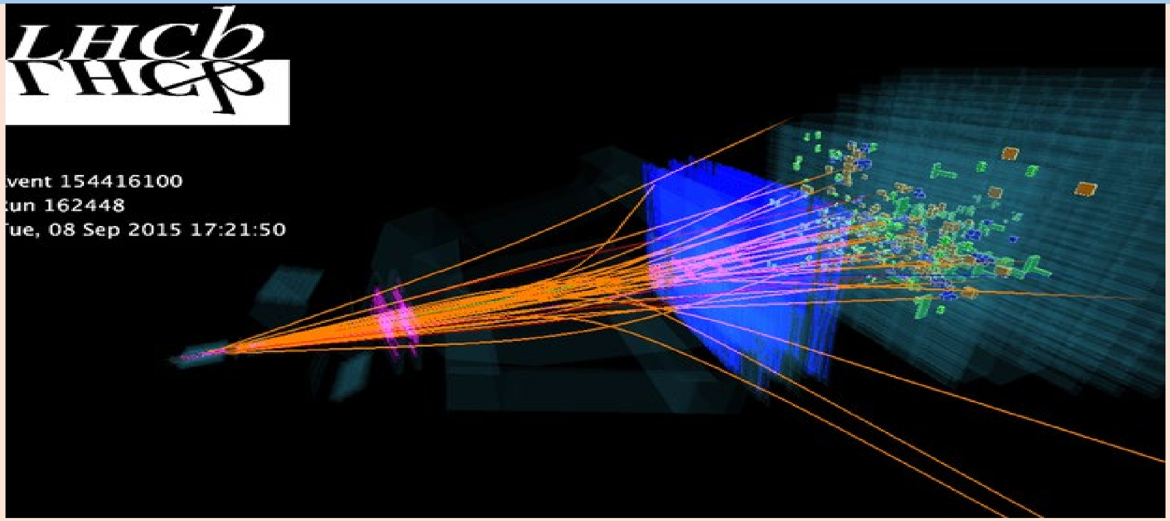
New developments and applications of a Deep-learning-based Full Event Interpretation (DFEI) in proton-proton collisions



Marta Calvi[1], Simone Capelli[1], Jonas Eschle[2], Julian Garcia Pardinas[1,3], Abhijit Mathad[3], Andrea Mauri[4], Nicola Serra[5], Rafael Silva Coutinho[2], **Felipe Luan Souza De Almeida[2]**, William Sutcliffe[5], Azusa Uzuki[5]

Univ. degli Study Milano-Bicocca[1], Syracuse University[2], CERN[3], Imperial College[4], University of Zurich[5]

22nd International Workshop on Advanced Computing and Analysis Techniques in Physics Research

Outline (I)

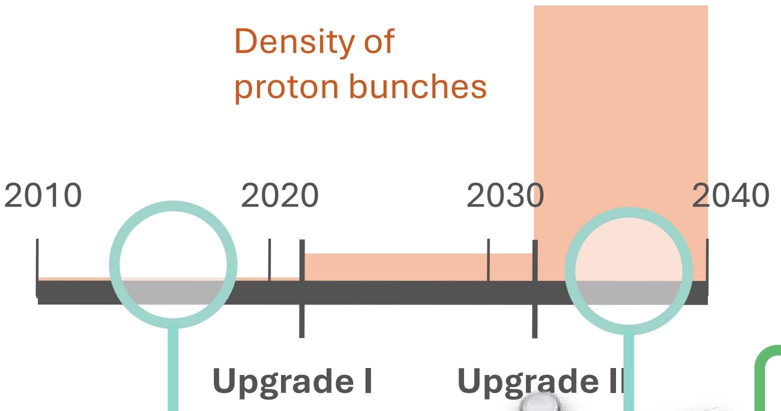
- Introduction: DFEI
- Performance improvements
- Applications
- Summary

Introduction

Increased particle multiplicities for LHCb Upgrades I and II bring big challenges, both for trigger and offline analysis

Which events are interesting?

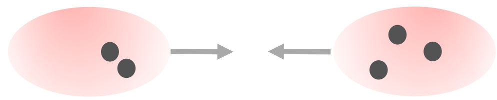
"Signal-centric" trigger strategy



Which *parts* of the event are interesting?

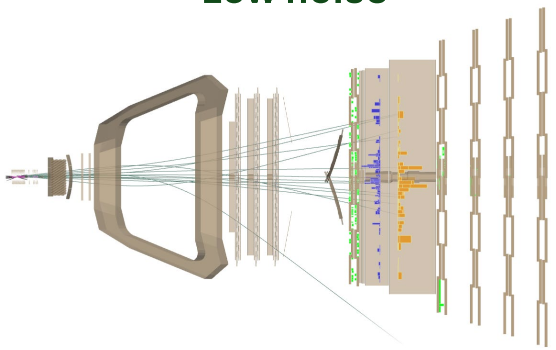
⇩

New trigger paradigm

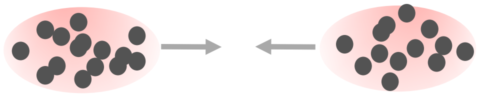


≪ 1 signal per event, ~50 tracks

Low noise

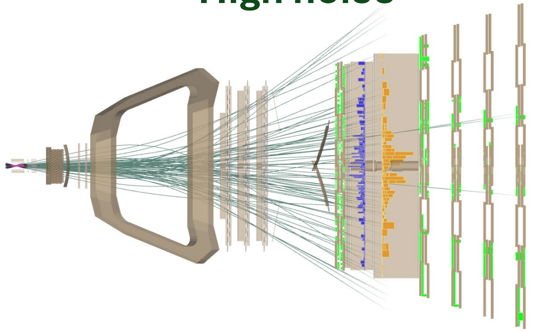


Need for ML

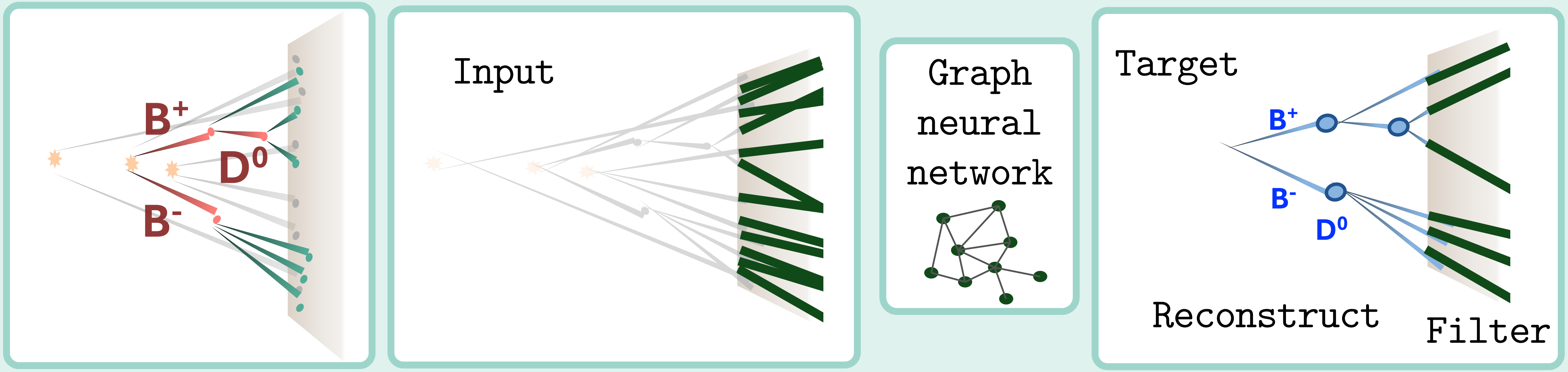


Multiple signals per event, ~1000 tracks

High noise



Deep-learning based Full Event Interpretation (DFEI)



