A Deep-learning based Full-Event Interpretation (DFEI) algorithm

for the identification and hierarchical reconstruction of heavy-hadron decay chains in proton-proton collisions

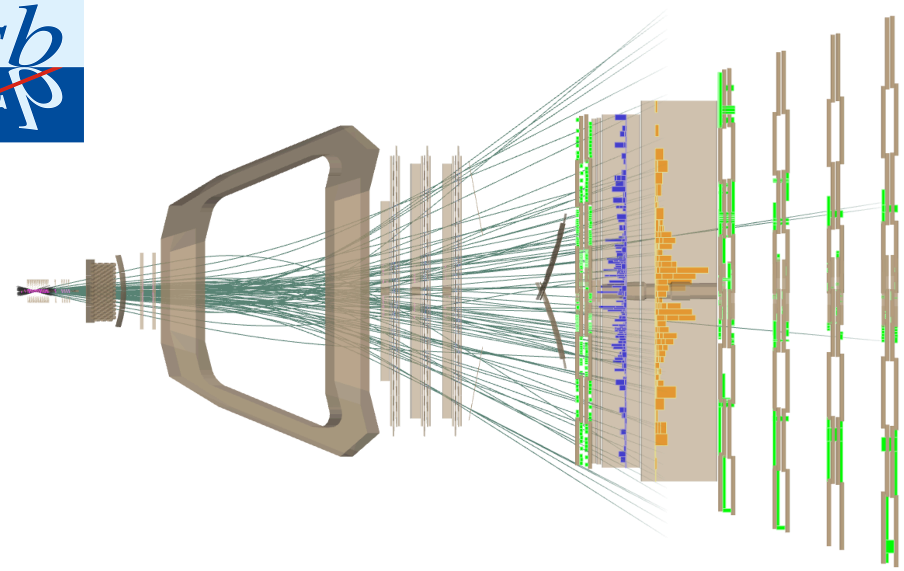
Julián García Pardiñas[1], Andrea Mauri[2], Marta Calvi[1], Jonas Eschle[3], Simone Meloni[1], Nicola Serra[3]

1 University and INFN Milano-Bicocca, Italy
2 NIKHEF, The Netherlands
3 University of Zürich, Switzerland

ICHEP 2022

Bologna, Italy
8th of July 2022

The LHCb trigger: entering a new era

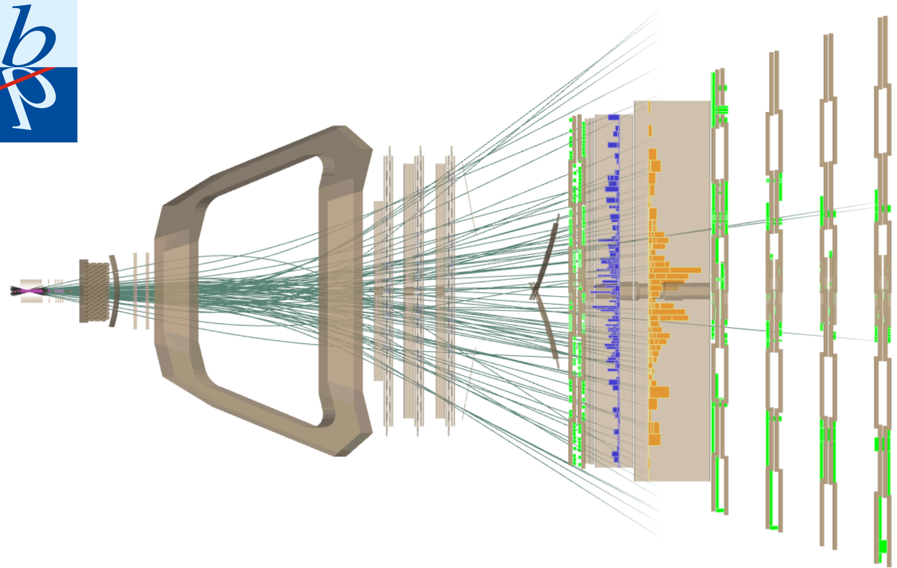


The LHCb detector

Forward spectrometer, studying the decays of **beauty and charm hadrons**.

Current trigger strategy: **signal based**.

The LHCb trigger: entering a new era



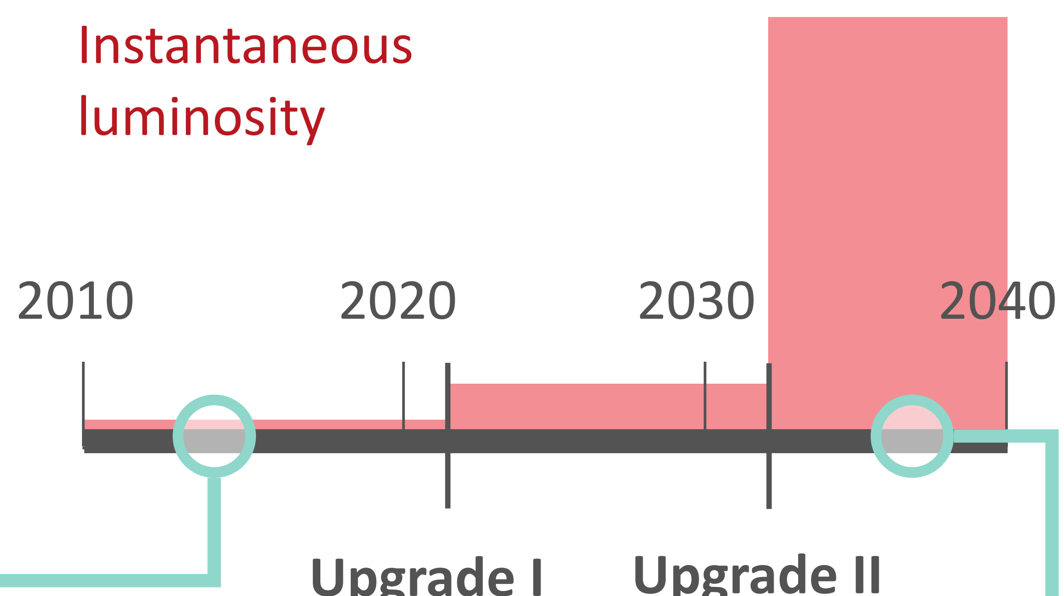
The LHCb detector

Forward spectrometer, studying the decays of **beauty and charm hadrons**.

Current trigger strategy: **signal based**.

≪ 1 signal
~ 50 tracks

Which events are interesting?

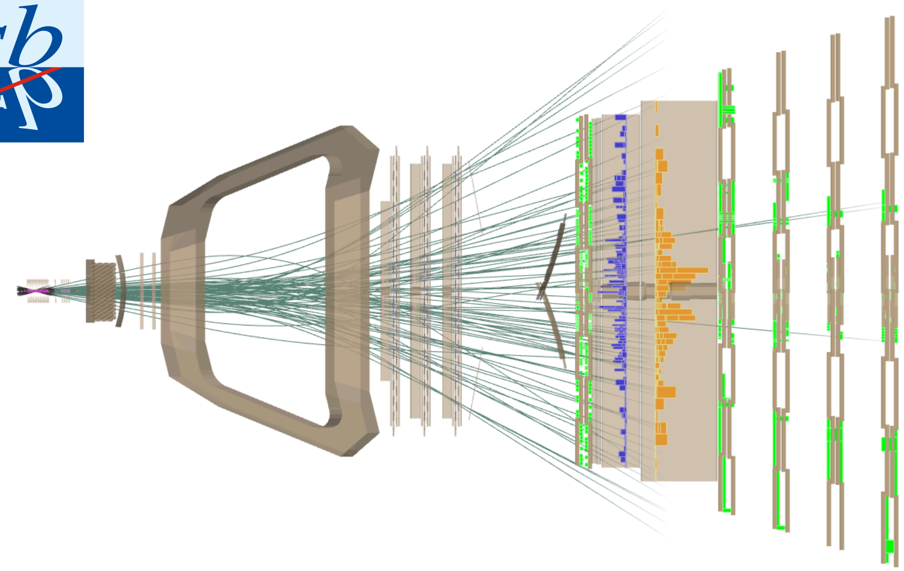


~ 5 signals
~ 1000 tracks

Which parts of the event are interesting?

[LHCb-PUB-2014-027]

The LHCb trigger: entering a new era



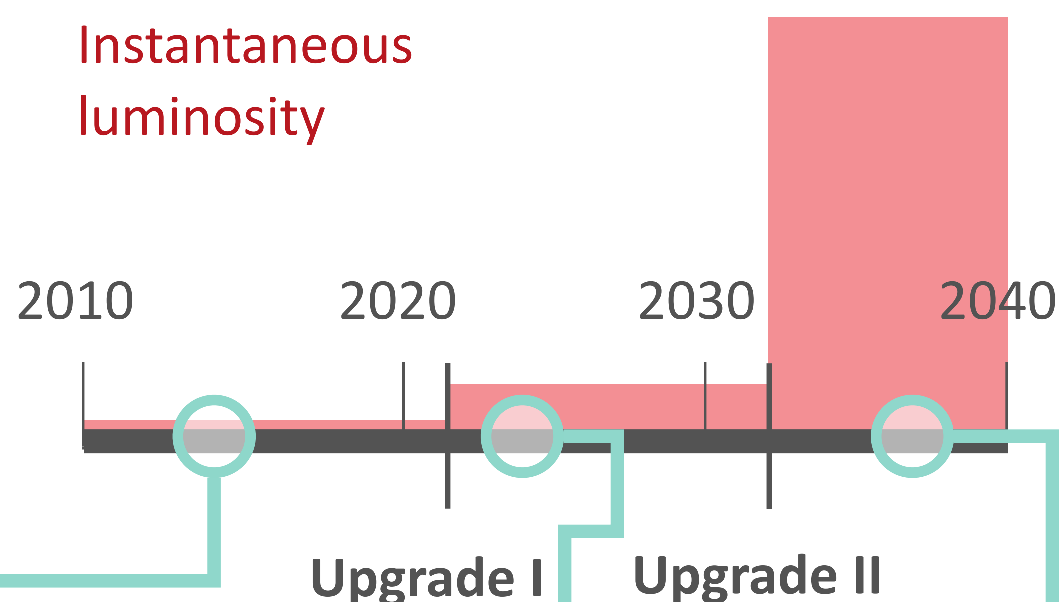
The LHCb detector

Forward spectrometer, studying the decays of **beauty and charm hadrons**.

Current trigger strategy: **signal based**.

≪ **1 signal**
~ **50 tracks**

Which events are interesting?



~ **5 signals**
~ **1000 tracks**

Which parts of the event are interesting?

[LHCb-PUB-2014-027]

~ **0.5 signals**
~ **140 tracks**

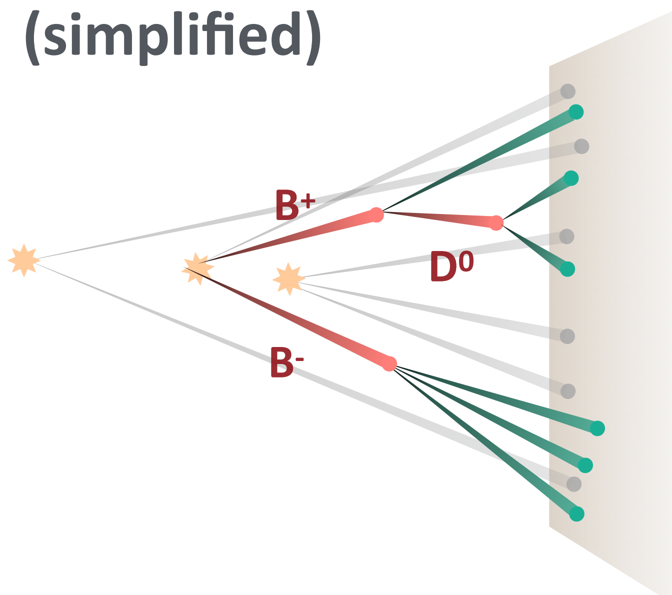
- **Fully software trigger, CPU + GPU** [JINST 14 (2019) 04, P04006].
- Data buffer to enlarge the time window for online processing.
  ➡ Online alignment and calibration, **offline-quality online reconstruction**.

