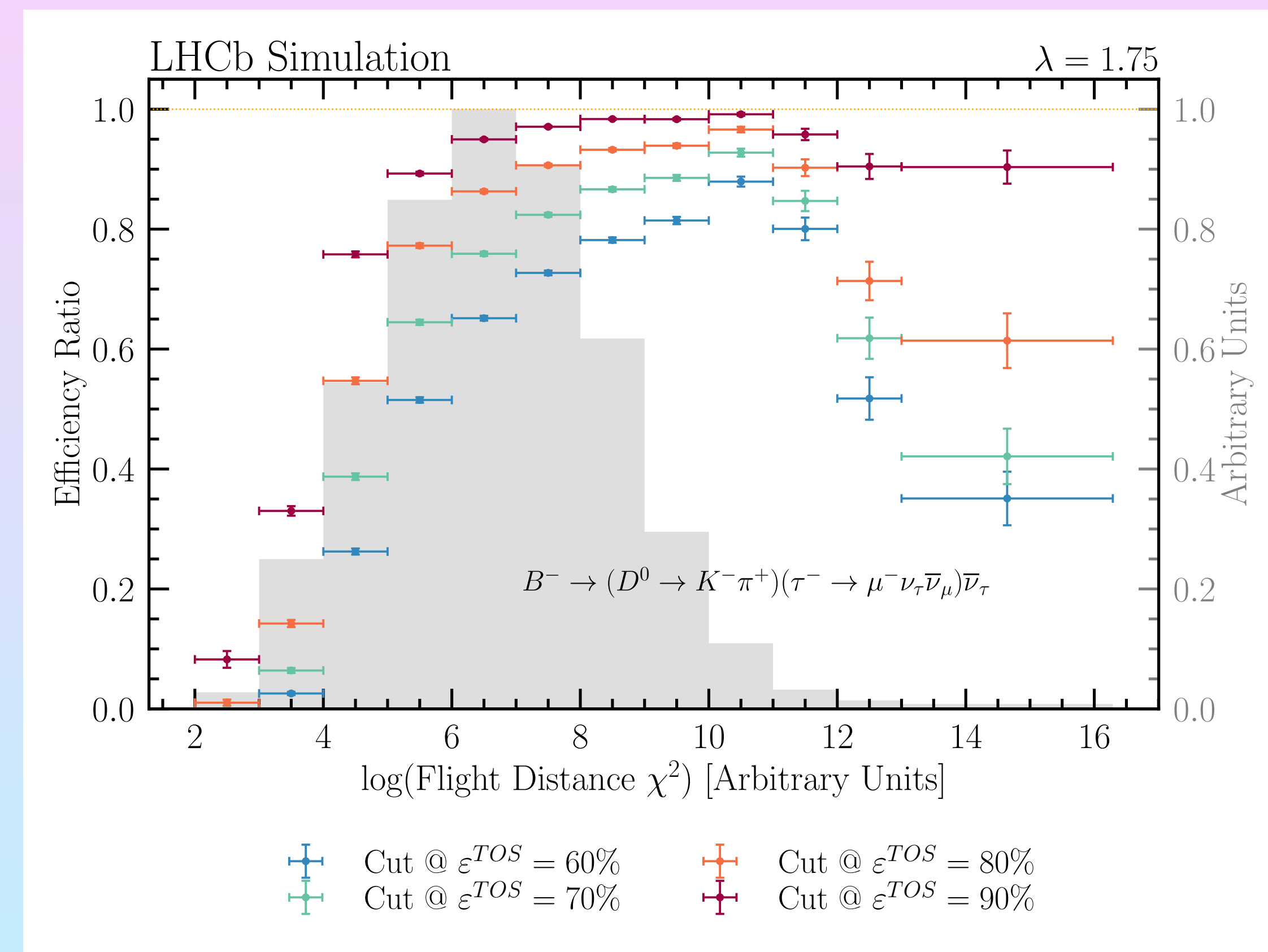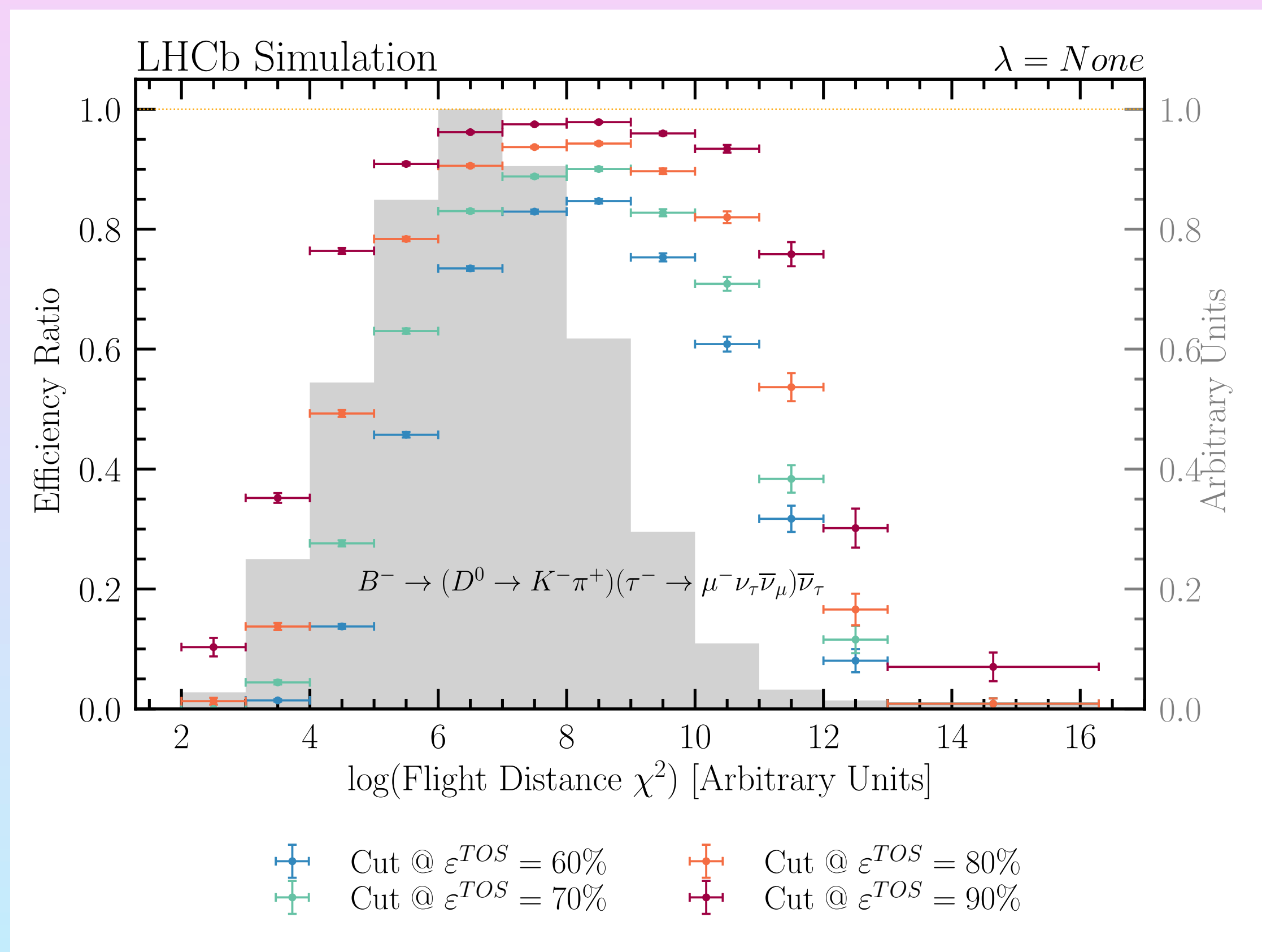