- **Input features**: charged particles and its measured properties (nodes) and their relations (edges)
- Hierarchical, automatized and inclusive reconstruction of heavy-hadron decay chains
- **Trigger**: Safely discard rest of event, with minimal loss for analyses → powerful event size reduction in a multi-signal environment
- **Analysis**: tool for background classification & suppression

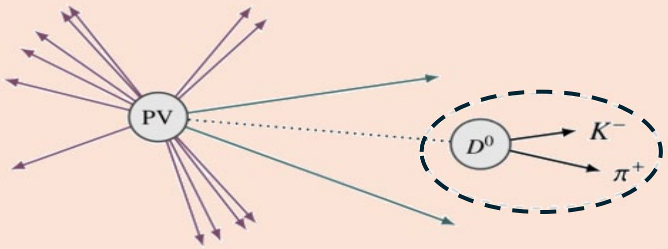
SIGNAL-BASED TRIGGER VS FEI

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 that 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

Reconstruct **the b- and c- hadron decay chains in the event**, in a hierarchical-clustering manner (cluster → unstable particle), **and discard the rest**.

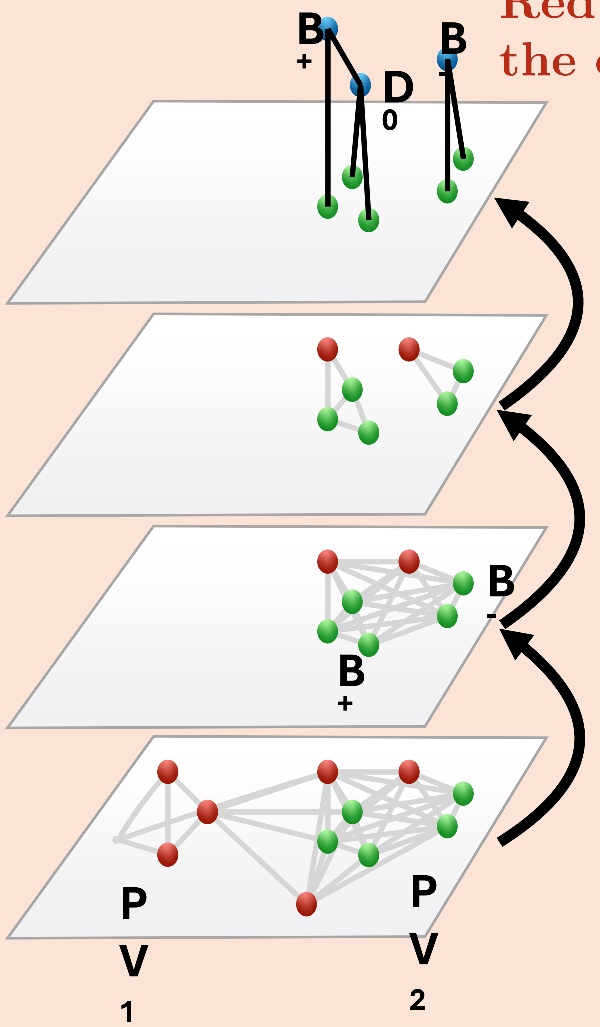
Advantages:

- **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 …

DFEI

[Comput.Softw.Big Sci. 7 (2023) 1, 12]

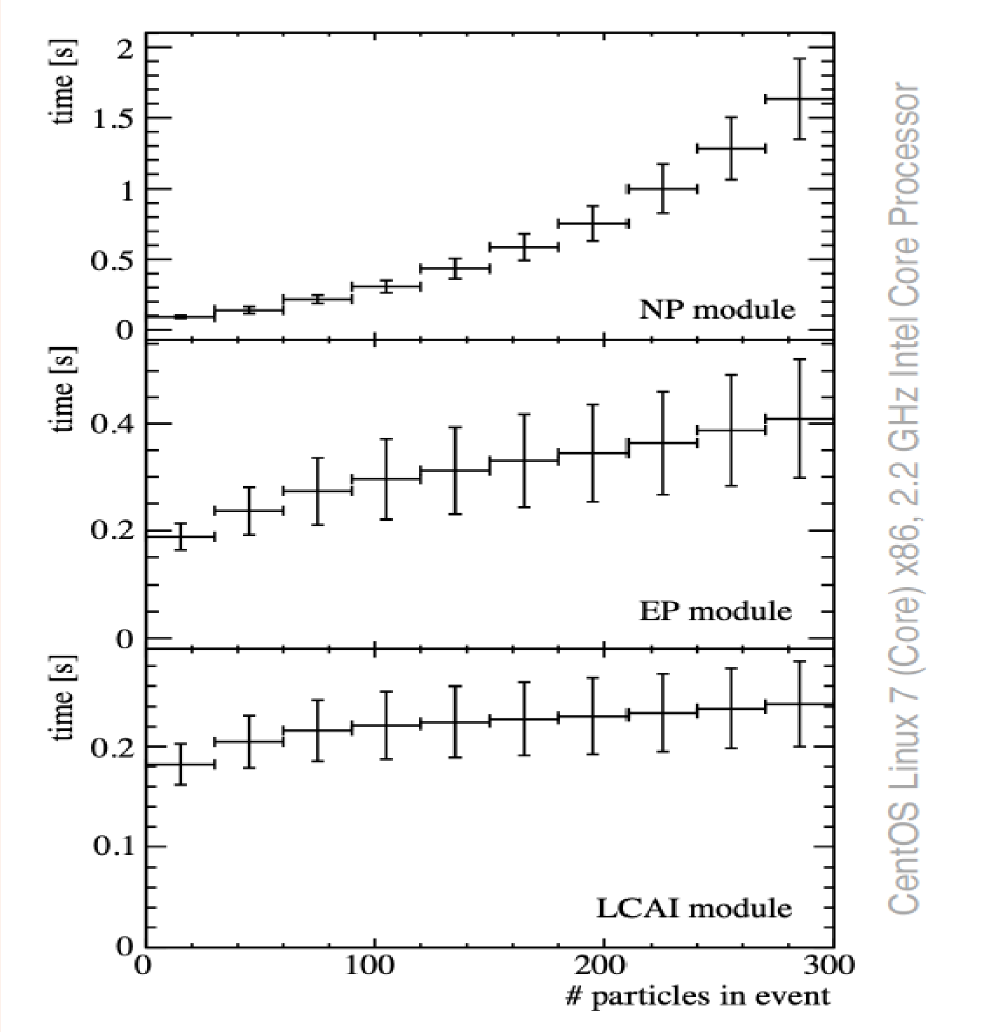
Blue: reconstructed ancestors
Green: particles from a b-hadron
Red: particles from the rest of the event