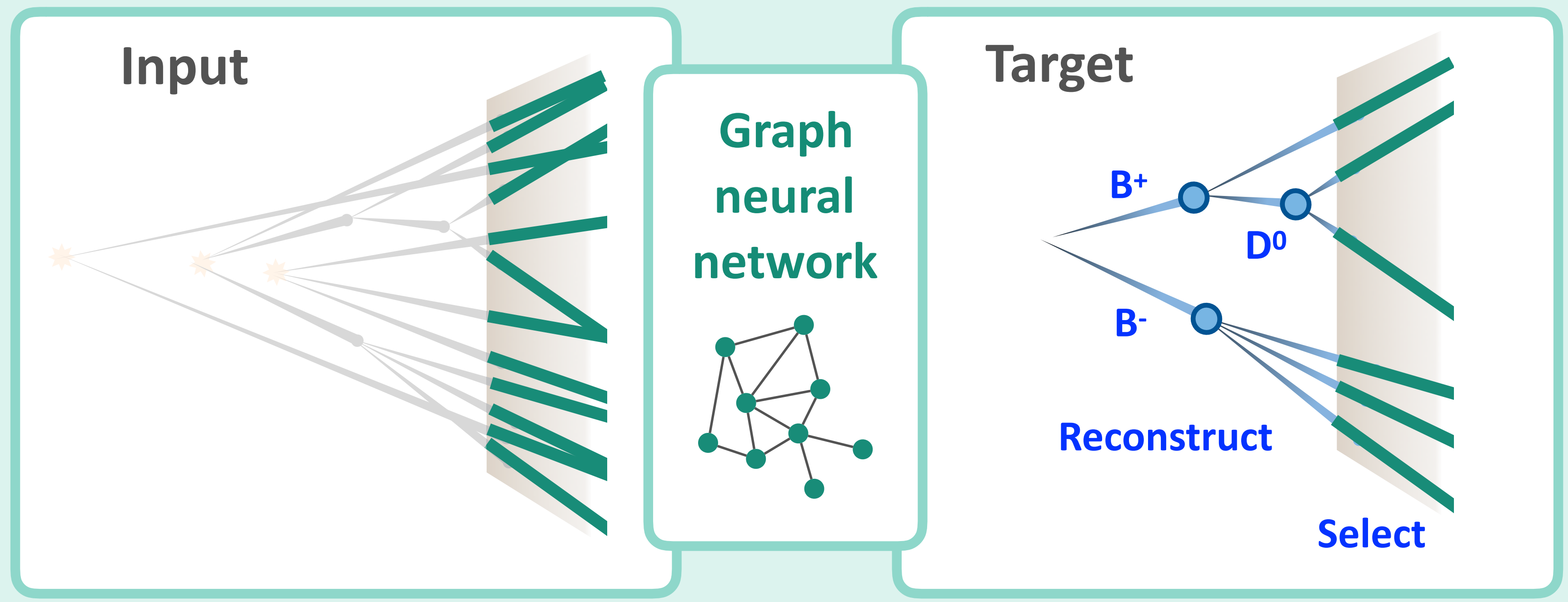
Facing the new era with machine learning

Novel approach proposed

DFEI:
Deep-learning based Full Event Interpretation

LHCb event (simplified)

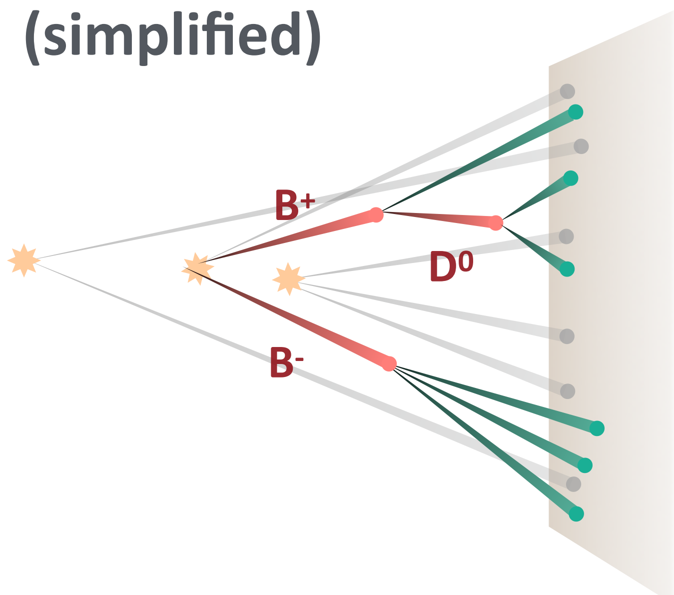




"Maximally efficient" trigger.

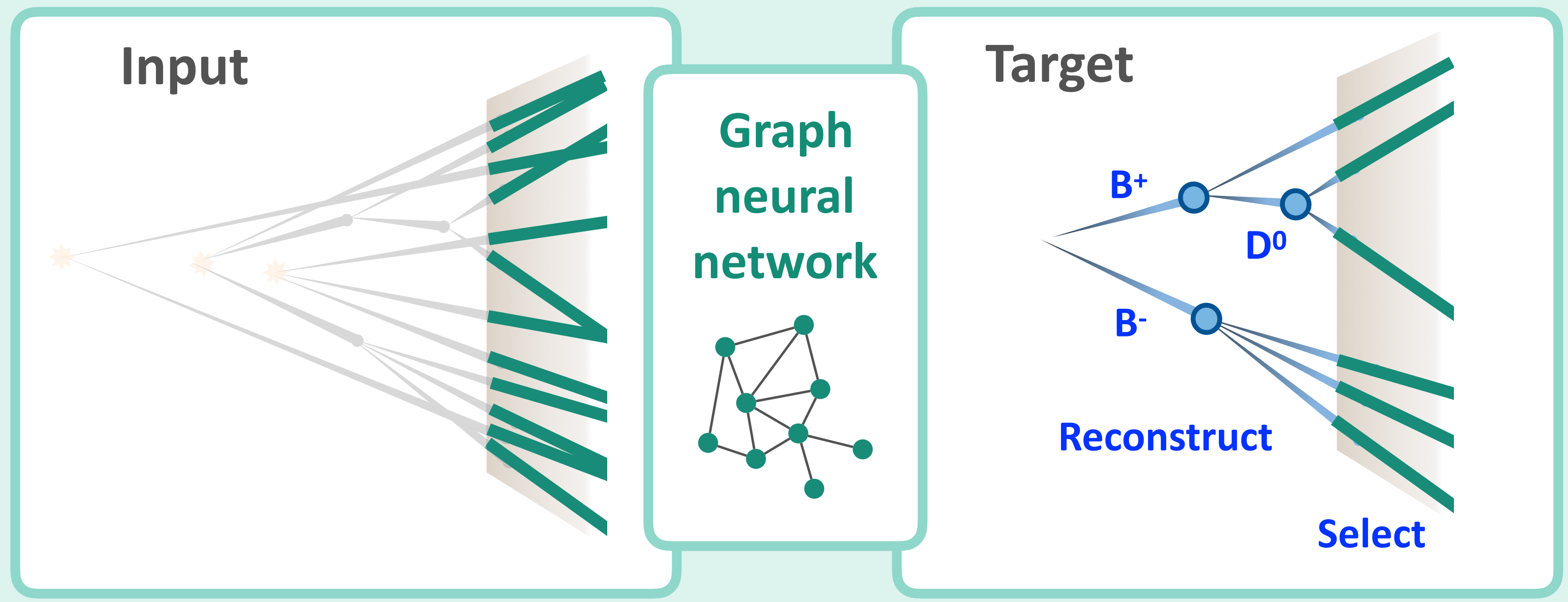
Facing the new era with machine learning

LHCb event (simplified)



Novel approach proposed

DFEI:
Deep-learning based Full Event Interpretation



"Maximally efficient" trigger.

Similar developments in other experiments

Full Event Interpretation algorithm at an e+e- collider
[Comput.Softw.Big Sci. 3 (2019) 1 6], BELLE2-MTHESIS-2020-006].

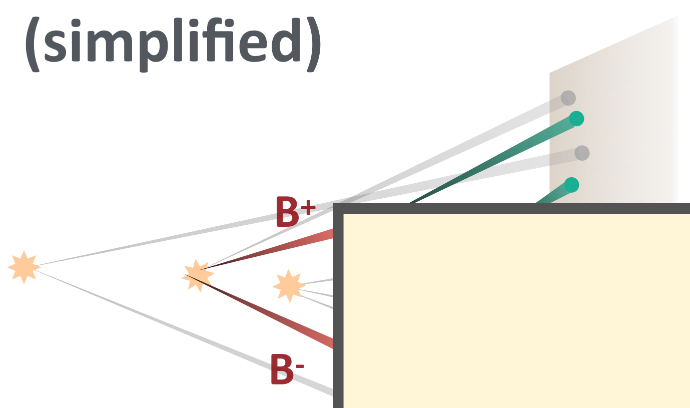
GNNs for trigger purposes
[see e.g. Eur.Phys.J.C 81 (2021) 5, 381, Frontiers in Big Data 3 (2021) 44].

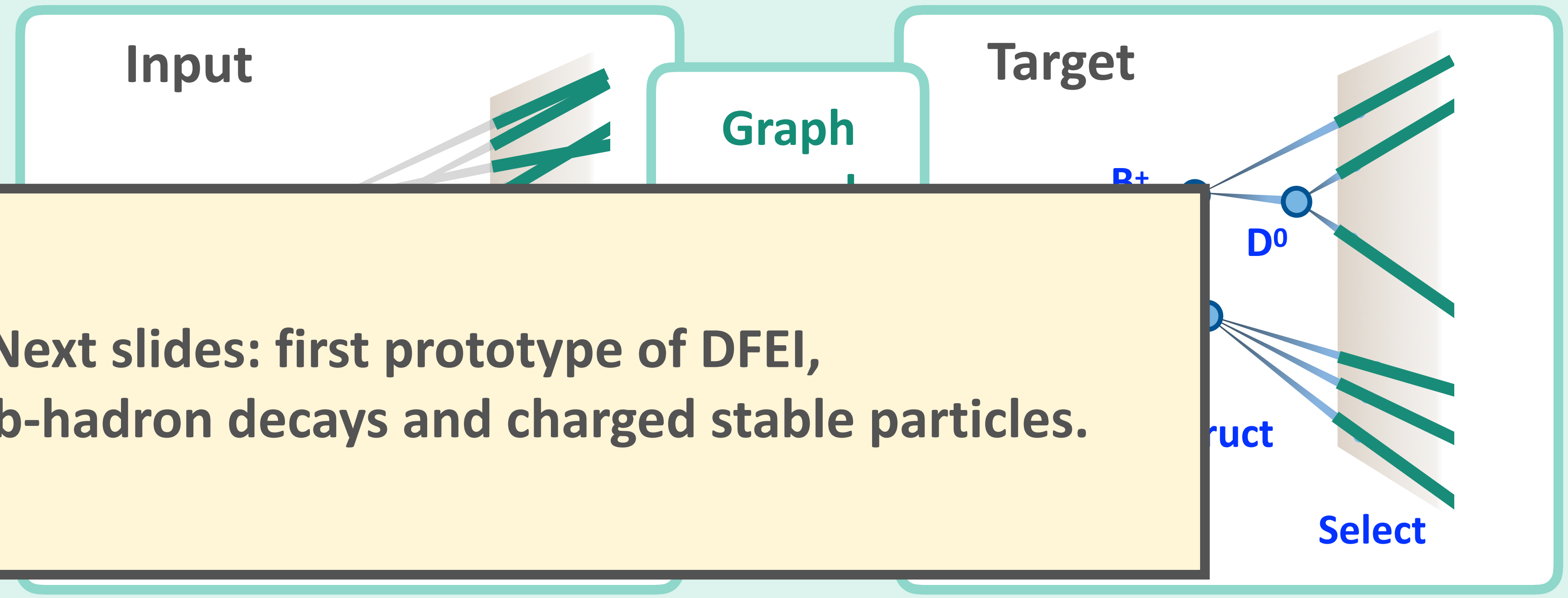
Facing the new era with machine learning

Novel approach proposed

DFEI:
Deep-learning based
Full Event Interpretation

LHCb event (simplified)





Next slides: first prototype of DFEI,
focused on b-hadron decays and charged stable particles.

"Maximally efficient" trigger.

Similar developments in other experiments

Full Event Interpretation algorithm at an e+e- collider
[Comput.Softw.Big Sci. 3 (2019) 1 6], BELLE2-MTHESIS-2020-006].

GNNs for trigger purposes
[see e.g. Eur.Phys.J.C 81 (2021) 5, 381, Frontiers in Big Data 3 (2021) 44].