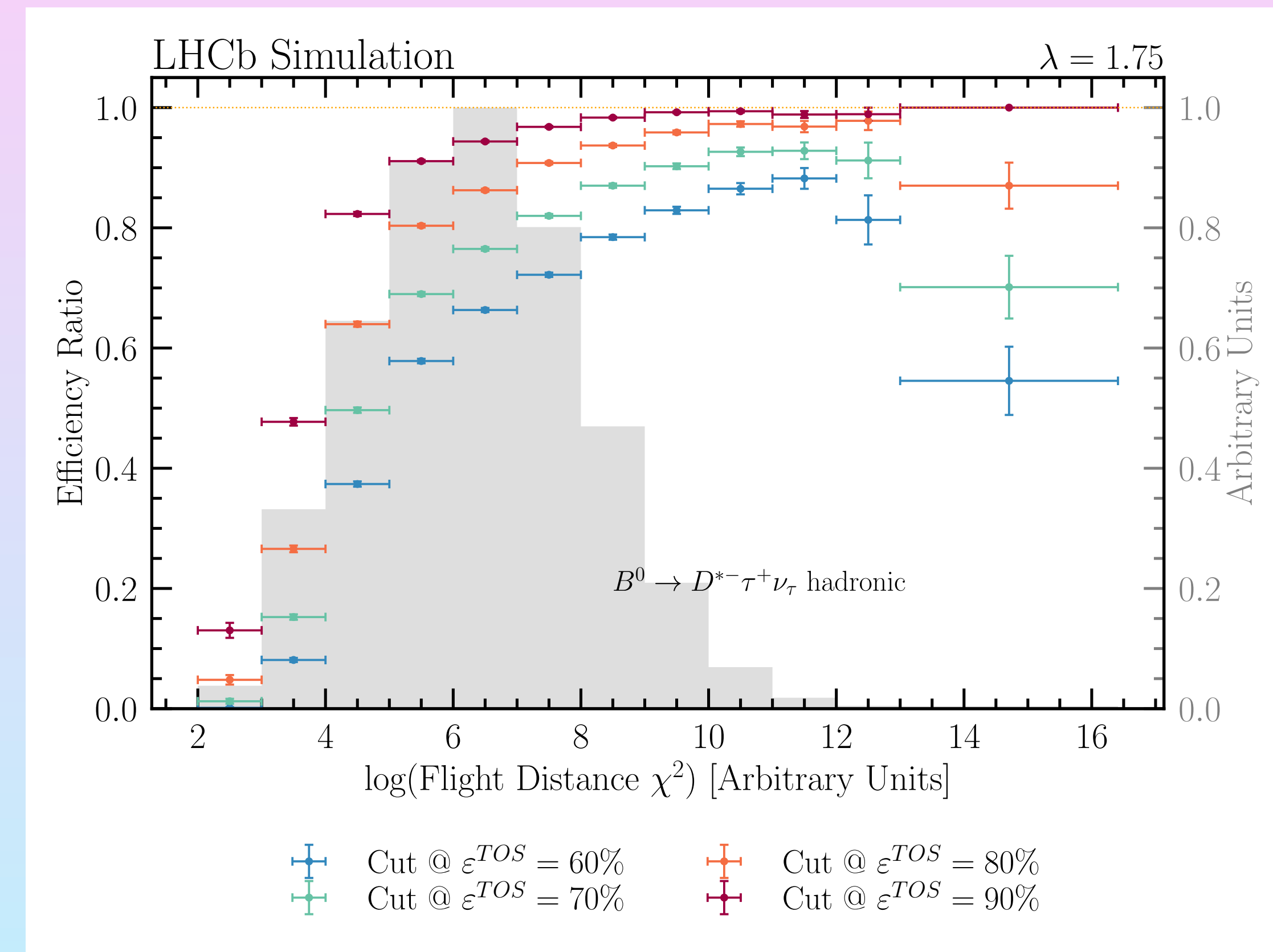Unconstrained Neural Network (~BDT)