3. “Lowest common ancestor” inference

2. Edge pruning

1. Node pruning

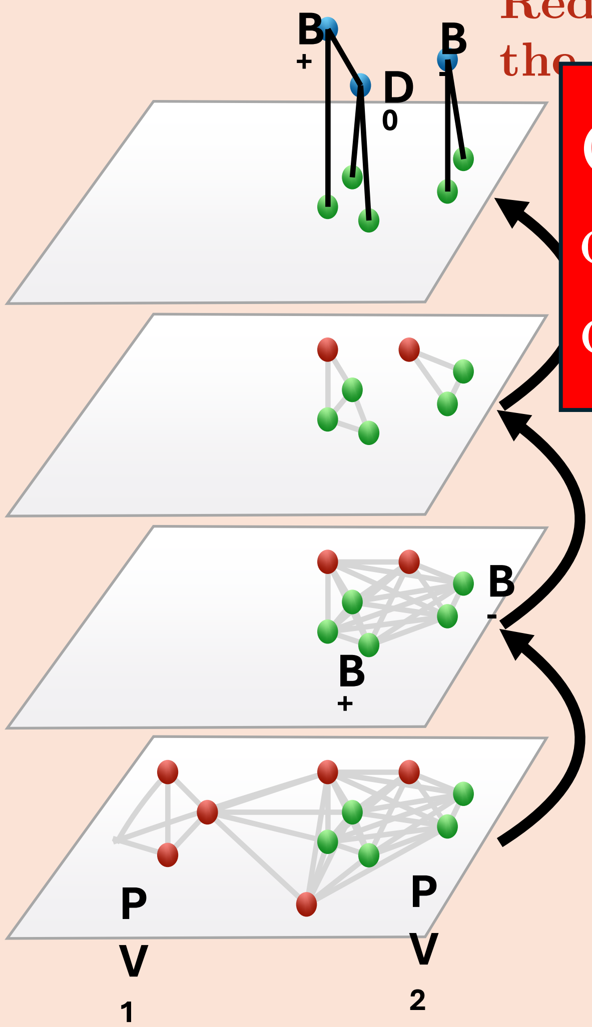


+ about 2 s per event for pre/postprocessing of the multiple modules

DFEI

[Comput.Softw.Big Sci. 7 (2023) 1, 12]

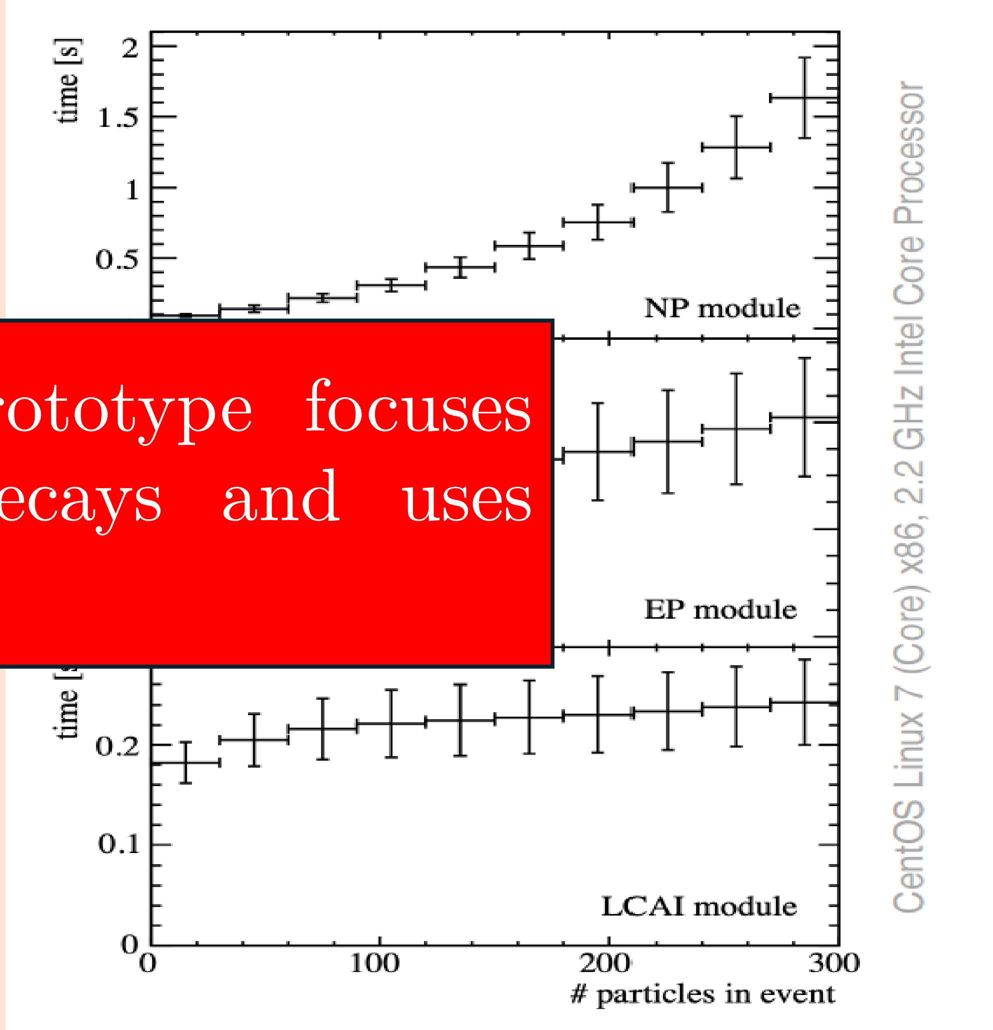
Blue: reconstructed ancestors
Green: particles from a b-hadron
Red: particles from the rest of the event



2. Edge pruning

1. Node pruning

Our current DFEI prototype focuses only on b-hadrons decays and uses charged stable particles



+ about 2 s per event for pre/postprocessing of the multiple modules

Outline (II)

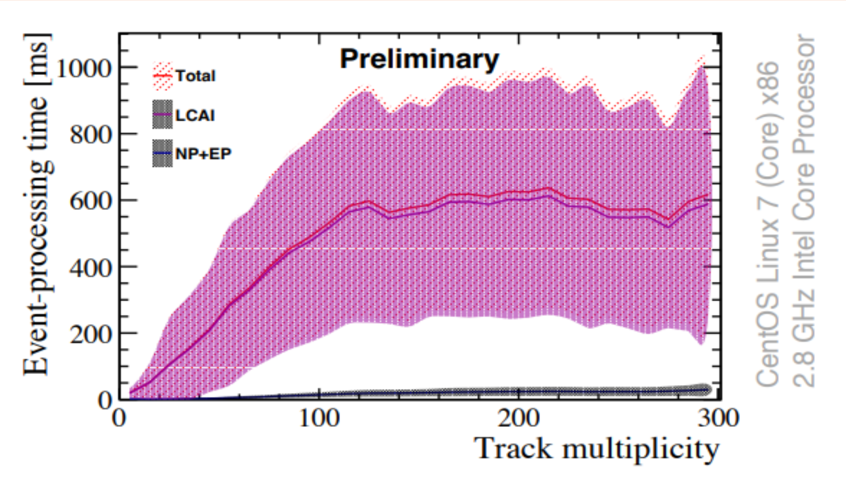
- Introduction: DFEI
- Performance improvements
- Applications
- Summary

