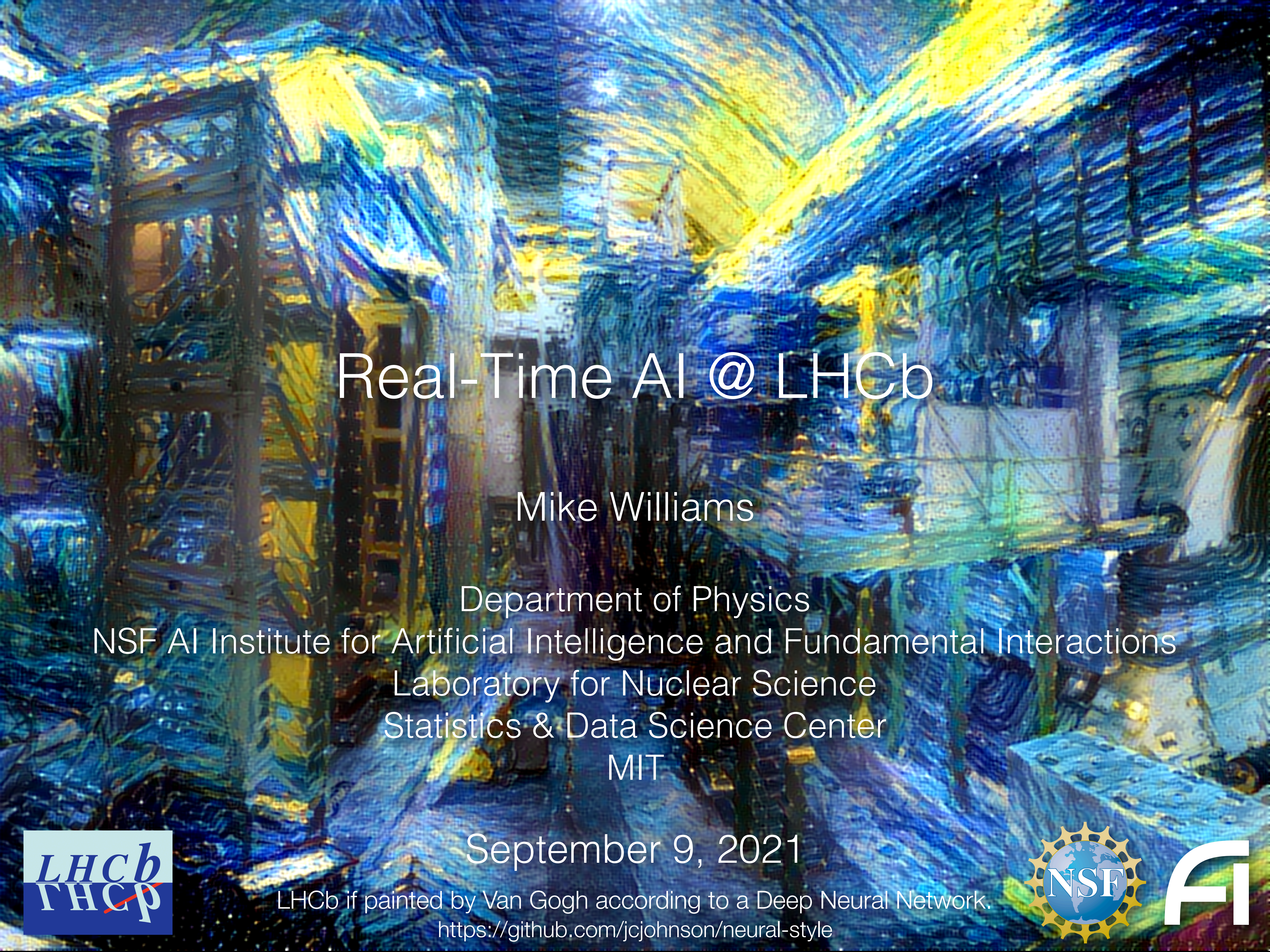
Real-Time AI @ LHCb

Mike Williams

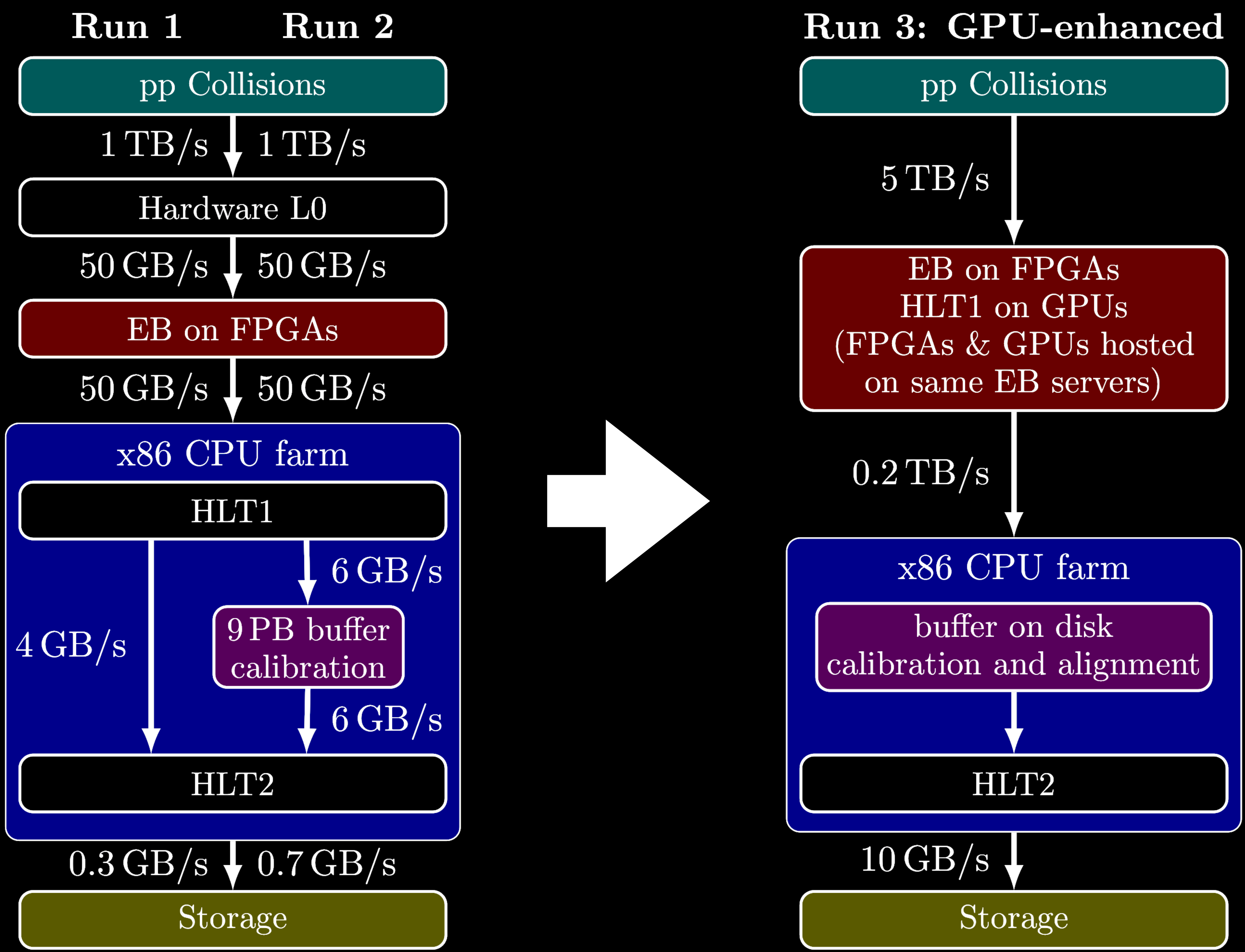
Department of Physics
NSF AI Institute for Artificial Intelligence and Fundamental Interactions
Laboratory for Nuclear Science
Statistics & Data Science Center
MIT

September 9, 2021



LHCb if painted by Van Gogh according to a Deep Neural Network.
https://github.com/jcjohnson/neural-style