Constrained Neural Network (Monotonic Net)

➡ Monotonic Lipschitz Neural Network provides **higher signal efficiency** for candidates with a **higher flight distance** (more displaced tracks)

The Monotonic Net in the Topological Trigger

Unconstrained Neural Network (~BDT)

Constrained Neural Network (Monotonic Net)

➡ Run 3 model of the Topological Trigger **sensitive to long-lived beauty candidates** and therefore also to possible BSM physics

Implementation in the LHCb Software

Bandwidth [MB/s] = Event Size [kB] x Output Rate [kHz]

Limited!

Output Rate

- Control the rate by cutting on the response of the neural network output

- Connected to the efficiency of the selection

  ➡ Tighter cut on response lower the efficiency

 ➡ Maximise physics output using minimal computational resources

Implementation in the LHCb Software

Bandwidth [MB/s] = Event Size [kB] x Output Rate [kHz]

Event Size

- Current Implementation:
  Found Signal Candidate in Event?

  ➡ Save entire event

  ➡ High event size

- Future **Possibilities**:

  ➡ Save the parts of the event that correspond to the signal candidate only

  ➡ Lowered event size
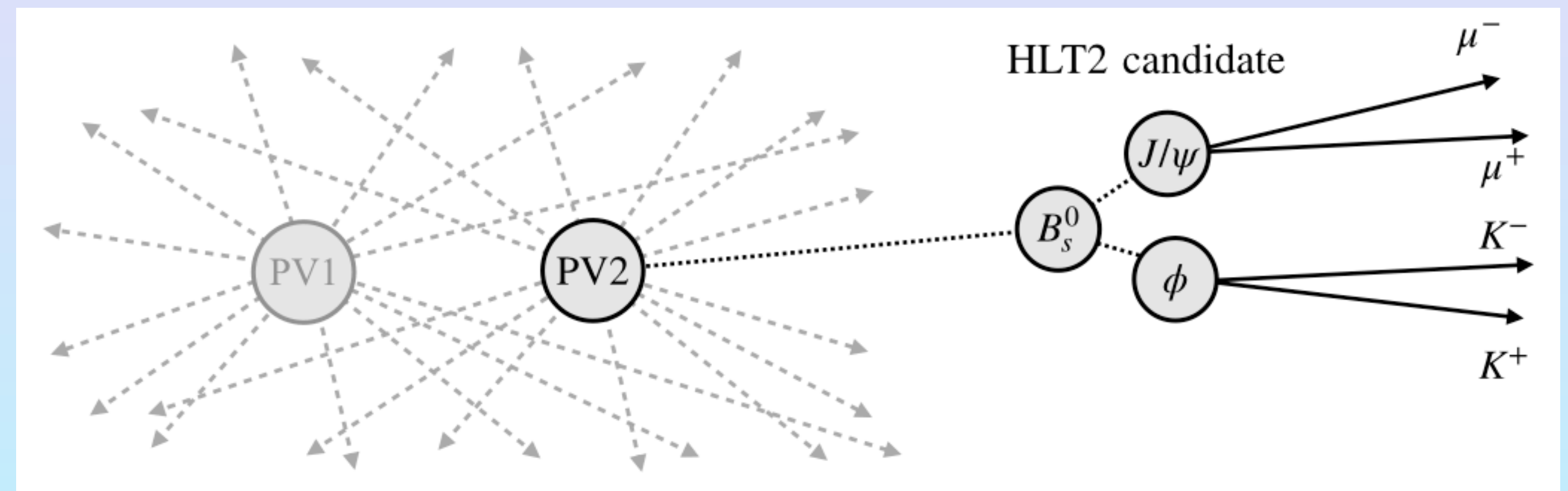
Summary

- Topological Trigger in Run 3 will be based on Neural Networks:

  - Provides robustness against detector effects and sensitivity to long-lived objects

  - Possibility to select BSM physics which would otherwise be disregarded

- Implementing the Neural Networks into the Software requires careful tuning considering limited computational resources

Thank you for your attention!