First speed-up round and C++ inference pipeline

- First DFEI prototype: quadratic scaling of the inference time with the track multiplicity, overall evaluation time on the order of few seconds per event on CPU. Evaluation pipeline on python with TensorFlow.
- To improve the scaling with the track multiplicity per event, substitute the GNN models of the NP and EP by CatBoost BDT models:
  - Evaluated independently for each particle/edge, using information from kinematics, topology, PV association, and summary of particle multiplicity in the event.
  - Cut at 99% signal efficiency for nodes/edges.
- Construct a C++ inference pipeline that takes as input the information of the set of all charged stable particles per event.
  - Using the C API of CatBoost and TMVA::SOFIE for the LCAI GNN inference

First speed-up round and C++ inference pipeline

- First DFEI prototype: quadratic scaling of the inference time with the track multiplicity, overall evaluation time on the order of few seconds per event on CPU. Evaluation pipeline on python with TensorFlow.
- To improve the scaling with the track multiplicity per event, substitute the GNN models of the NP and EP by CatBoost BDT models:
  - Evaluated independently for each particle/edge, using information from kinematics, topology, PV association, and summary of particle multiplicity in the event.
  - Cut at 99% signal efficiency for nodes/edges.
- Construct a C++ inference pipeline that takes as input the information of the set of all charged stable particles per event.
  - Using the C API of CatBoost and TMVA::SOFIE for the LCAI GNN inference



- Time dominated by the LCAI algorithm. **Very good scaling achieved thanks to the fast pre-filtering by the NP and EP.** Significant overall speed up, but final number pending on a hyper-parameter tuning of the LCAI.
- Next: physics performance studies for the new configuration, hyper-parameter tuning.
- In parallel: study of GNN architecture developments to gain further speed-ups.

Outline (III)

- Introduction: DFEI
- Performance improvements
- Applications
- Summary

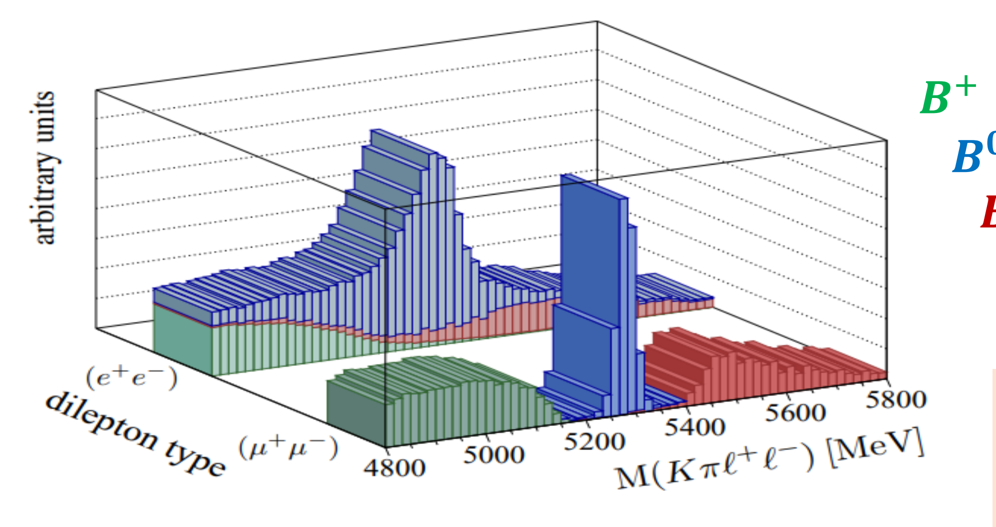
Standard vs DFEI-based analysis

- Standard exclusive analysis: identification of signal decay → fine-tuned selection
  - Possible contamination from similar partially-, over-reconstructed or poorly identified decays

- In DFEI, the algorithm is trained to inclusively identify and reconstruct decay chains originating from a b-hadron → independently of the type of the decay chain or of the number of particles in the final state
  - Partially- and over-reconstructed decays in standard analysis are fully reconstructed in DFEI
  - These contributions are classified in reconstructed clusters of different numbers of particles
  - DFEI seems to mitigate partially- and over-reconstructed background

Example case: $B^0 \to K^+\pi^-\ell^+\ell^-$

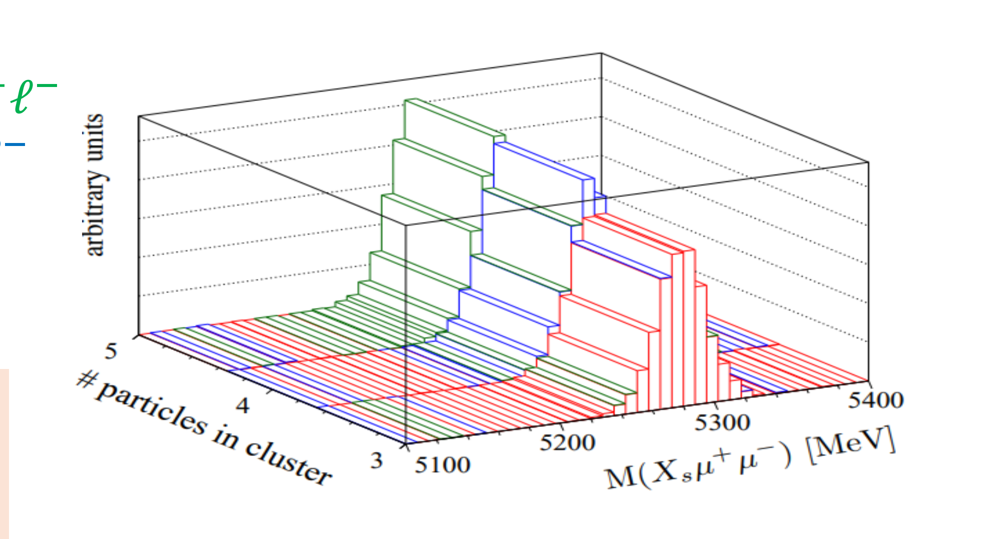
- Partially- and over-reconstructed backgrounds are challenging to disentangle from the signal in exclusive, conventional analysis → loose of sensitivity
  - Particularly important for electron modes, impacting R_X measurements
- With DFEI, these contributions are fully reconstructed and classified in clusters of different numbers of particles