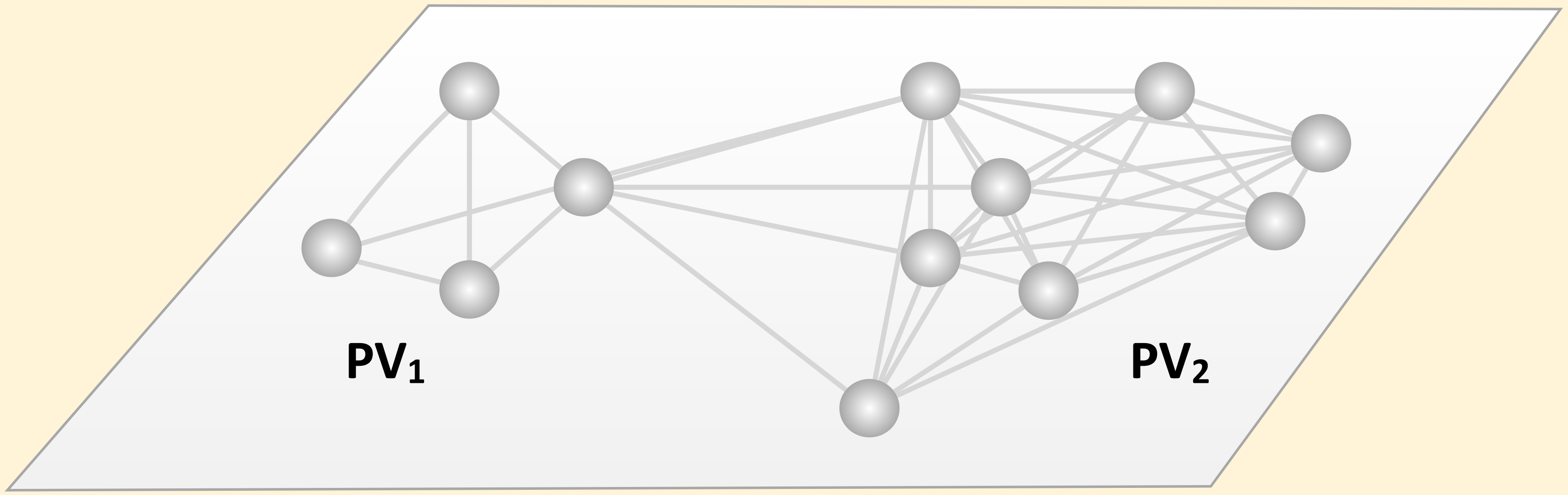
Input graph construction

Nodes: all the charged particles in the event.

➥ On average **~140**.

Edges: connect particles which are topologically close (see backup for details).

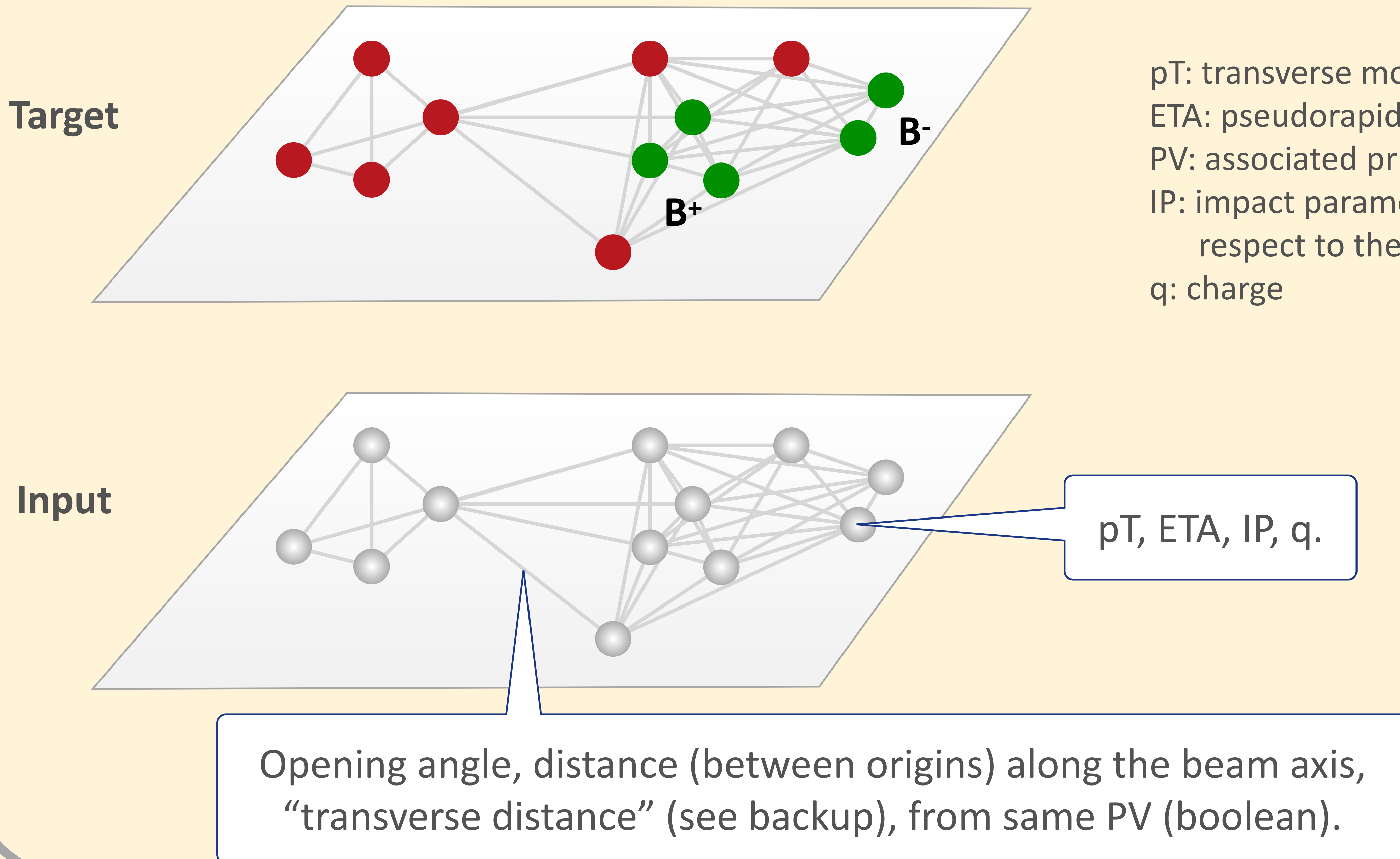
➥ On average **~10 000**.



PV_1, PV_2: different proton-proton primary vertices.

1^{st} module: node pruning

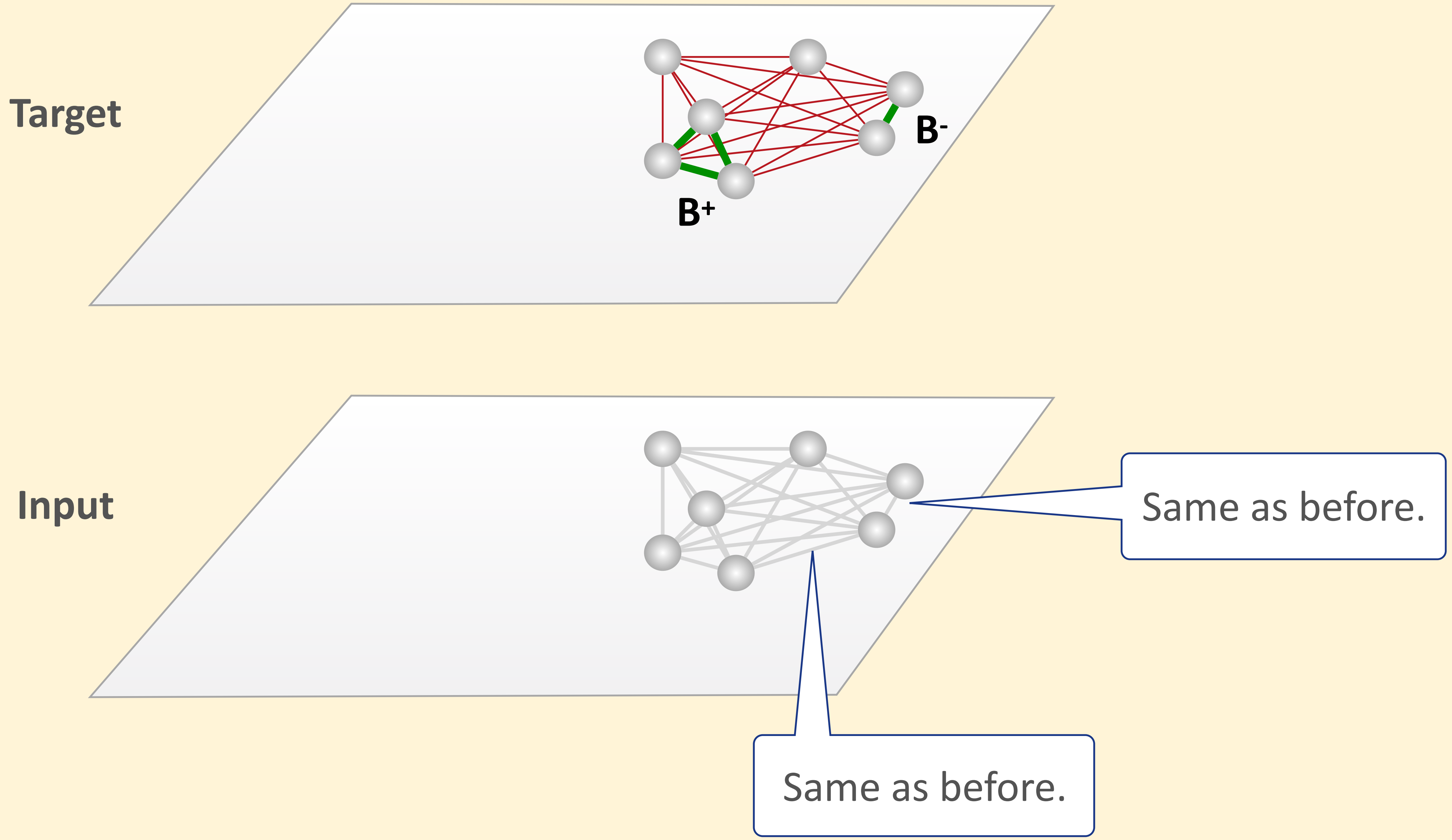
Signal nodes: particles from a b-hadron (any of them)
Background nodes: particles from the rest of the event



pT: transverse momentum
ETA: pseudorapidity
PV: associated primary vertex
IP: impact parameter with respect to the PV
q: charge

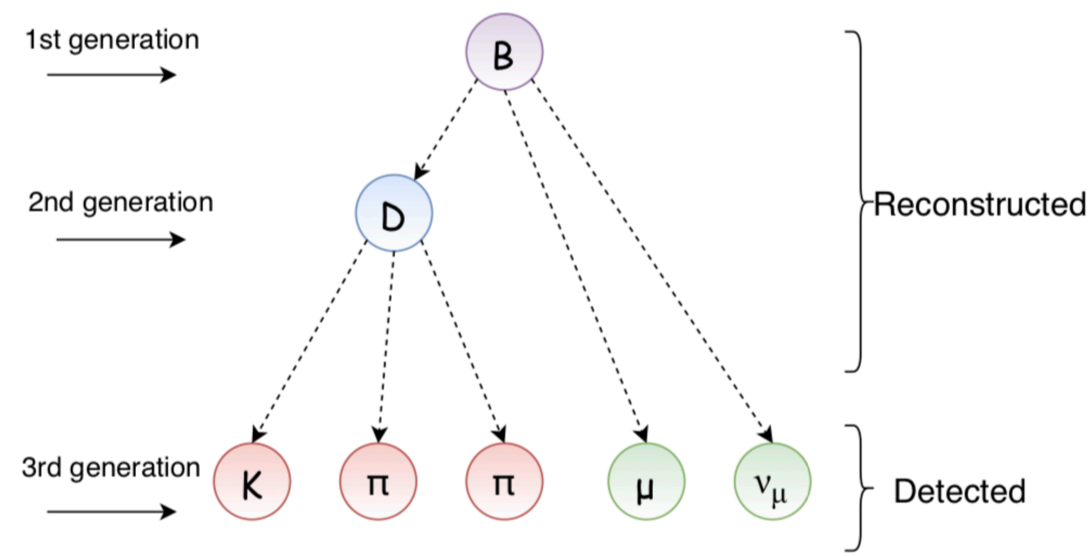
2nd module: edge pruning

Signal edges: pairs of particles with the same b-hadron ancestor
Background edges: any other pair of particles



3rd module: Lowest Common Ancestor (LCA) inference

From [BELLE2-MTHESIS-2020-006]:



Adjacency Matrix

	B	D	K	π	π	μ	ν_μ
B	0	1	0	0	0	1	1
D	1	0	1	1	1	0	0
K	0	1	0	0	0	0	0
π	0	1	0	0	0	0	0
π	0	1	0	0	0	0	0
μ	1	0	0	0	0	0	0
ν_μ	1	0	0	0	0	0	0