Backup

Data flow in Run 3

RTA and DPA dataflow diagrams for Run 1, Run 2,
and the upgraded LHCb detector
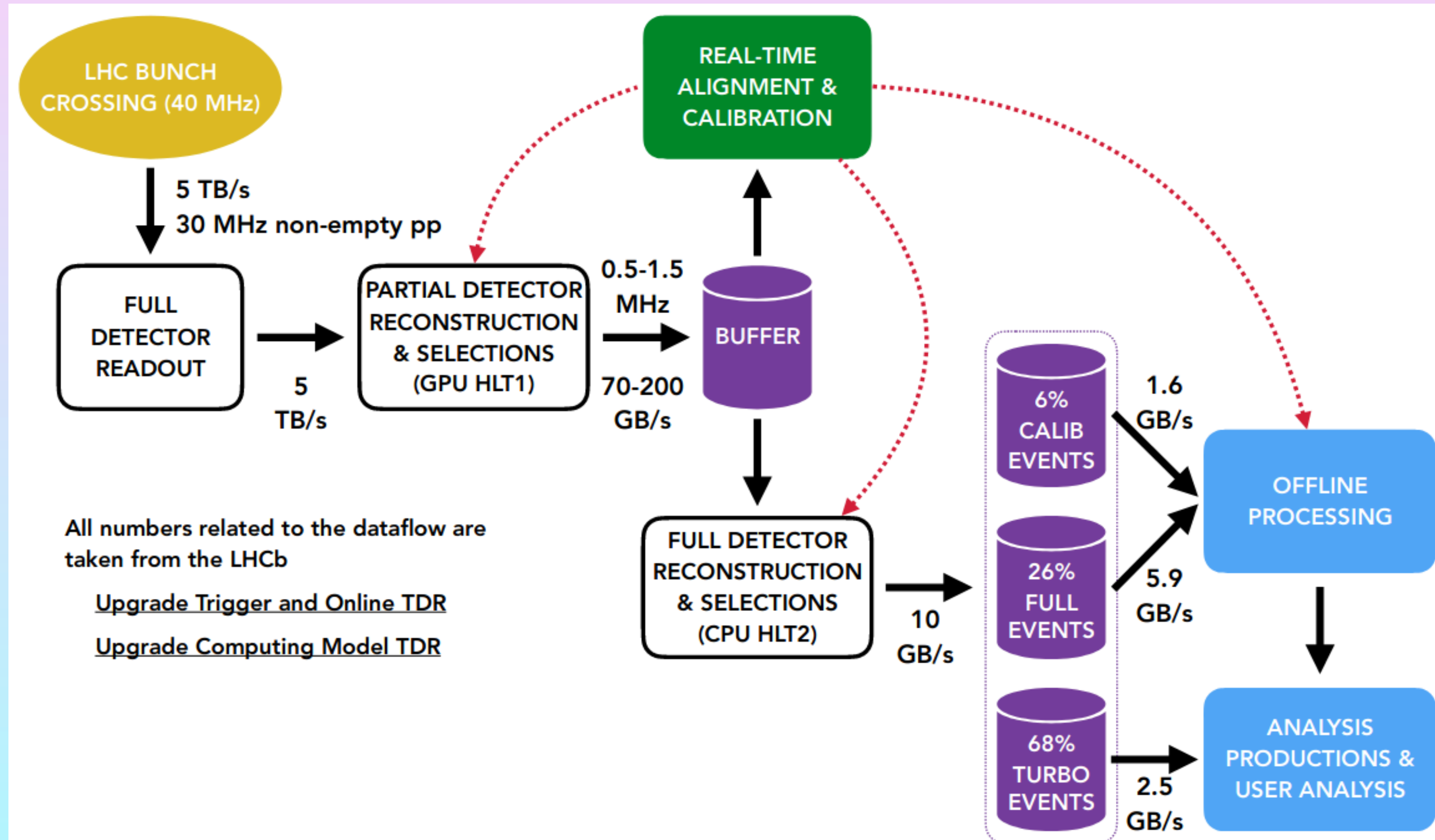
Reconstruction Flow in HLT2

- Two important types of algorithms: Particle Filters & Particle Combiners

- Particles and Tracks are reconstructed and then filtered according to certain properties (minimal transverse momentum etc.)