$B^+ \to K^+\pi^-\pi^+\ell^+\ell^-$

$B^0 \to K^+\pi^-\ell^+\ell^-$

$B^+ \to K^+\ell^+\ell^-$



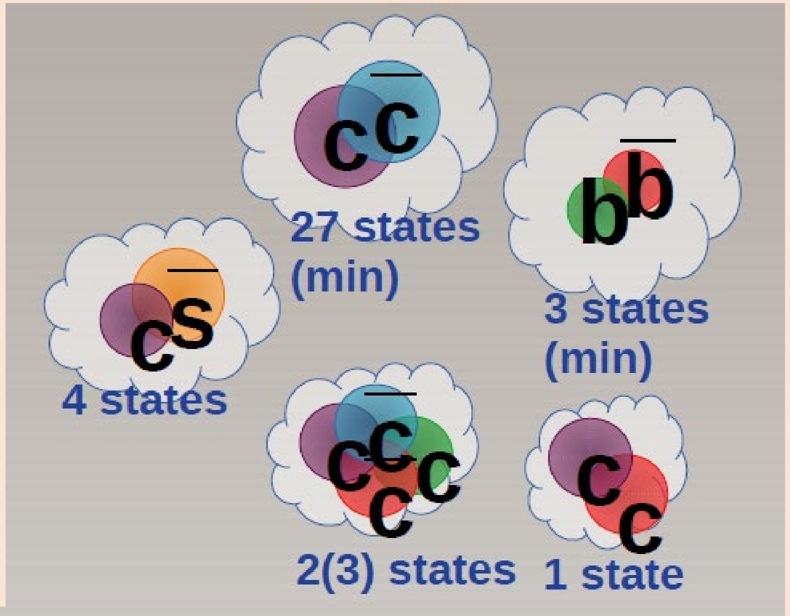
Conventional, exclusive analysis

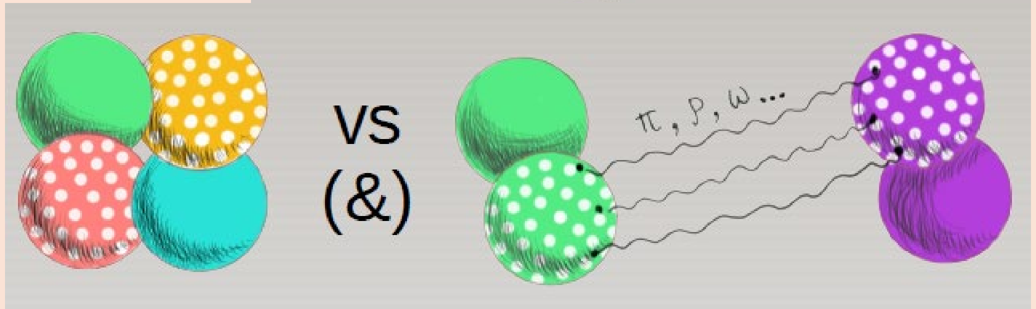
With DFEI inclusive reconstruction

Search for exotic states

- Several states observed (>35 since 2003)
- But: we still don't understand their nature
  - Bound or molecular states?
- Need unambiguous experimental evidence
- Other doubly-heavy states $[QQ\overline{ud}]$:
  - T_{bb} $[bb][\overline{ud}]$ → $\sim 10^{-3}$ events in Run3&4
  - T_{bc} $[bc][\overline{ud}]$ → may be below $\bar{B}D$ threshold, but opposite expectations in some molecular models

Karliner, Rosner, 2017, Semay, SIlvestre-Brac, 1994, Carames, Vijande, Valcarce, 2019, Meng et al., 2021
Li, Sun, Liu, Zhu, 2012, Liu et al., 2019, Hudspith et al., 2020

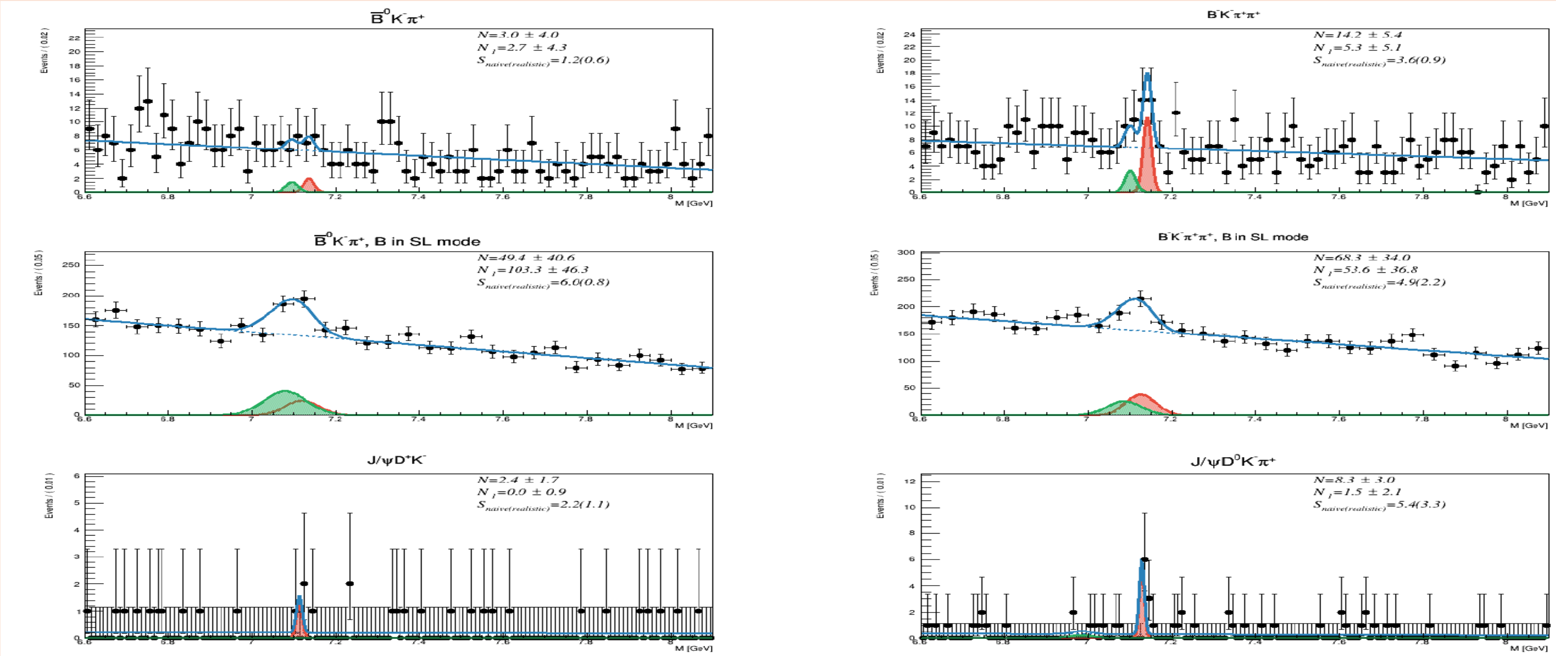




Search for Tbc – prospects for Run3, I. Polyakov,
Hunting for the charming beauty tetraquark Tbc: LHCb meets theory, 5 October 2023, CERN

Search for exotic states

- Current ideas for searching for exotic states include simultaneous analysis of multiple exclusive decay chains
  - **E.g.: the search for the tetraquark T_{bc} could involve the simultaneous mass fit of 20 to 40 channels!**