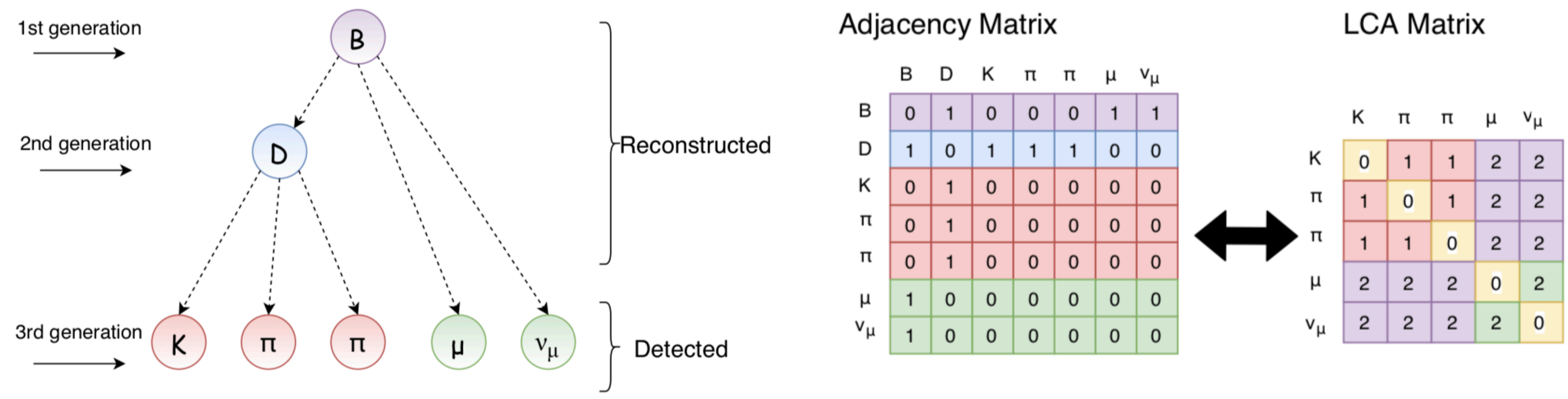
LCA Matrix

	K	π	π	μ	ν_μ
K	0	1	1	2	2
π	1	0	1	2	2
π	1	1	0	2	2
μ	2	2	2	0	2
ν_μ	2	2	2	2	0

Problem reduced to **multi-class classification on edges**.

3rd module: Lowest Common Ancestor (LCA) inference

From [BELLE2-MTHESIS-2020-006]:

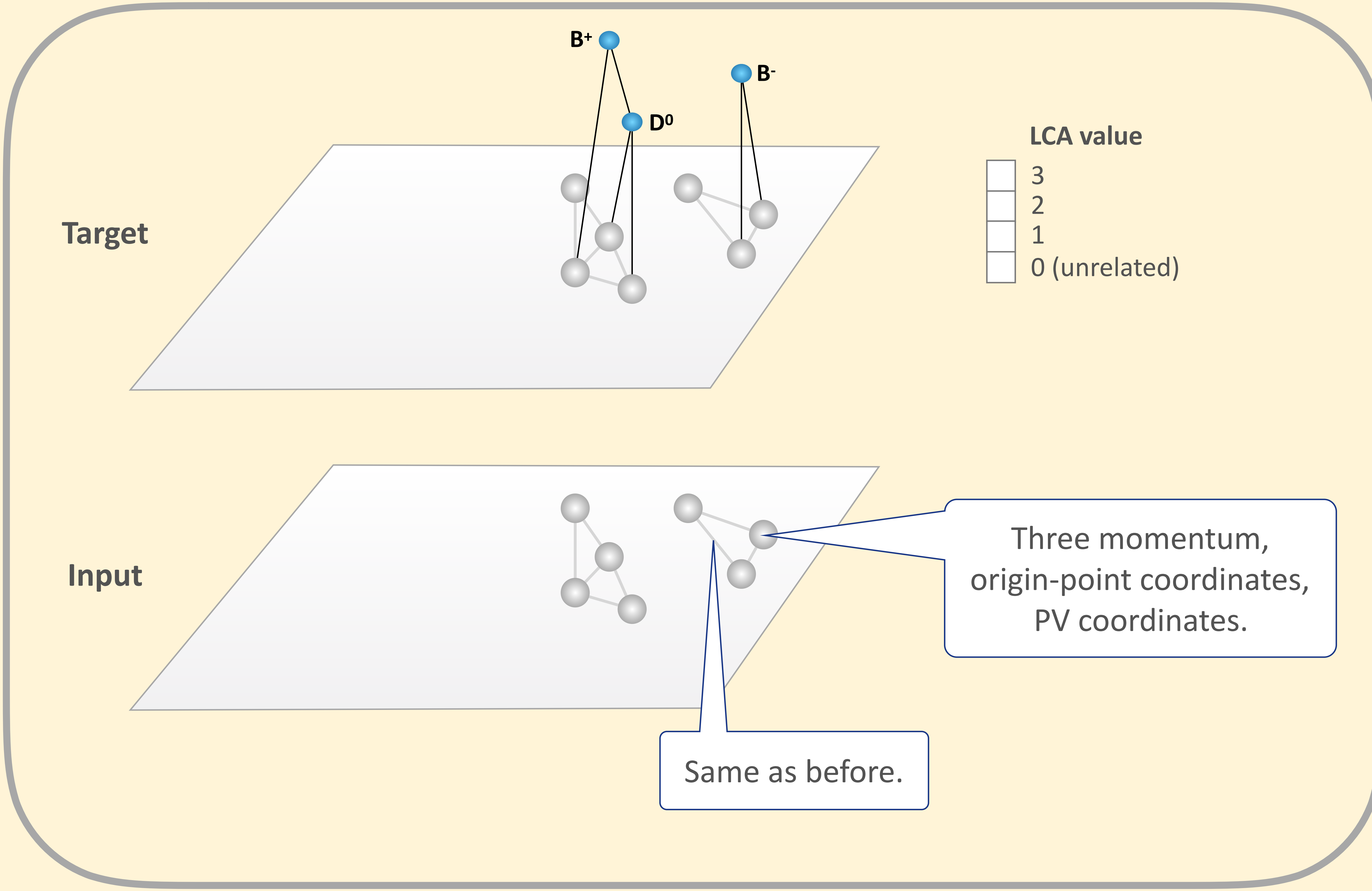


Problem reduced to **multi-class classification on edges**.

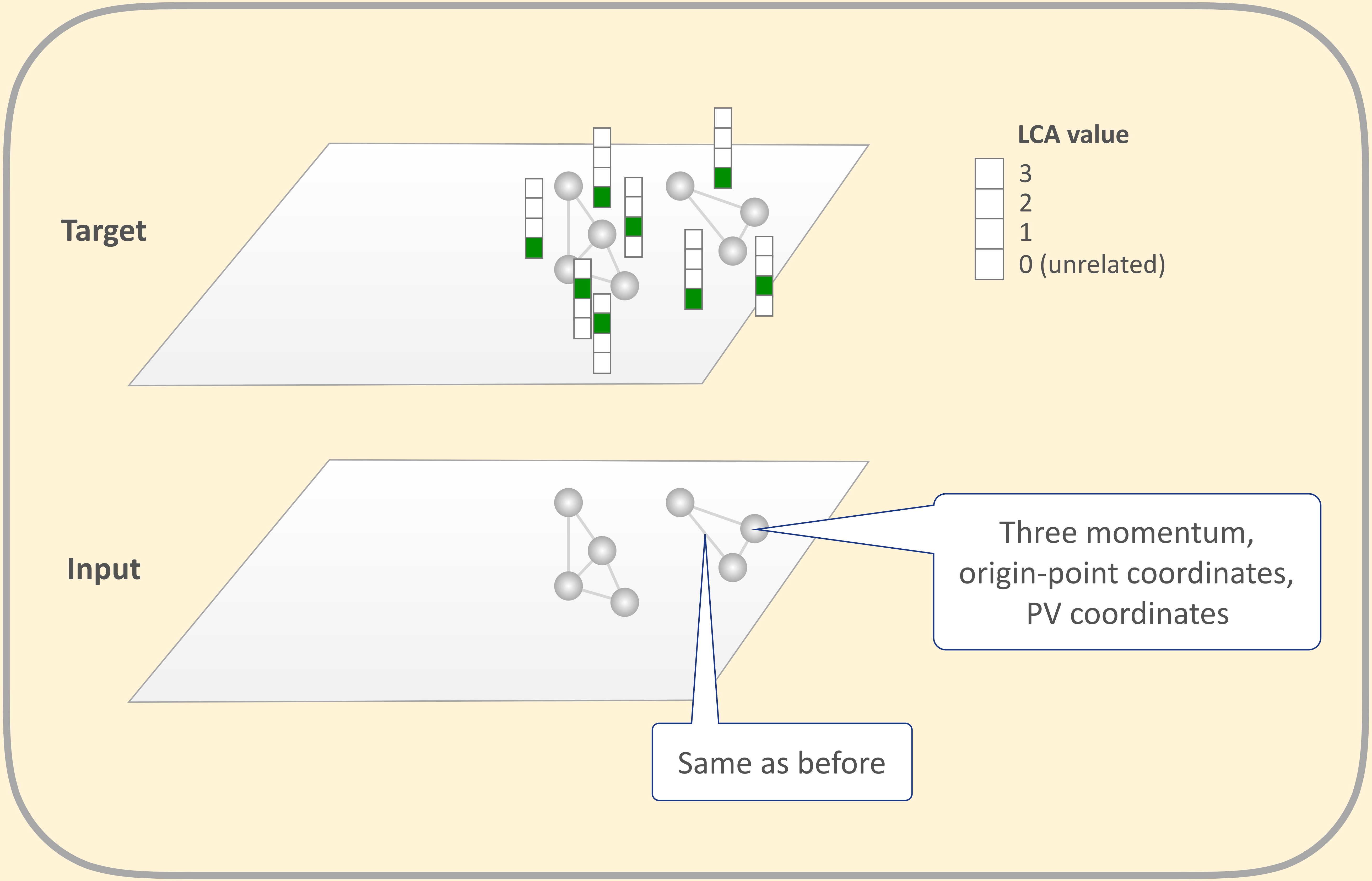
For the prototype, use as target a **simplified version of the decay chain, based on the reconstructible vertices**.

- Very-short-lived resonances merged with the previous ancestor.
- Resonances with less than two charged descendants merged with the previous ancestor.