- Particles are then combined to a candidate of a decaying particle

- Combinations and vertices are again filtered according to certain properties

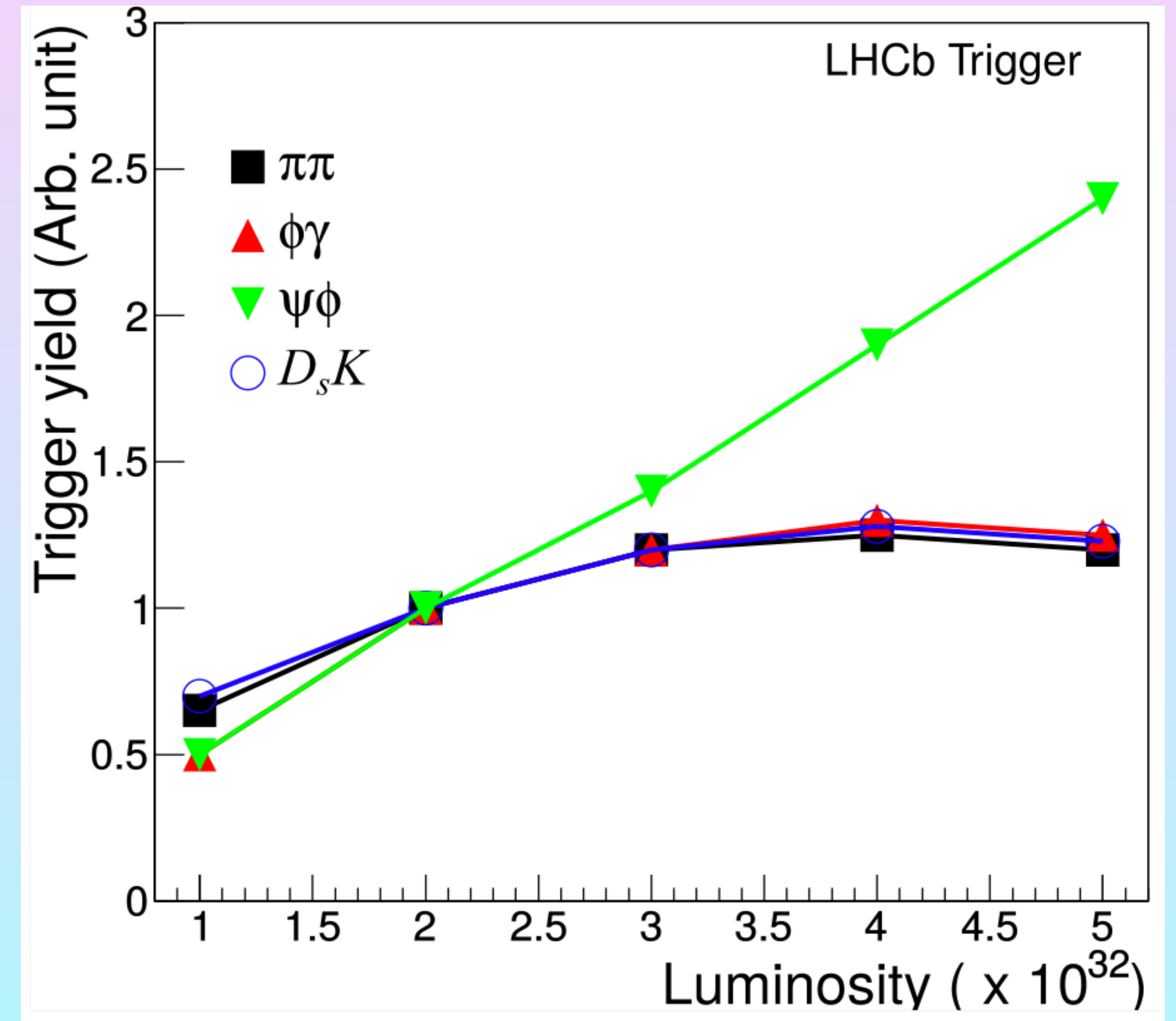- Combinations are then passed to selection algorithm to obtain Trigger decision
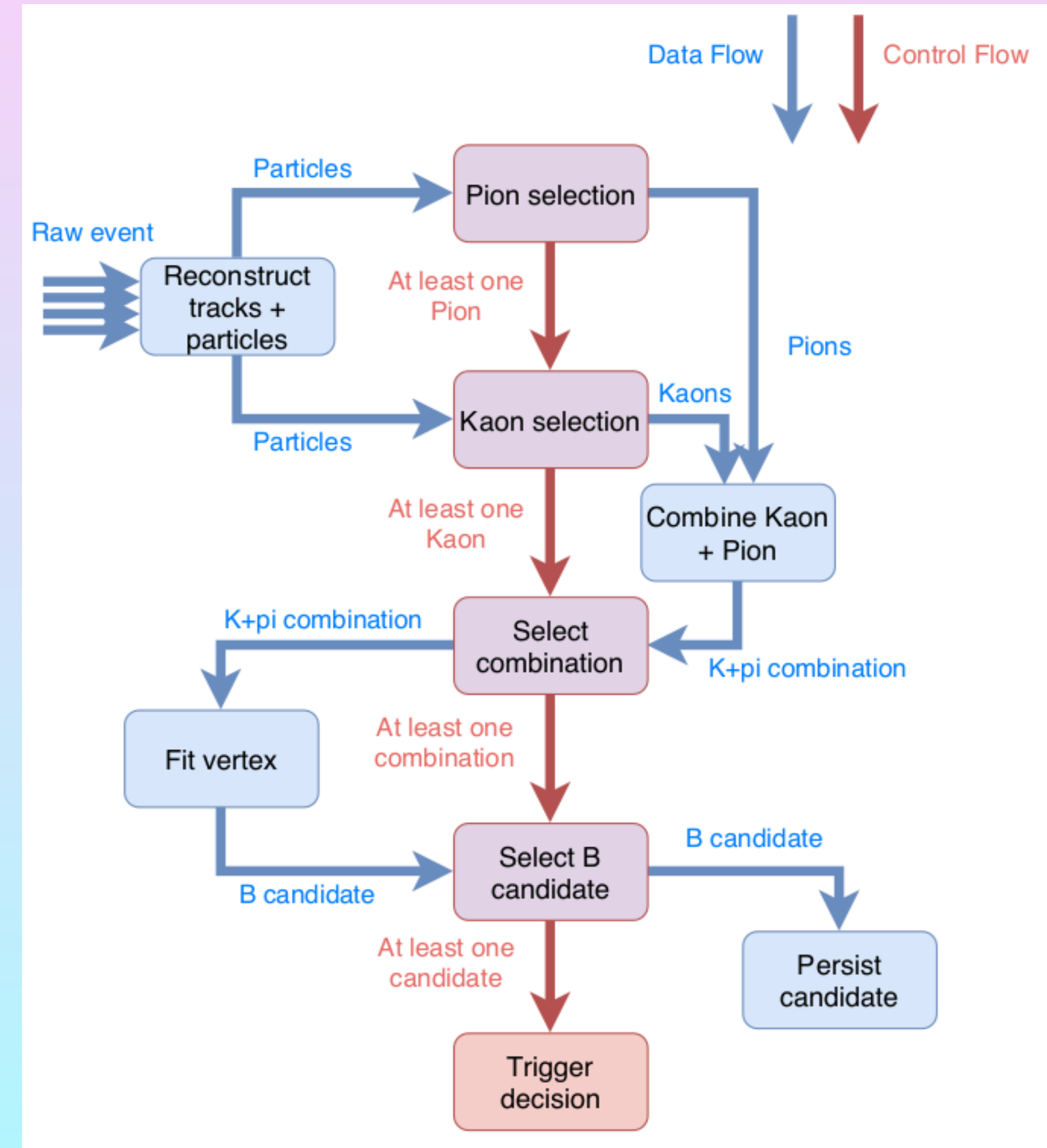
Samples - Detailed Overview