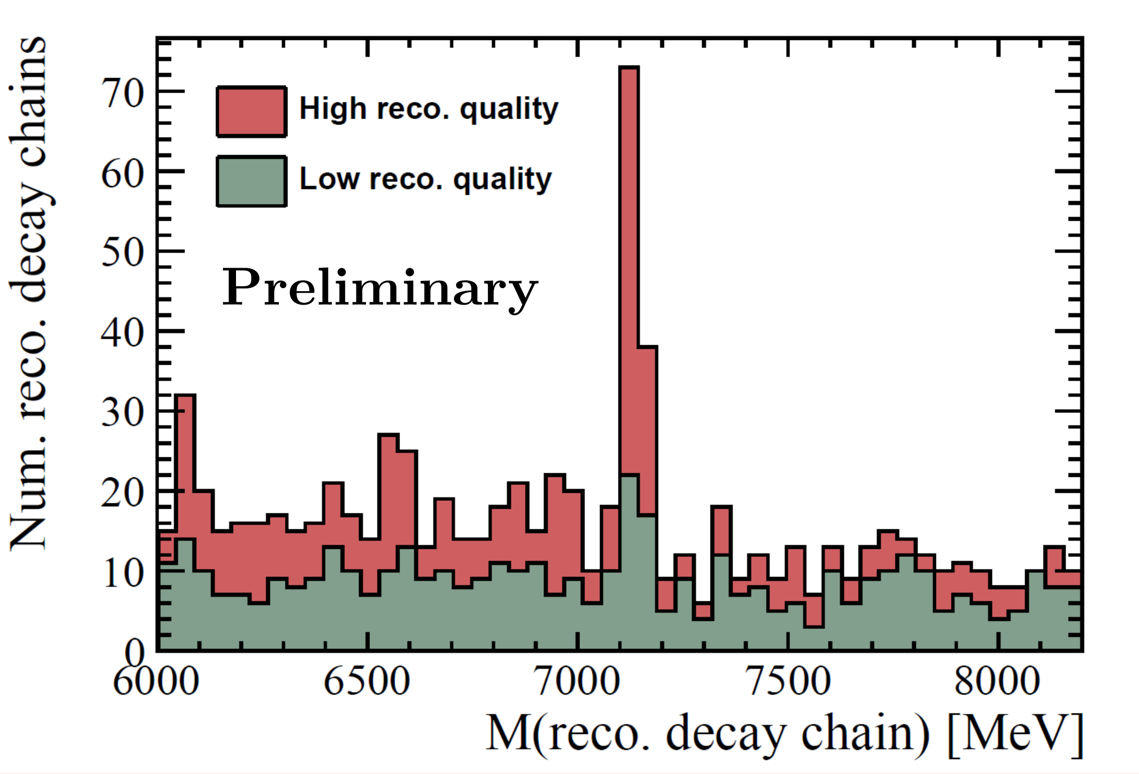
Search for Tbc – prospects for Run3, I. Polyakov,
Hunting for the charming beauty tetraquark Tbc: LHCb meets theory, 5 October 2023, CERN

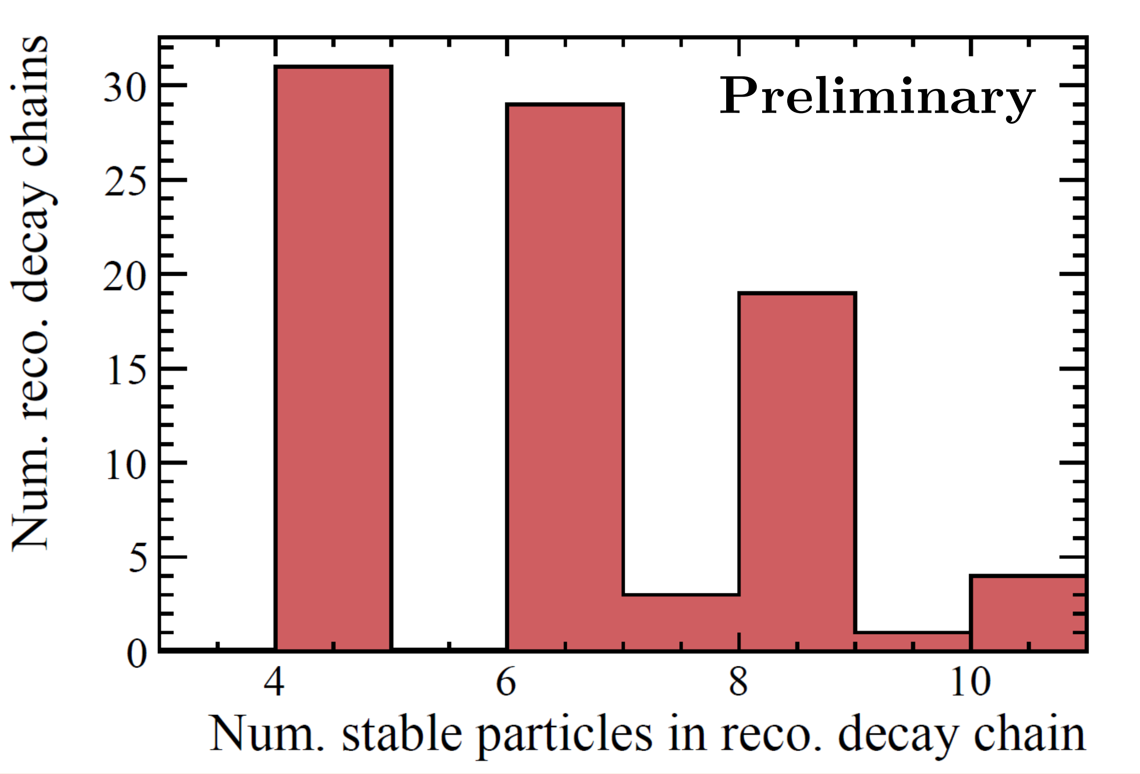
decay channel	$\varepsilon_{tot} \times \mathcal{B}$ $[10^{-9}]$	Expected yield
fully reconstructed channels		
D^0D^0	7.8	3.1
$D^0D^+\pi^-$	9.2	3.7
$D^0D^0\pi^+\pi^-$	3.4	1.4
$D^+D^+\pi^-\pi^-$	3.5	1.4
sum		10
$J/\psi D^+K^-$	2.3	0.9
$J/\psi D^0K^-\pi^+$	3.1	1.2
sum		2.1
$\overline{B}^0K^-\pi^+$	32.9	13.2
$B^-K^-\pi^+\pi^+$	33.6	13.4
$\overline{B}^0K^-\pi^+\pi^+\pi^-$	6.7	2.7
sum		29
$\overline{B}^0{}_{SL}K^-\pi^+$	188	75
$B^-{}_{SL}K^-\pi^+\pi^+$	94	38
$\overline{B}^0{}_{SL}K^-\pi^+\pi^+\pi^-$	61	24
sum		137
$D^0D^+\mu^-\nu$	56	22
$D^0D^0\pi^+\mu^-\nu$	43	17
$D^0D^+\mu^- + X$	163	65
$D^0D^0\mu^- + X$	108	43
sum		147
$\overline{B}^0K^-\mu^+\nu$	24	9.5
$B^-K^-\pi^+\mu^+\nu$	16	6.5
sum		16
$D^0K^-\pi^+$	68	27
D^+K^-	134	54
$D^0\pi^+\pi^-$	2.5	1
$D^+\pi^-$	4.9	2
sum		84

Search for exotic states

DFEI can simultaneously reconstruct the different decay chains, allowing for a broader and more direct search for exotic states

Simulated sample of several weakly decaying modes of T_{bc} analyzed simultaneously using DFEI:

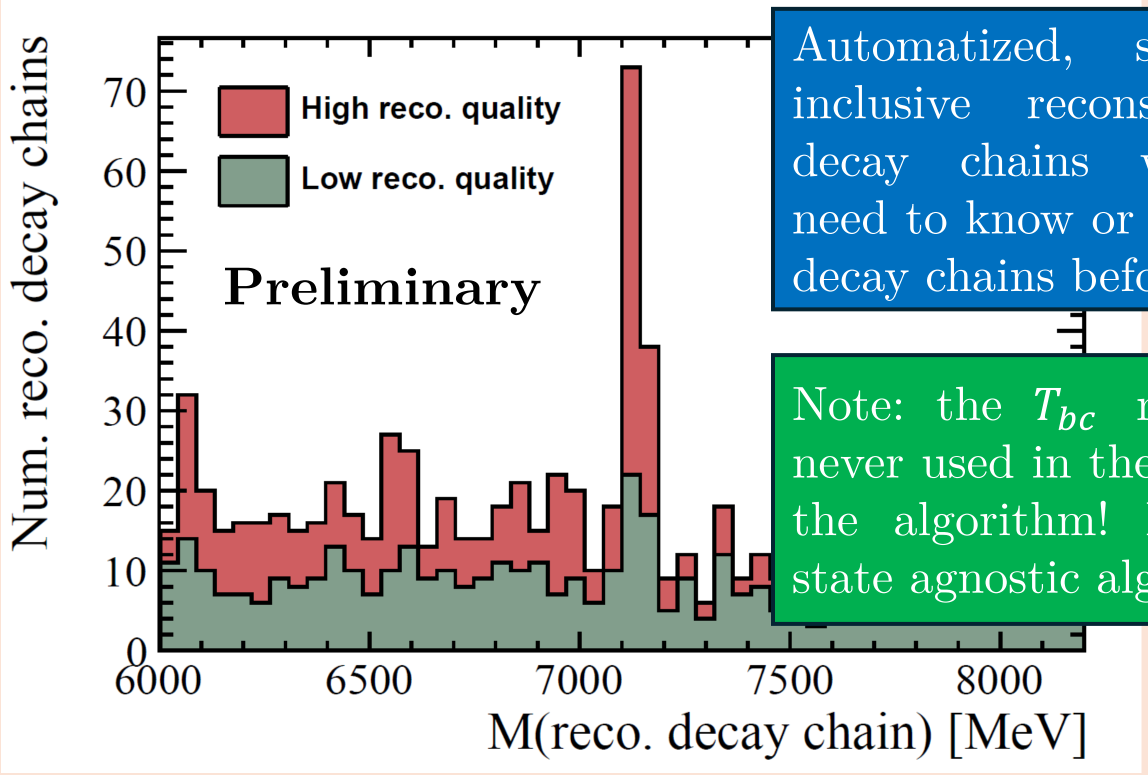




Search for exotic states

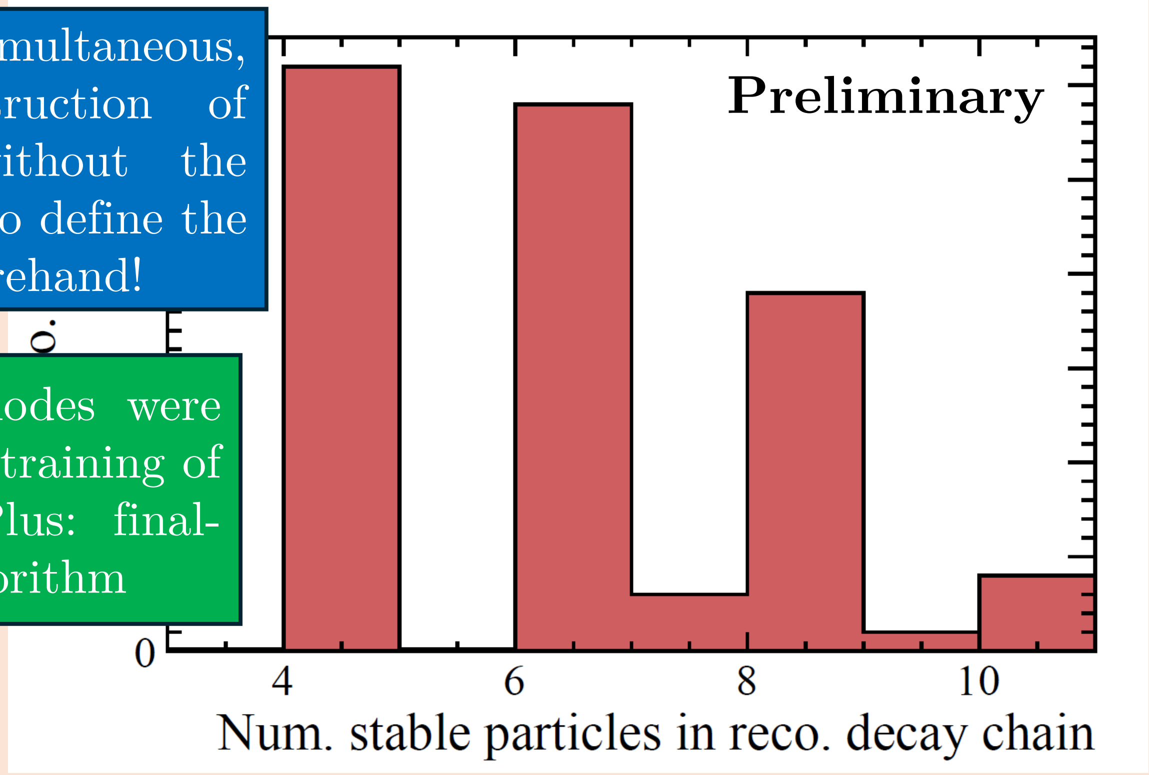
DFEI can simultaneously reconstruct the different decay chains, allowing for a broader and more direct search for exotic states

Simulated sample of several weakly decaying modes of T_{bc} analyzed simultaneously using DFEI:



Automatized, simultaneous, inclusive reconstruction of decay chains without the need to know or to define the decay chains beforehand!

Note: the T_{bc} modes were never used in the training of the algorithm! Plus: final-state agnostic algorithm



Outline (IV)