3rd module: Lowest Common Ancestor (LCA) inference

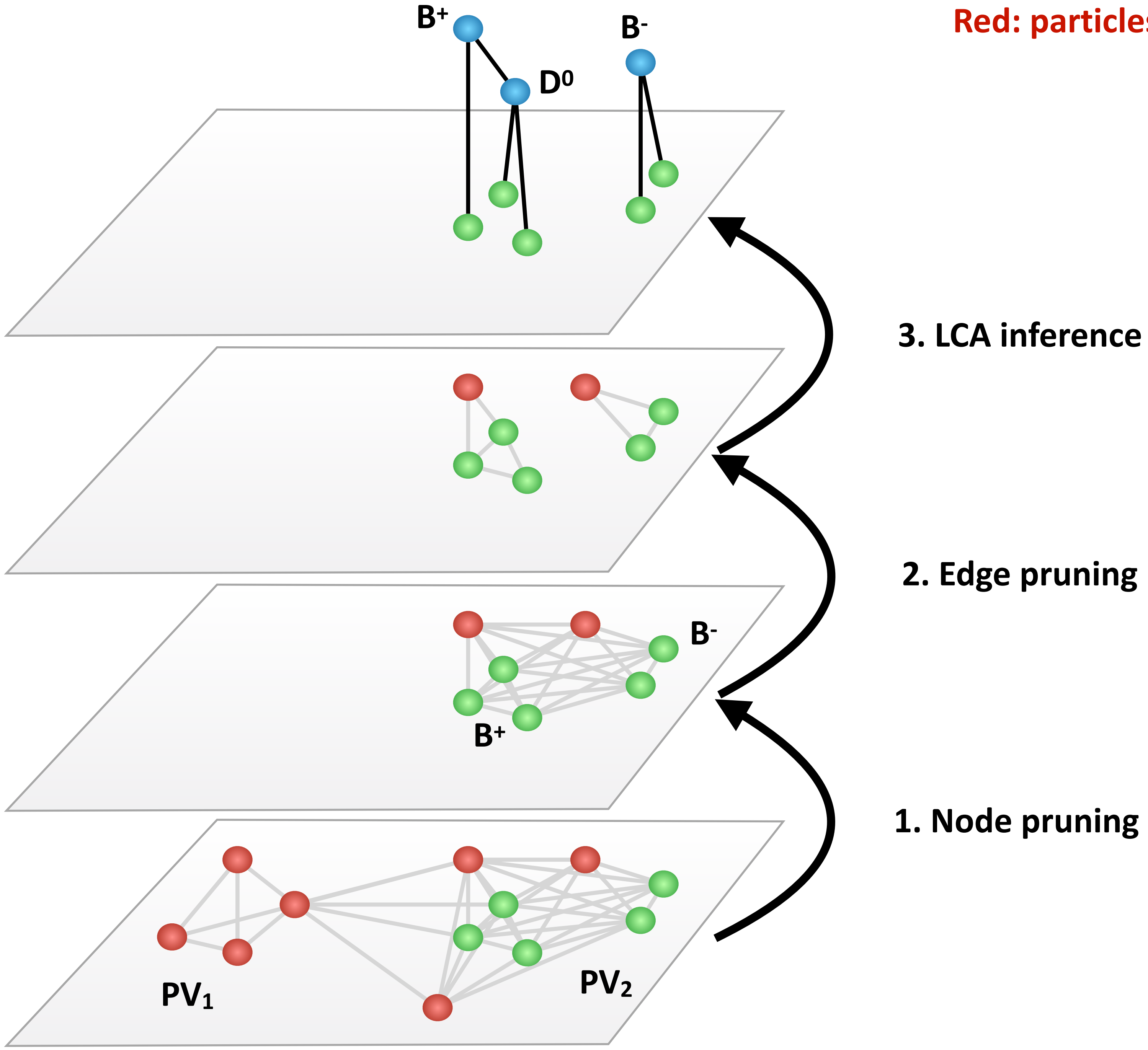


3rd module: Lowest Common Ancestor (LCA) inference



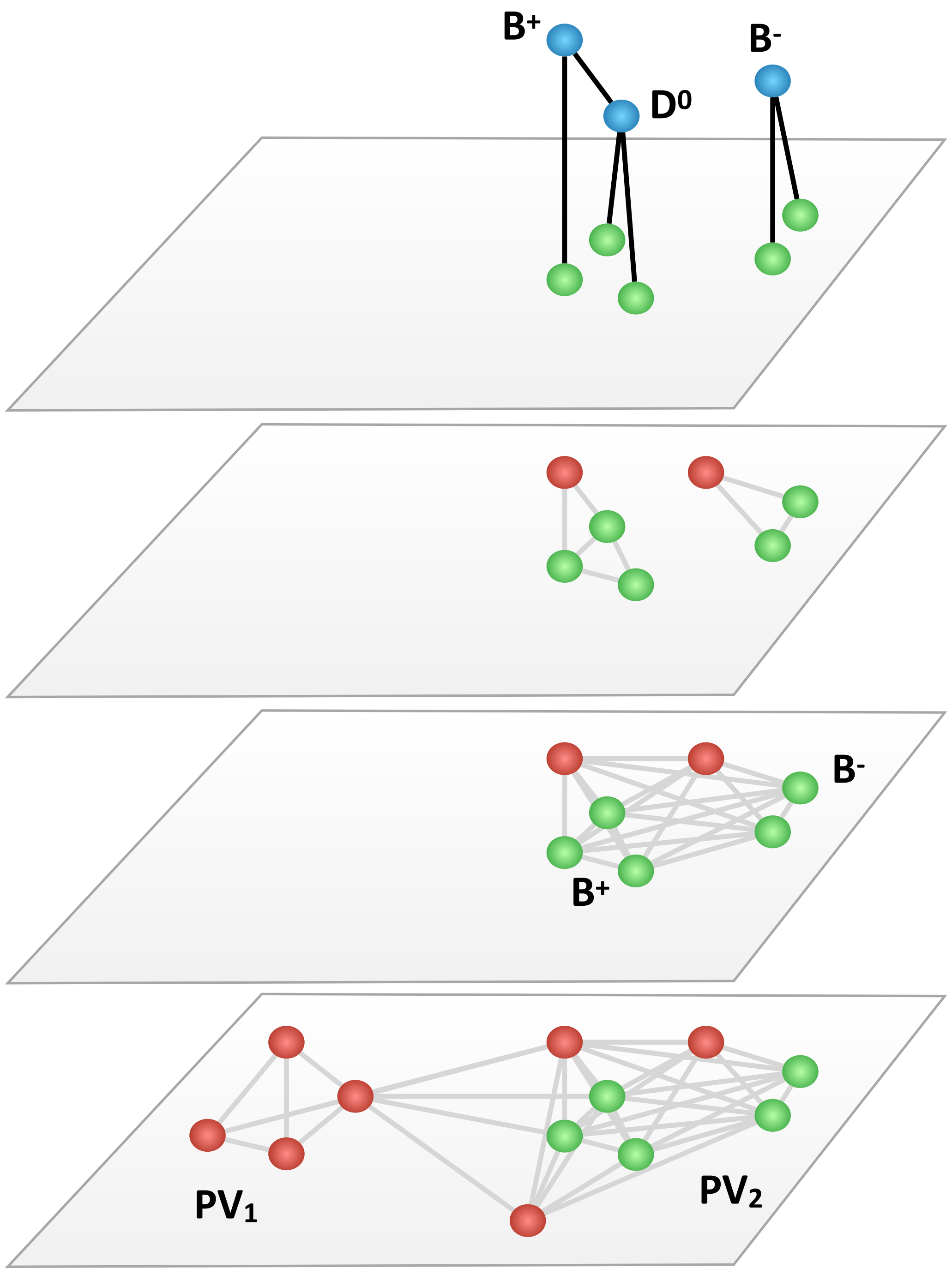
Global overview of the algorithm

Blue: reconstructed ancestors.
Green: particles from a b-hadron
Red: particles from the rest of the event



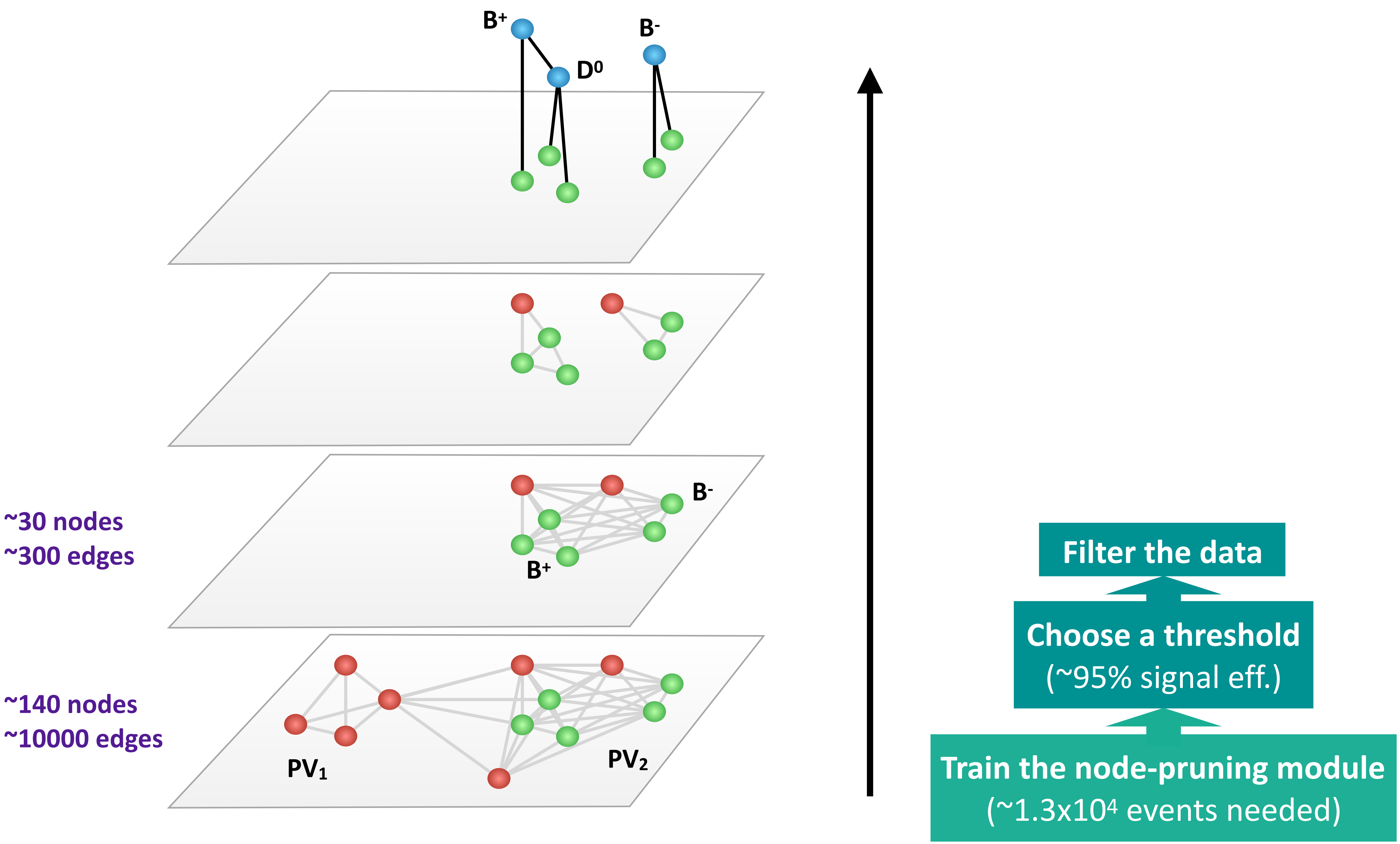
Training

Dataset:
- PYTHIA-based simulation, Run 3 conditions, approximated emulation of LHCb reconstruction.
- Events required to contain at least one b-hadron (inclusive decay).



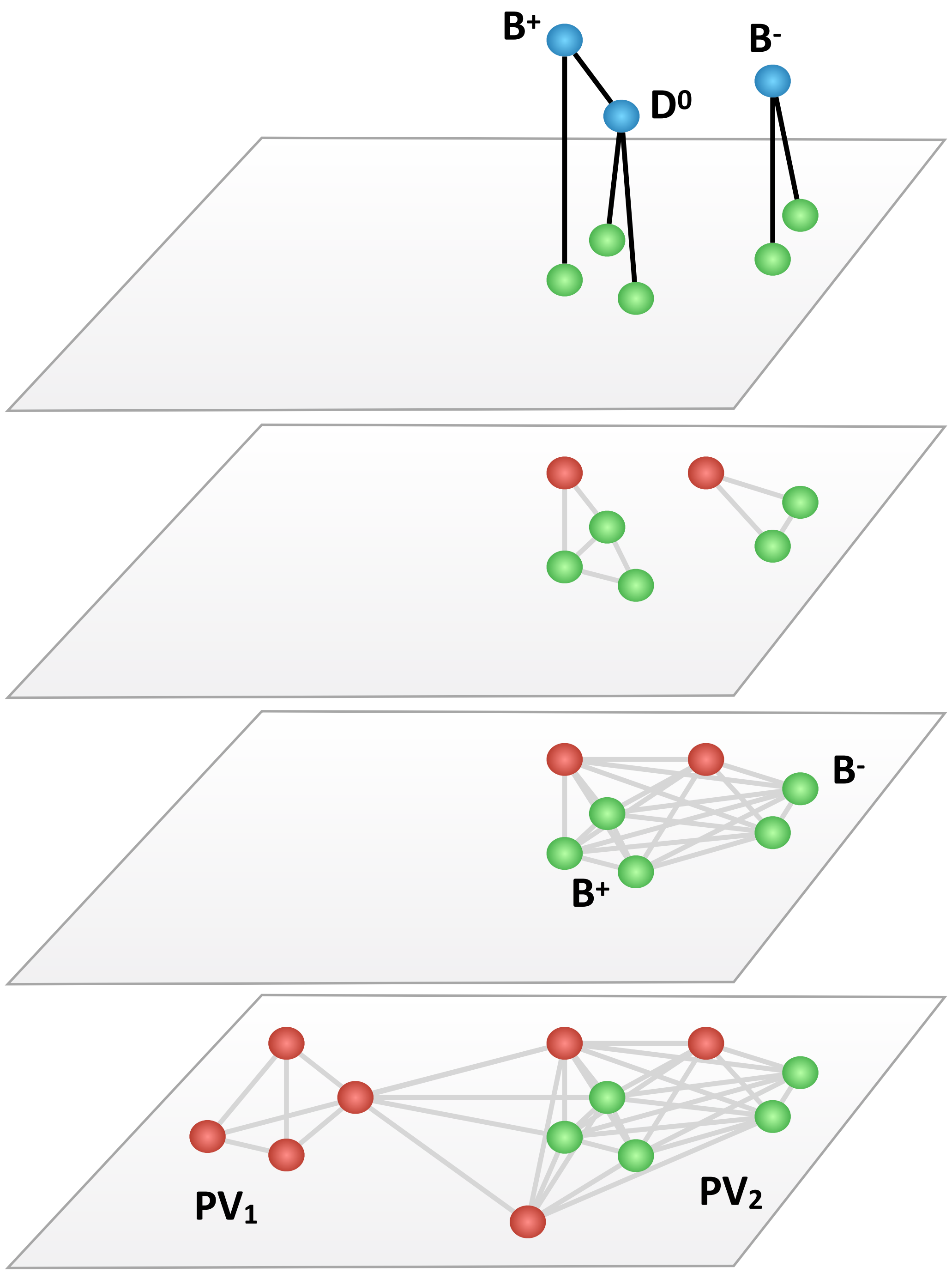
Training

Dataset:
- PYTHIA-based simulation, Run 3 conditions, approximated emulation of LHCb reconstruction.
- Events required to contain at least one b-hadron (inclusive decay).