Training Samples	
$\Lambda_b^0 \to (\Lambda_c(2625)^+ \to (\Lambda_c^+ \to p^+ K^- \pi^+)\pi^+\pi^-)\mu^-\overline{\nu}_\mu$	$B^+ \to \mu^+\mu^-\mu^+\nu_\mu$
$B+ \to (\overline{D}^0 \to K^+\pi^-)(K_S^0 \to \pi^+\pi^-)\pi^+$	$\Lambda_b^0 \to p^+\mu^-\overline{\nu}_\mu$
$B^0 \to (D^0 \to K^-\pi^+\pi^+\pi^-)(\overline{D}^0 \to K^+\pi^-)$	$B^+ \to K^+\mu^+\mu^-\mu^+\mu^-\mu^+\mu^-$
$B^+ \to (\overline{D}^0 \to K^+\pi^-)\pi^+\pi^-\pi^+$	$B_c^+ \to (J/\psi(1S) \to \mu^+\mu^-)\mu^+\nu_\mu$
$B_s^0 \to (D_s- \to K^+K^-\pi^-)\nu_\mu\mu^+$	$B_s^0 \to K^-\nu_\mu\mu^+$
$B^+ \to (D^{\star+}(2010) \to (D^0 \to K^-\pi^+)\pi^+)(\overline{D}^0 \to K^+\pi^-)$	$B_s^0 \to K^-(\tau^+ \to \mu^+\nu_\mu\overline{\nu}_\tau)\nu_\tau$
$B^0 \to (D^{\star-} \to \pi^-(\overline{D}^0 \to K^+\pi^-))(\tau^+ \to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$	$B^+ \to \pi^+\pi^-K^+$
$B_c^+ \to (J/\psi(1S) \to \mu^+\mu^-)(\tau^+ \to \mu^+\nu_\mu\overline{\nu}_\tau)\nu_\tau$	$B^0 \to (K^{\star 0}(892) \to K^+\pi^-)\gamma$
$B^- \to (D^0 \to K^-\pi^+)(\tau^- \to \mu^-\nu_\tau\overline{\nu}_\mu)\overline{\nu}_\tau$	$B^- \to p^+\overline{p}^-(\tau^- \to \mu^-\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$
$B^+ \to (\overline{D}^{\star 0} \to \pi^0((\overline{D}^0 \to K^+\pi^-))(\tau^+ \to \pi^+\pi^-\pi^+\overline{\nu}_\tau)\nu_\tau$	$B^- \to p^+\overline{p}^-\mu^-\overline{\nu}_\mu$