Real-Time Analysis

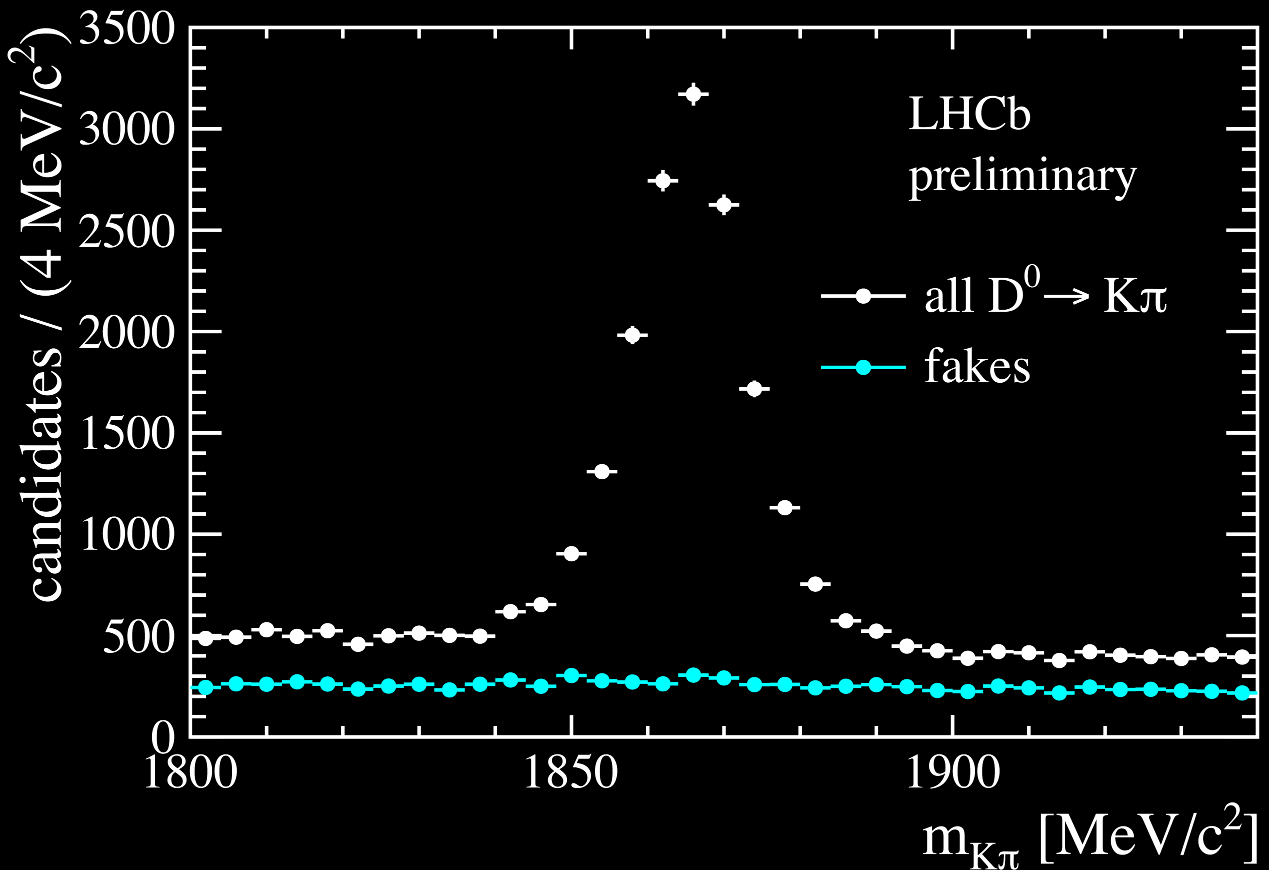


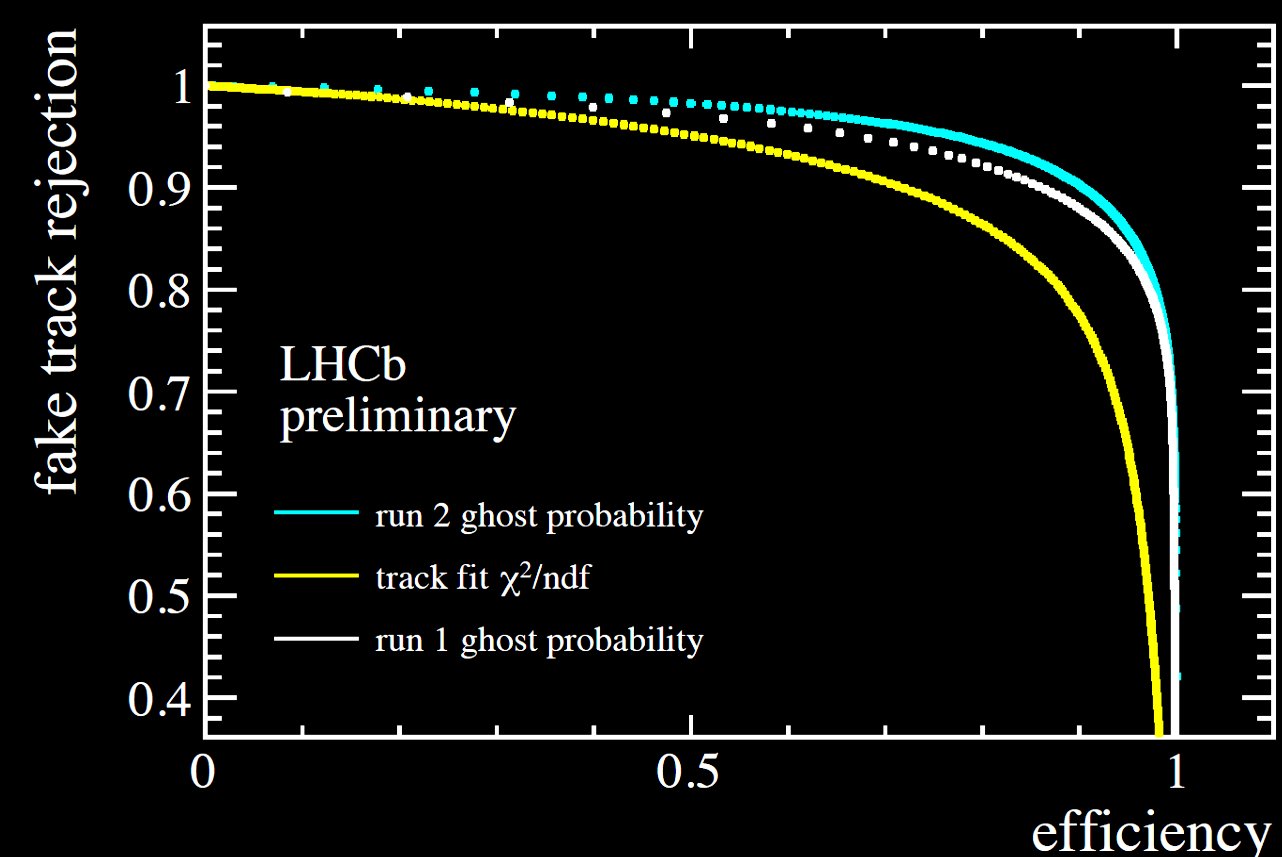
GPU-enhanced option greatly increases our discovery potential in Run 3!

Fake Tracks

Fake-track-killing NN based on 21 features, most important are hit multiplicities and track-segment chi2 values from tracking subsystems. Significantly reduces the rate of events selected in the first software-trigger stage. (also reject fake clusters with NNs)

LHCb-PUB-2017-011

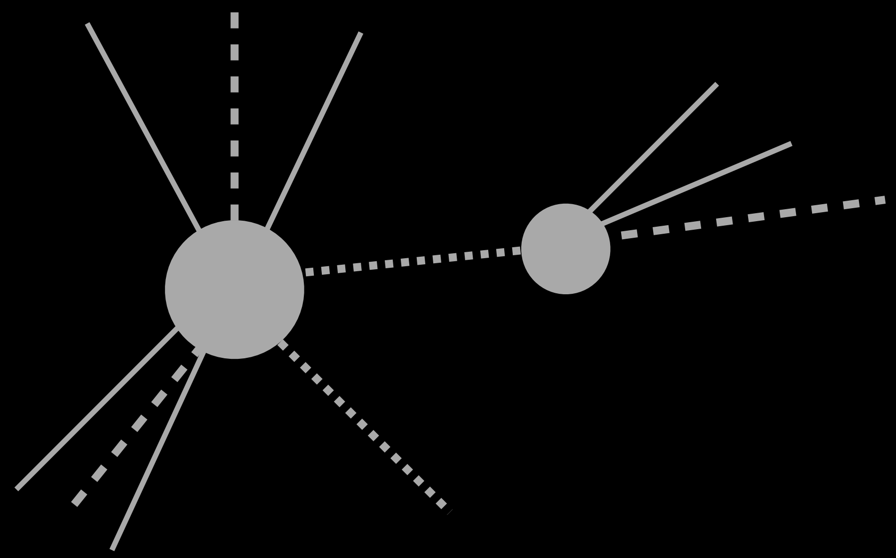




Performance evaluated using standard candle signals with and without applying a criterion on the fake-tracking-filling NN. Run on all tracks in real time so must be fast; uses custom activation function and highly optimized C++ implementation.

Real-Time Inclusive Selections

The vast majority of the LHCb trigger bandwidth (in each HLT stage) is dedicated to inclusive heavy-flavor selections, where the typical signature is a secondary vertex (SV) displaced from the primary pp collision point.