Training

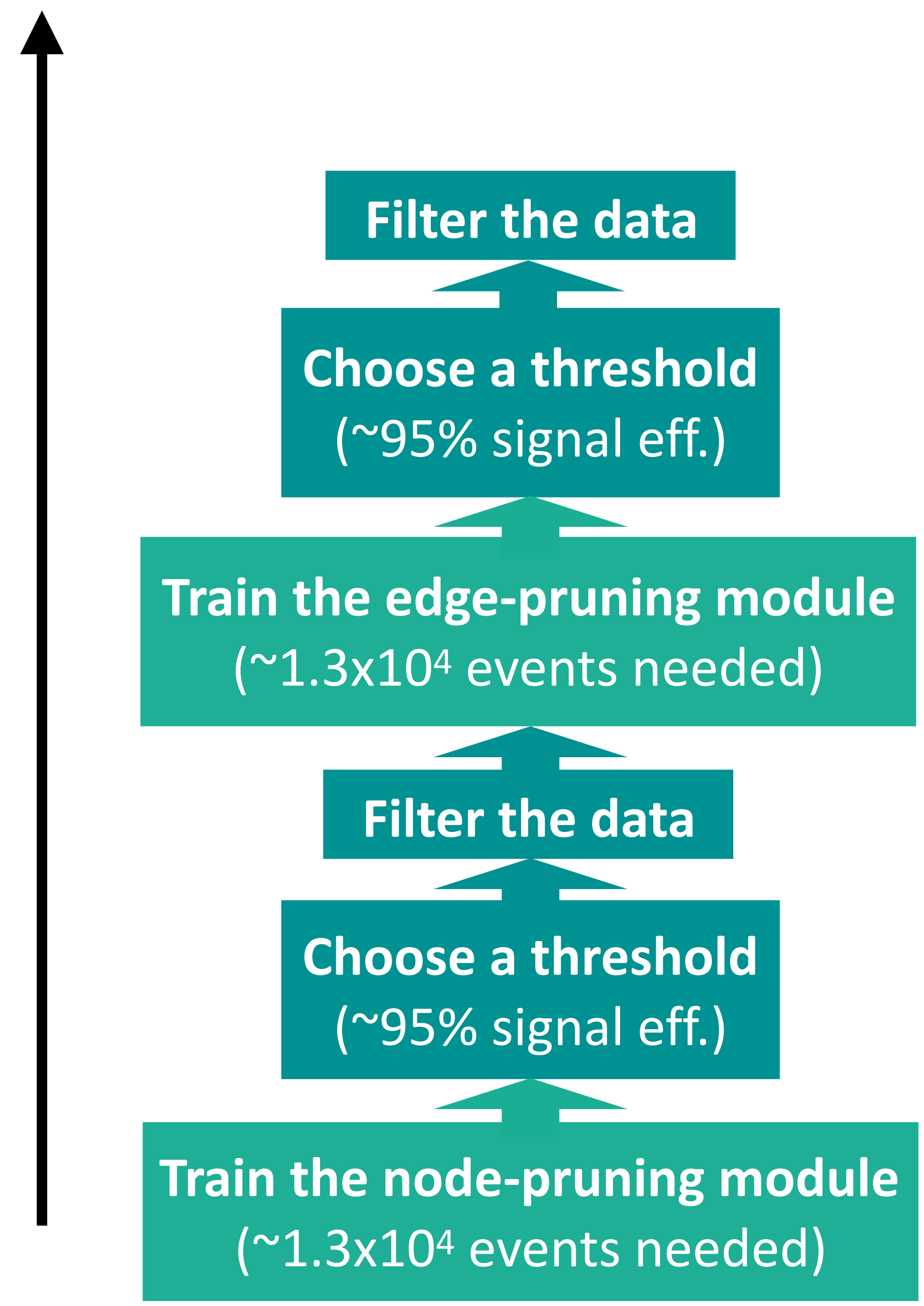
Dataset:
- PYTHIA-based simulation, Run 3 conditions, approximated emulation of LHCb reconstruction.
- Events required to contain at least one b-hadron (inclusive decay).



~20 nodes
~75 edges

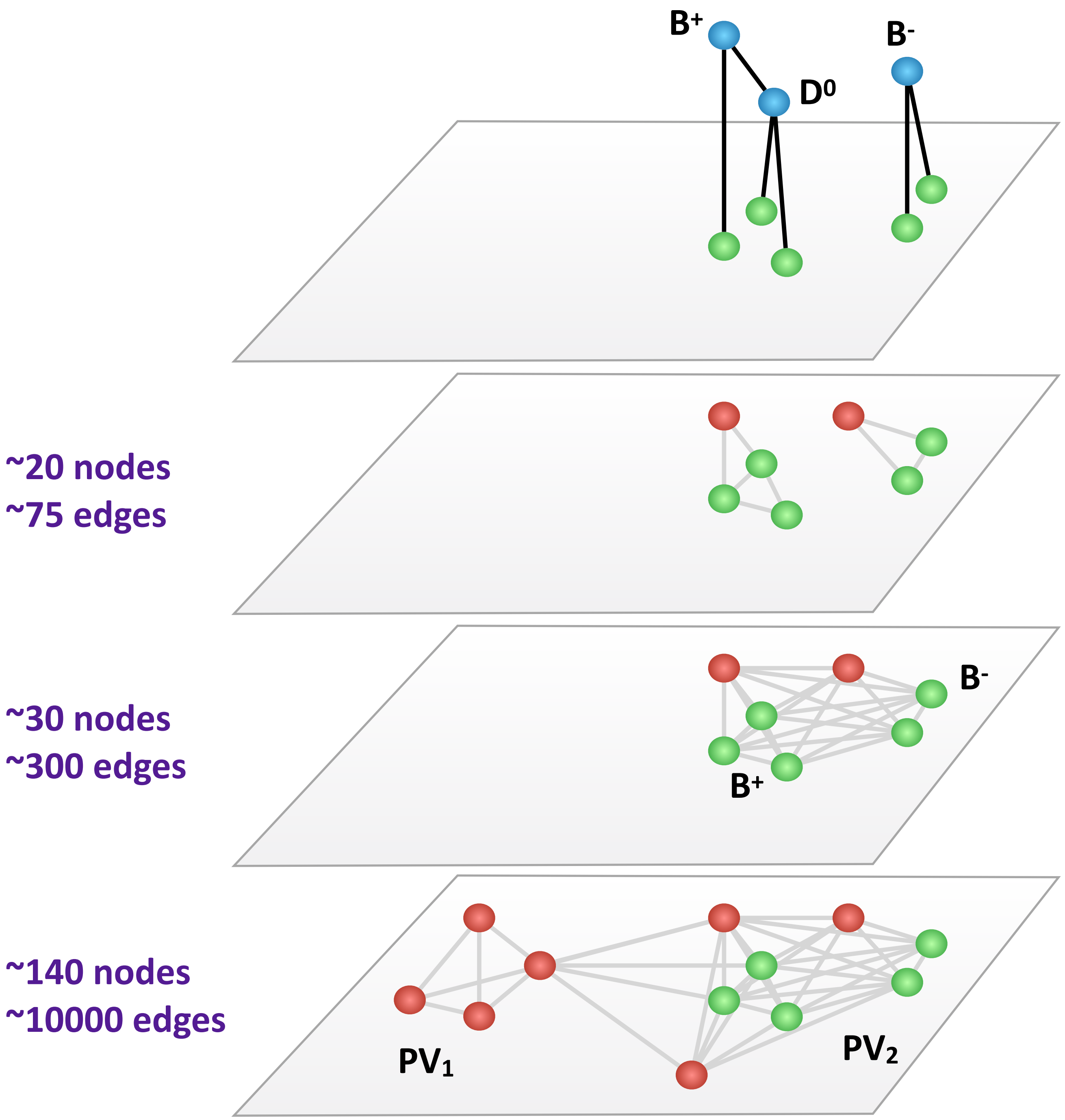
~30 nodes
~300 edges

~140 nodes
~10000 edges



Training

Dataset:
- PYTHIA-based simulation, Run 3 conditions, approximated emulation of LHCb reconstruction.
- Events required to contain at least one b-hadron (inclusive decay).



Train the LCA inference module
(~3.1x10^6 events needed)

Filter the data

Choose a threshold
(~95% signal eff.)

Train the edge-pruning module
(~1.3x10^4 events needed)

Filter the data

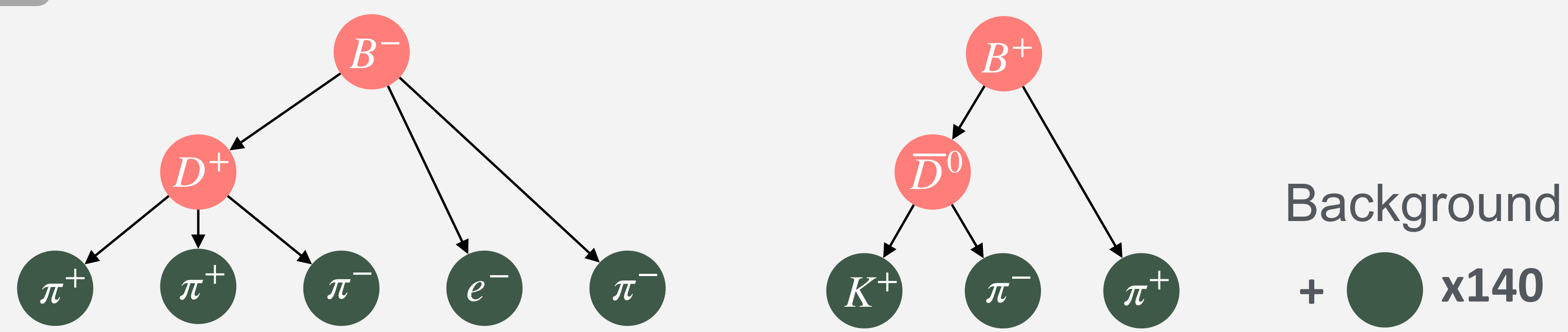
Choose a threshold
(~95% signal eff.)

Train the node-pruning module
(~1.3x10^4 events needed)

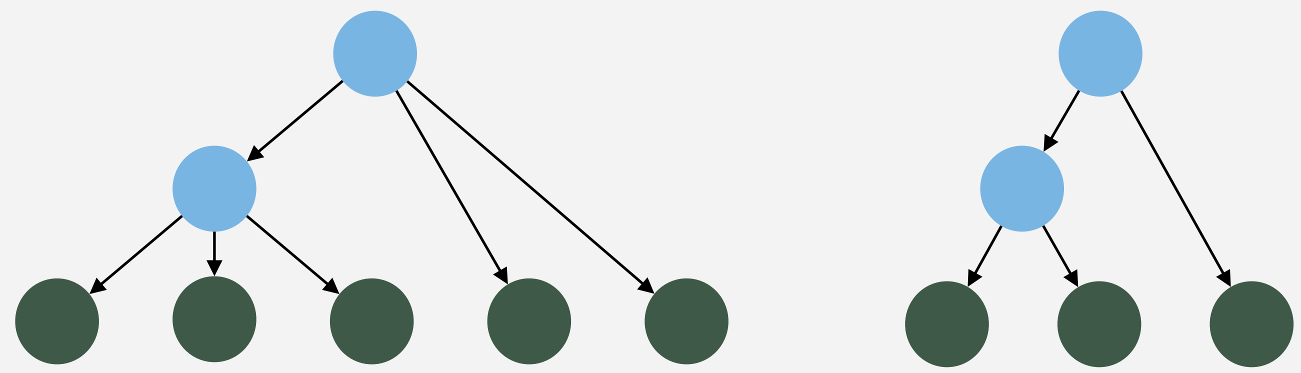
Preliminary performance: single chosen event

Example of a perfectly reconstructed simulation event.