Removal of the Hardware Trigger L0

- L0 used information from the calorimeters and muon chambers

- Based selection on transverse energy information

- Hadronic events required a high threshold in transverse energy to obtain reasonable output rate

➡ High threshold cut out high amount of beauty decays

➡ Signal yield saturates for increased luminosity

➡ Higher luminosity would not provide benefit for hadronic decays

CERN-LHCC-2011-001
Letter of Intent for the LHCb Upgrade

23

J.Phys: Conf. Ser. 878 012012
The LHCb Upgrade

Features - Detailed Table

TwoBody Features	ThreeBody Features
$\min\left(p_{\text{T, FS particles (1,2)}}\right)$	$\min\left(p_{\text{T, FS particles (1,2,3)}}\right)$
$\text{sum}\left(p_{\text{T, FS particles (1,2)}}\right)$	$\text{sum}\left(p_{\text{T, FS particles (1,2,3)}}\right)$
$p_{\text{T, B-Hadron}}$	$p_{\text{T, B-Hadron}}$
$\log\left(\min(\chi^2_{\text{IP, FS particles (1,2)}})\right)$	$\log\left(\min(\chi^2_{\text{IP, FS particles (1,2,3)}})\right)$
$\log\left(\max(\chi^2_{\text{IP, FS particles (1,2)}})\right)$	$\log\left(\max(\chi^2_{\text{IP, FS particles (1,2,3)}})\right)$
$\log\left(\chi^2_{\text{FD, B-Hadron}}\right)$	$\log\left(\chi^2_{\text{FD, B-Hadron}}\right)$
$\log\left(\chi^2_{\text{Vertex, B-Hadron}}\right)$	$\log\left(\chi^2_{\text{Vertex, B-Hadron}}\right)$
DOCA (B-Hadron)	DOCA (B-Hadron)
	$\min\left(p_{\text{T, FS particles (1,2)}}\right)$
	$\text{sum}\left(p_{\text{T, FS particles (1,2)}}\right)$
	DOCA (TwoBody)
	$\log\left(\chi^2_{\text{FD, TwoBody}}\right)$
	$\log\left(\chi^2_{\text{Vertex, TwoBody}}\right)$
	$\log\left(\chi^2_{\text{IP, TwoBody}}\right)$
	$p_{\text{T, TwoBody}}$