- Introduction: DFEI
- Applications
- Performance improvements
- Summary

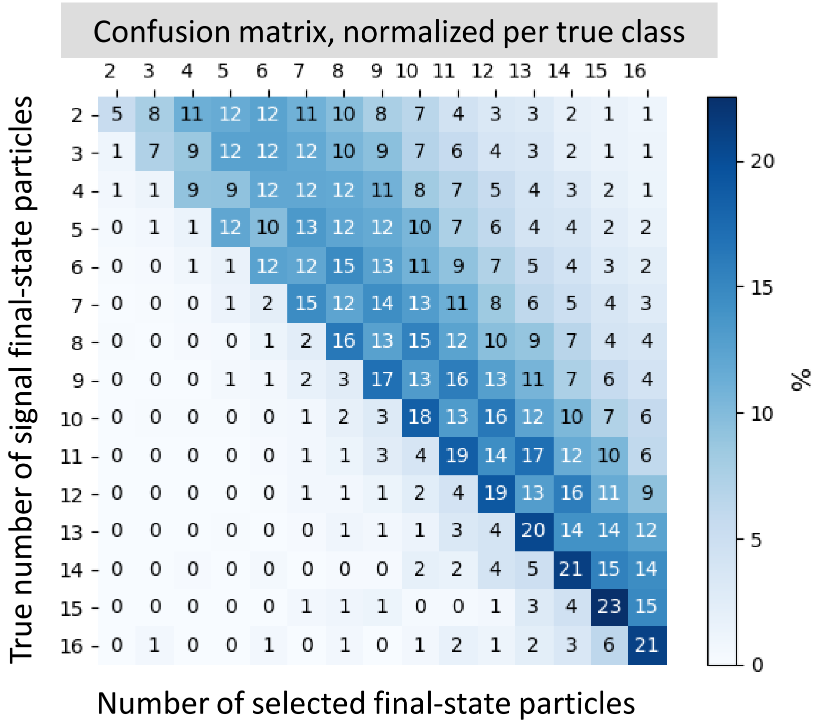
Summary

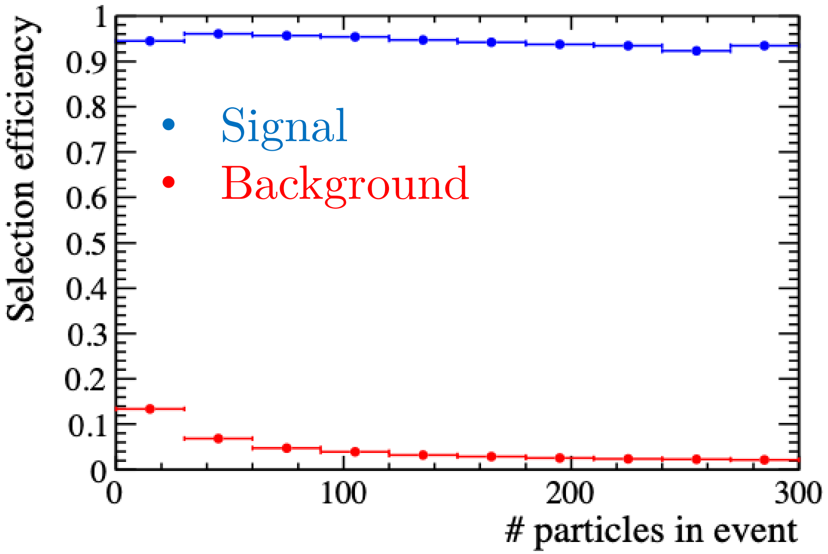
- Changing from GNN to BDT models for NP and EP and a new C++ inference pipeline resulted **in significant speed-up**
- **Timing scaling flat at high track multiplicity**
- Further speed-ups expected from hyper-parameter tuning of LCAI
- DFEI opens the door for **exciting physics applications** in hadronic machines
  - Seems to mitigate contamination from partially- and over-reconstructed decays
  - Can be used for inclusive searches for exotic states
  - And more
- Stay tuned: further developments are on the way

Thank you!

Backup

Performance: final-state particle filtering





Consistent performance with different number of signals

"single-b-hadron-signal" approach performance ***comparable*** to the envisaged nominal LHCb strategy for Run 3 [JINST 14 (2019) 04, P04006]

LHCb: 90% sig eff, 90% bkg rej. power

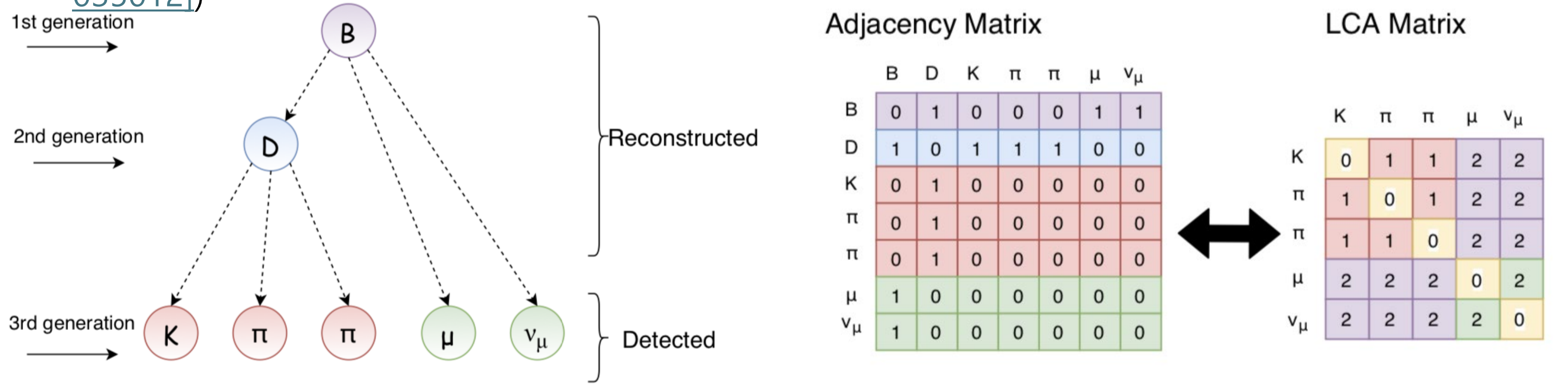
DFEI: 94% sig eff, 96% bkg rej. power

DFEI selects all of them simultaneously

DFEI capability #1
Powerful event size (~ x14) reduction in a multi-signal environment.

3rd module: Lowest Common Ancestor (LCA) inference

From [BELLE2-MTHESIS-2020-006]:
(see also [James Kahn et al 2022 Mach. Learn.: Sci. Technol. 3 035012])



	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

	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**.

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.

Decay-level performance

Decay mode	Perfect (%)	Wrong hierarchy (%)	Not iso. (%)	Part. reco. (%)
Inclusive H_b decay	4.6 ± 0.1	5.9 ± 0.1	76.0 ± 0.2	13.4 ± 0.1
$B^0 \to K_0^*[K\pi]\mu^+\mu^-$	35.8 ± 0.7	19.2 ± 0.6	44.9 ± 0.7	<0.02
$B^0 \to K^+\pi^-$	38.0 ± 0.7	–	54.7 ± 0.7	7.2 ± 0.4
$B_s^0 \to D_s^-[K^-K^+\pi^-]\ \pi^+$	32.8 ± 0.7	7.1 ± 0.4	53.7 ± 0.8	6.4 ± 0.4
$B^0 \to D^-[K^+\pi^-\pi^-]D^+[K^-\pi^+\pi^+]$	22.7 ± 0.6	22.4 ± 0.6	54.9 ± 0.8	<0.02
$B^+ \to K^+K^-\pi^+$	35.7 ± 0.7	10.2 ± 0.4	46.4 ± 0.7	7.7 ± 0.4
$\Lambda_b^0 \to \Lambda_c^+[pK^-\pi^+]\ \pi^-$	21.7 ± 1.0	8.9 ± 0.7	36.8 ± 1.2	32.6 ± 1.1
$B_s^0 \to J/\psi[\mu^+\mu^-]\ \phi[K^+K^-]$	26.9 ± 0.6	20.5 ± 0.5	52.5 ± 0.6	<0.02

Different types of decay reconstruction

- wrong hierarchy: correct tracks but wrong hierarchy
- Not isolated: additional tracks that do not belong to the decay
- missing tracks of the true decay

Fraction of perfect signal reconstruction approximates the tag side efficiency for FEI at Belle (II) (order a few percent for semileptonic decays and a few per mille for hadronic decays.)