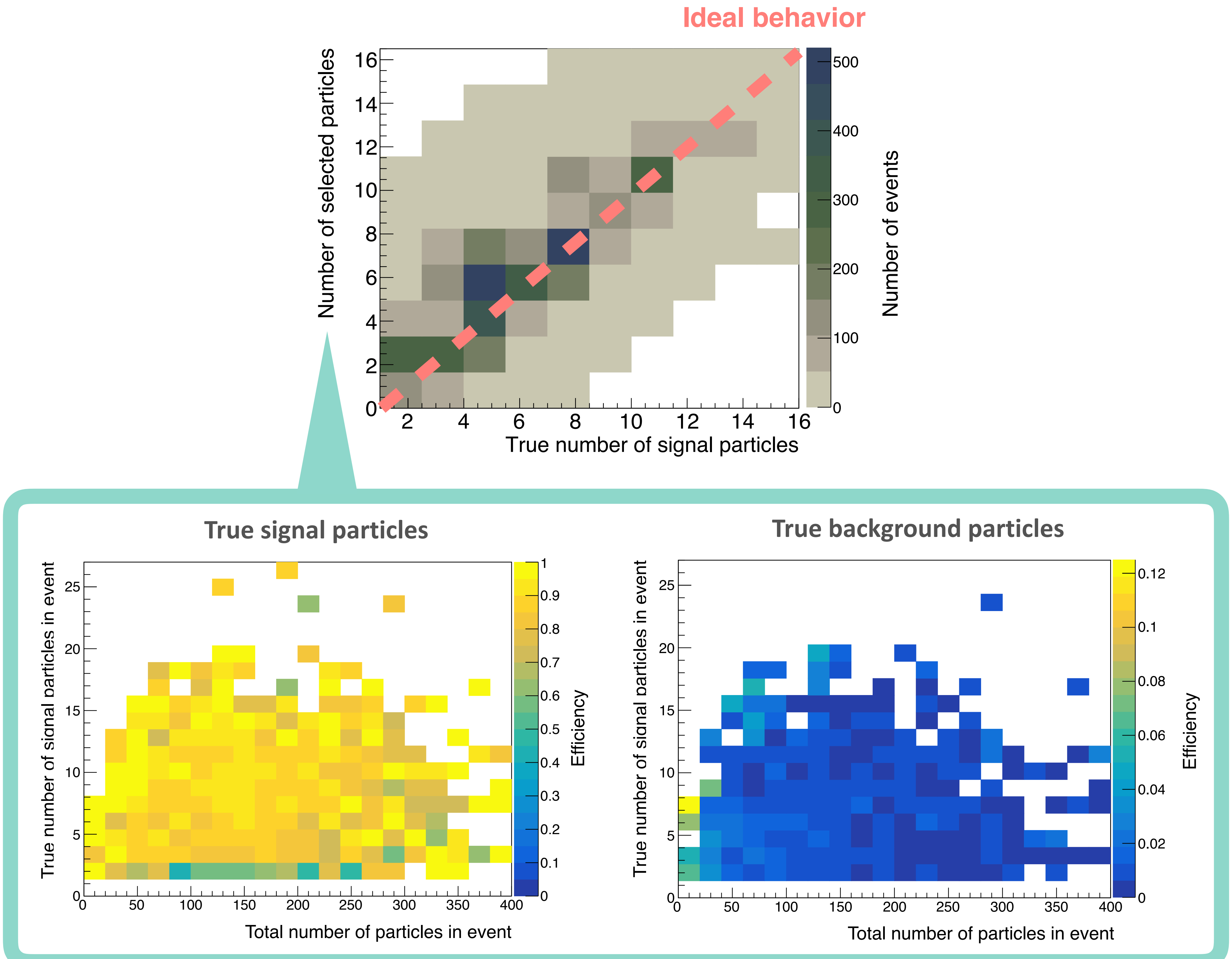
Simulated event



DFEI output



Preliminary performance: average over events



Summary and outlook

Unprecedented computational challenges for the future **Upgrade II of LHCb**.
Paradigm-change needed: from **“which events are interesting?”** to **“which parts of the event are interesting?”**.

As a solution, we propose a novel approach: change from the signal-based trigger strategy to a **Deep-learning based Full Event Interpretation**.

- ➡ Automatic and accurate identification and reconstruction of all the heavy-hadron decay chains per event.
- ➡ Allows to discard the rest of the event, with minimal loss for offline analyses.

We have developed the first prototype of the DFEI algorithm, focused on b-hadron decays and charged stable particles.

- ➡ **Very promising performance in realistic conditions!**

Future plans:

- ➡ Optimisation, expansion of functionality.
- ➡ Tests on Run 3 data.
- ➡ Full deployment in the Upgrade II.

Backup slides

Training dataset: emulating Run3 conditions

Particle collision&decay

The training and performance studies are currently done using **PYTHIA**, with the following configuration:

- Proton-proton collisions at 13 TeV.
- Average number of collisions per event: 7.6.
- Selecting **events with at least one b-hadron produced (inclusive decay)**.

"Detection and reconstruction"

We require all the tracks and the b-hadrons to be **inside the LHCb geometrical acceptance**.

In addition, we **emulate the reconstruction of the following quantities**, using publicly available expectations for the LHCb performance in Run3 (see backup):

- **Origin point of the tracks** (first measurement in the Vertex Locator).
- **Three-momentum of the tracks**.
- **Position of the primary vertices**.

Further bibliography

The LHCb Upgrades for Run3 and Run4: https://indico.cern.ch/event/868940/contributions/3813743/attachments/2081057/3495477/200725_ICHEP_LHCbUpgrades_v3.pdf

Performance estimates for Run3 conditions, used in our private simulation:

- Smearing of the true PV positions: https://indico.cern.ch/event/831165/contributions/3717129/attachments/2022791/3382986/ctd_2020_freiss.pdf
- Smearing of different reconstructed quantities: https://twiki.cern.ch/twiki/bin/view/LHCb/ConferencePlots and Computer Physics Communications 265, 108026 (2021).
- Geometry of the Vertex Locator: https://cds.cern.ch/record/2147229/files/10.1016_j.nima.2016.04.077.pdf

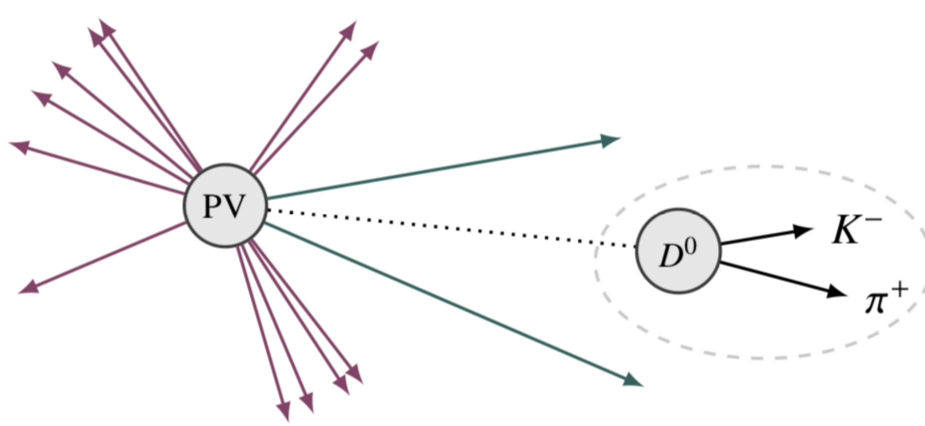
Signal-based trigger vs Full Event Interpretation (FEI)

Signal based

The current LHCb trigger is an **OR between many decay-mode selection lines**.

Since Run2, to reduce the event size, some lines **store only parts of the event which are related** to the specific signal. [JINST 14 (2019) 04, P04006]

E.g.: store the signal + the tracks in the same primary vertex (PV).



FEI

New proposal: try to **reconstruct the b- and c- hadron decay chains in the event**, in a hierarchical-clustering manner (cluster → unstable particle), **and discard the rest**.

Advantages:

- **Exploit extra correlations** between objects in the event.
- **Bandwidth oriented**: focus on storing as much "useful" information as possible.
- Case of several signals per event as an integral part of the approach.
- Establishment of a basis for an expanded functionality of the trigger: inclusive selections, study of anomalous events …

Differences between Belle II and LHCb

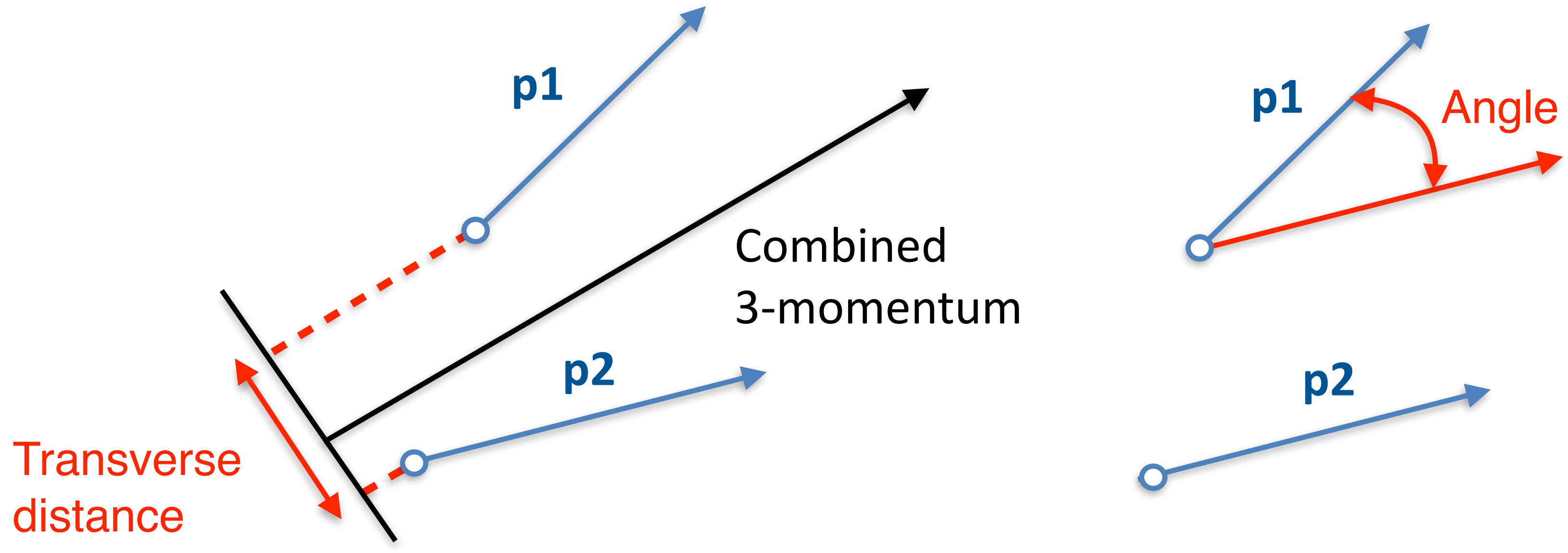
Belle II		LHCb
Only $B^0/B^\pm$ hadrons	→	All b-hadron species (+ c-hadrons)
List of possible decays large but limited and well identified.	→	Enormous list of possible decay chains.
e+e- collisions: clean environment.	→	pp collisions: important contamination from the rest of the event.
Hermetic detector: all the information (except for neutrinos) available.	→	Non-hermetic detector: partially-reconstructible decays in many cases.

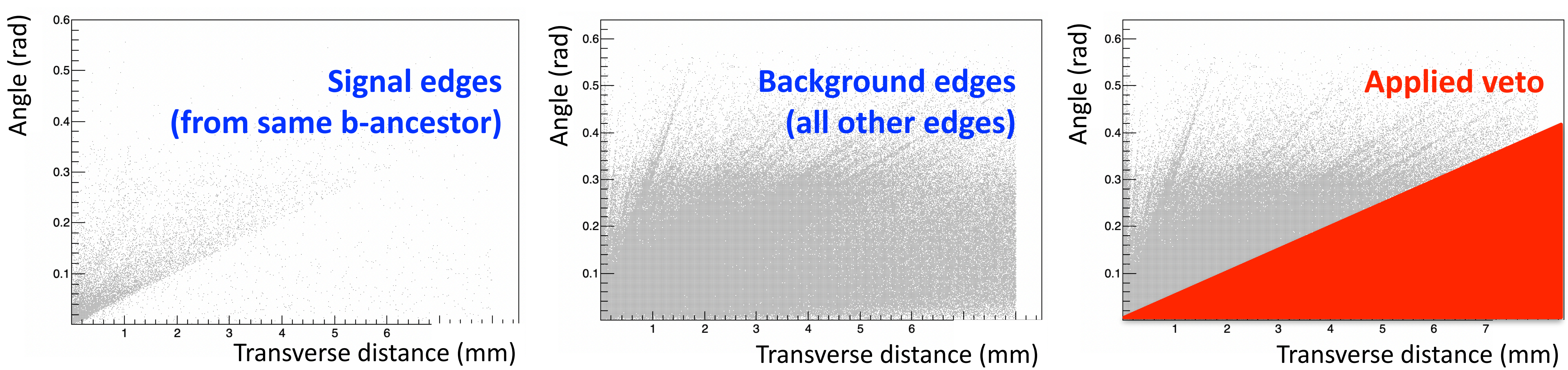
⇓

Leverage the full power of **deep-learning** and try to reconstruct as much as possible.
➥ **DFEI: Deep-learning based FEI for LHCb.**

Cut-based edge pruning

Define two adequate topological variables for each edge (pair of particles)





This veto reduces on average 60% of the total number of edges in the graph.
It also reduces connections between signal tracks, but it only leaves ~2% of the signal tracks fully disconnected.

Example of decay-tree simplification used in the prototype

Original chain of ancestors:

$\pi^+ \leftarrow \rho(770)^0 \leftarrow \phi(1020) \leftarrow D^+ \leftarrow B^0 \leftarrow B^{*0}$

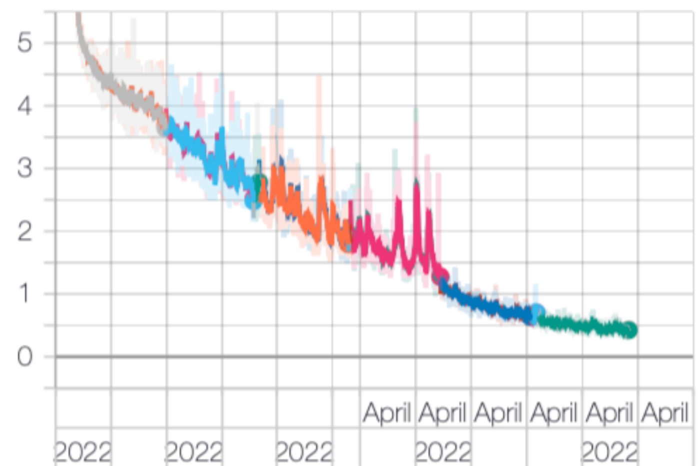
Simplified chain of ancestors (based on reconstructible vertices):

$\pi^+ \leftarrow D^+ \leftarrow B^0$

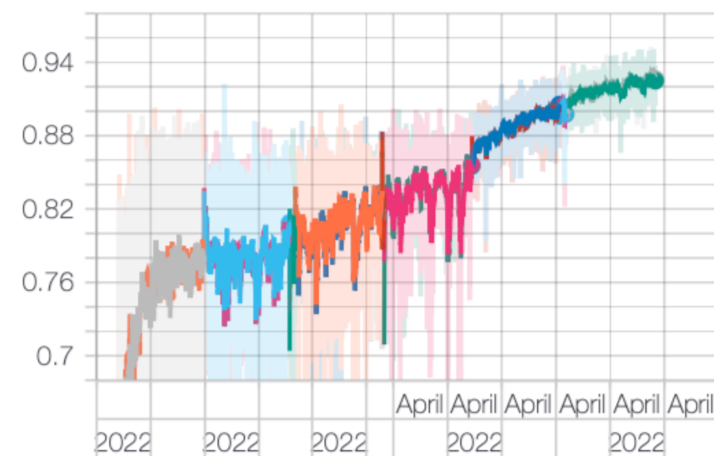
Training of the LCA reconstruction

Training split in 6 steps, each of them doing 4000 iterations in batches of 128 events.

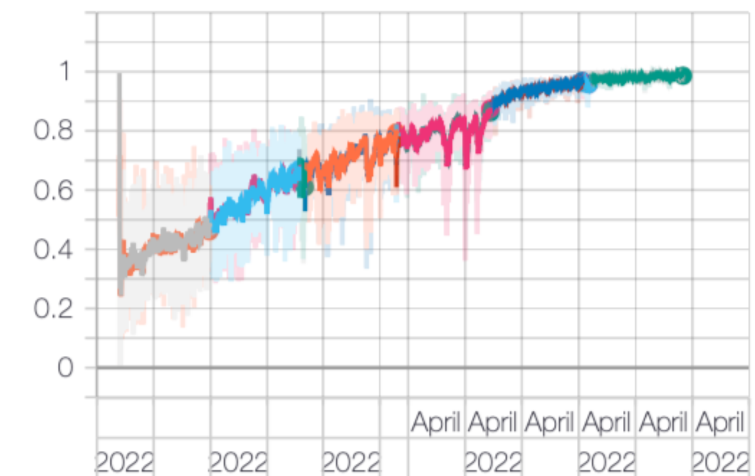
Loss



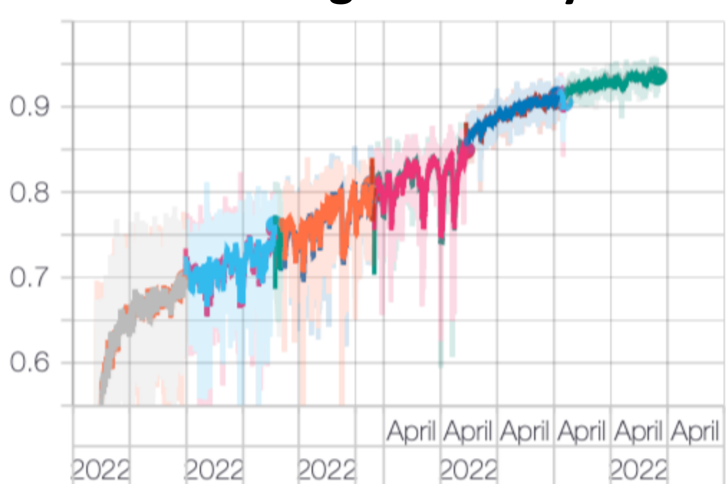
Accuracy (true LCA = 0)



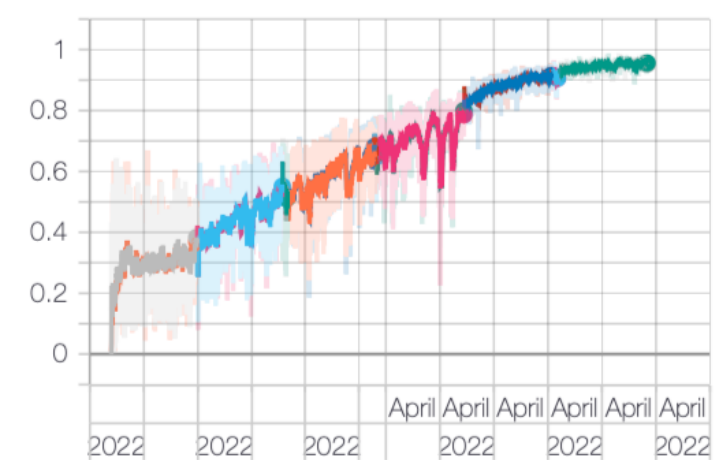
Accuracy (true LCA = 1)



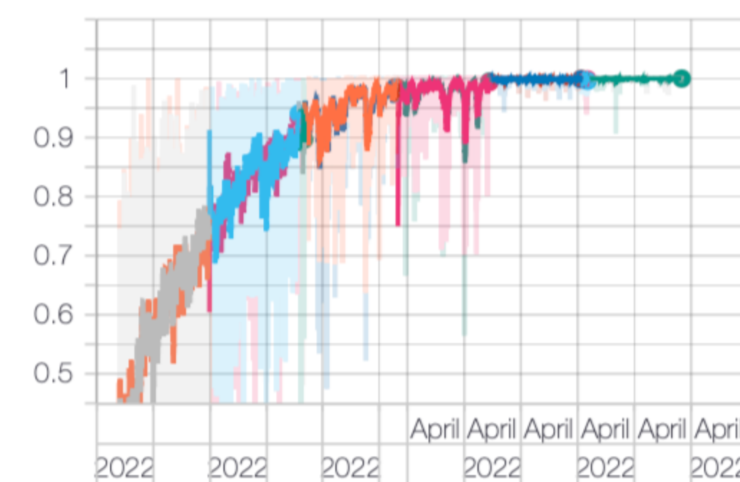
Average accuracy



Accuracy (true LCA = 2)



Accuracy (true LCA = 3)

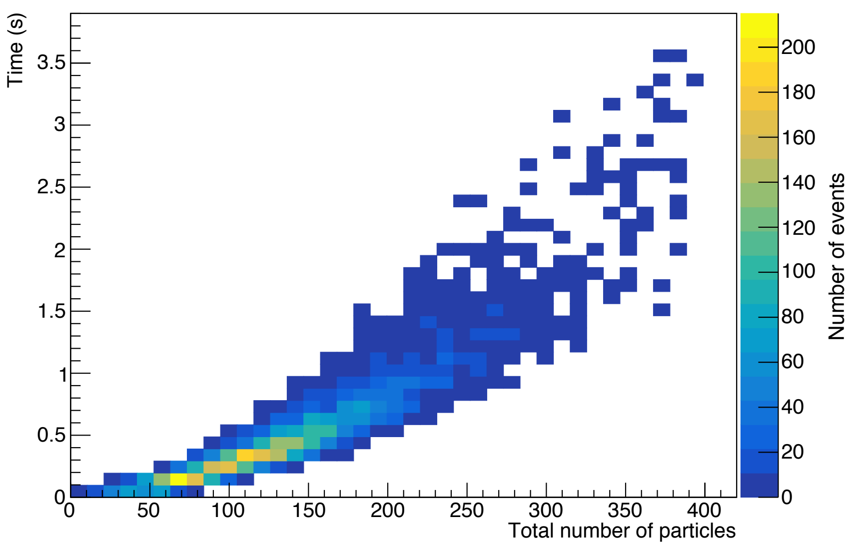


No signs of overtraining (training and test curves always ~overlapping).

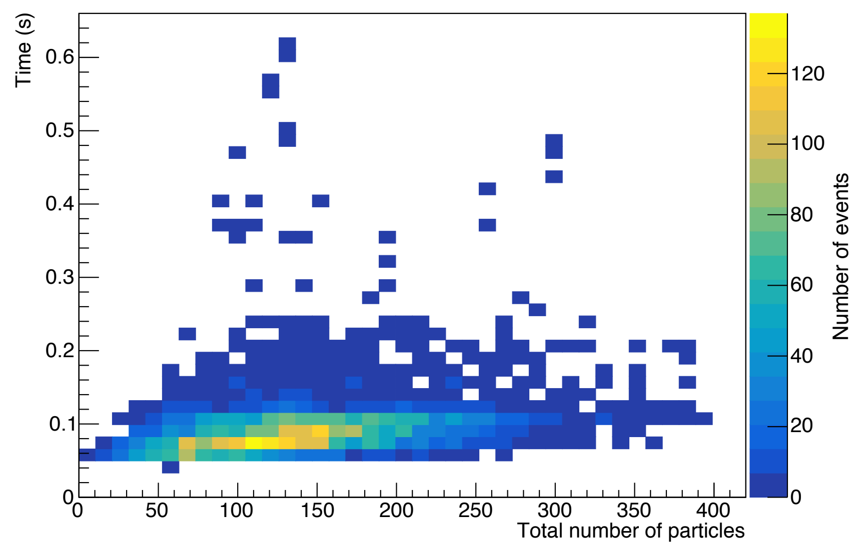
Performance: timing

Simplistic study (no parallelisation, no hardware accelerators*, algorithm to be further optimised), to **understand which are the slowest parts of the algorithm and how they scale with the total number of particles per event.**

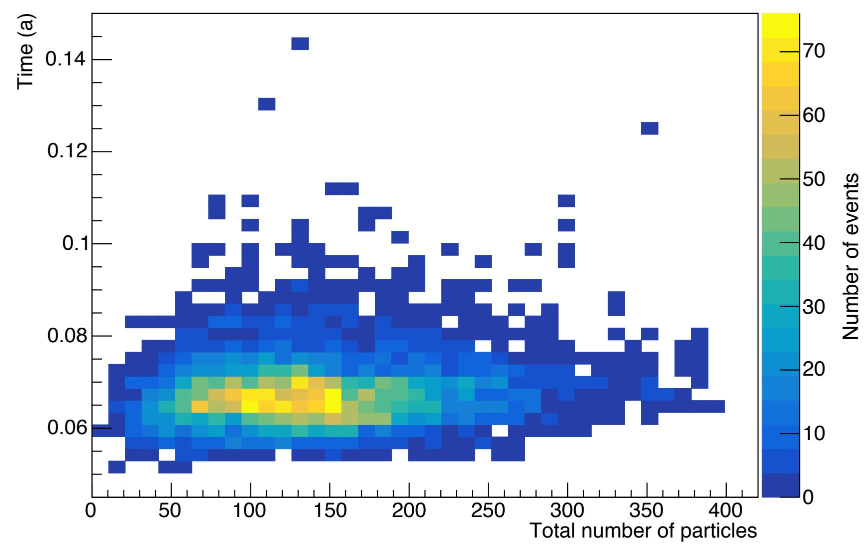
Node pruning



Edge pruning



LCA reconstruction



The slowest part is the node pruning, which also has the strongest dependency on the number of particles. → Many possible ways of optimisation.

The processing time of the subsequent algorithms is quite stable regarding changes in event complexity.

(*) Study done on a darwin-x86_64 architecture with a 2.8 GHz Intel Core i7 processor.