Discriminating features of beauty and charm SVs are (or related to):

- the number of tracks in the SV;
- the pT and impact parameter of the tracks;
- and the SV vertex chi2, pT, mass and *corrected* mass, flight distance, etc.

The rich feature space motivates using ML; however, uncertainties about both the run-time detector stability and deficiencies in training data raised some concerns about using ML here in real time — trigger decisions cannot be undone!

Real-Time ML-based Inclusive Selections

In Runs 1 & 2, inclusive triggers were based on (discretized) BDTs. Paraphrasing the motivations for the discretization from [1210.6861]:

- BDTs could define signal regions that are small relative to the run-time stability. Therefore, real-time miscalibration could result in lower real-world efficiencies that are hard to calibrate.
- Inclusive triggers are meant to select an entire class of signals, but are trained on O(10) examples; the training data doesn't precisely reflect the signal space.
- The response must be fast, which is not the case for large BDTs (or NNs).

Back in 2010, I proposed solving these by discretizing the input features. One can think of this as pixelating an N-D image:

- As long as the pixel size >> instability, pixel migration is rare and the BDT is robust.
- Classes of signals look more similar without fine-grained resolution.
- Discretization permits caching all responses in a LUT; the BDT is fast (at expense of RAM).

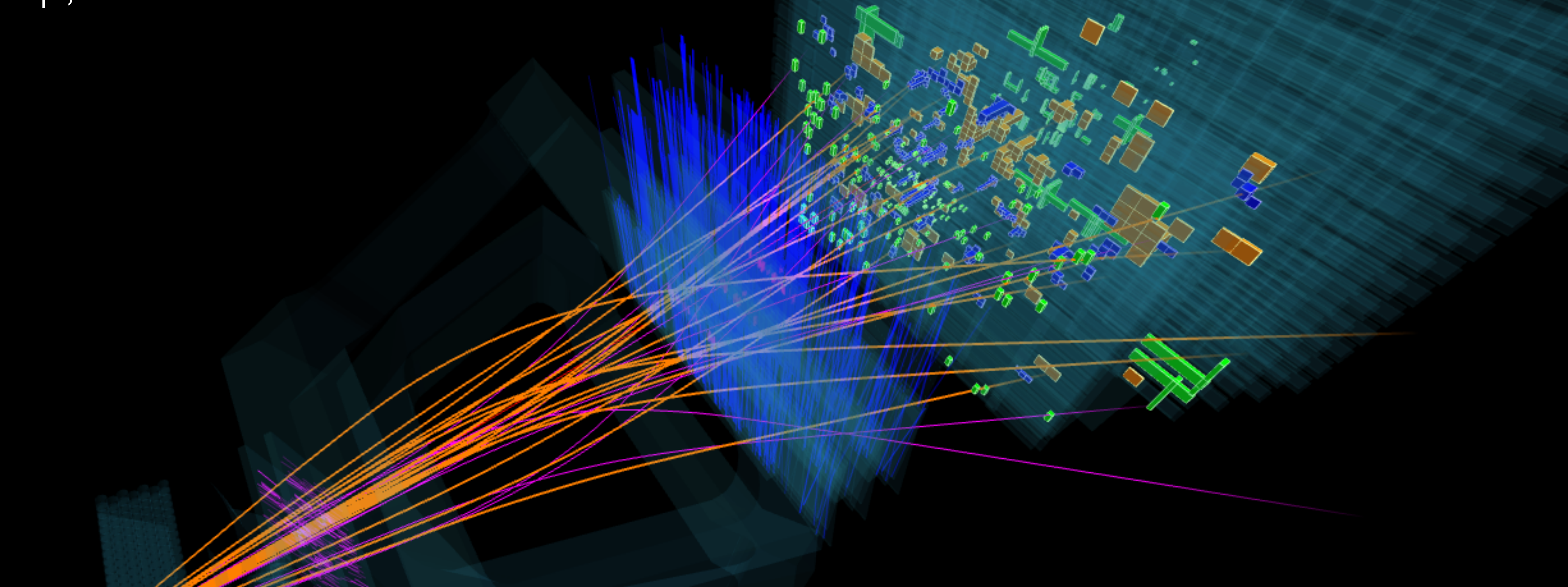
Incredibly successful (used in ~400 LHCb papers to date); however, for Run 3 we have developed a new NN architecture to satisfy all of these criteria, but in a smooth way and that also guarantees a monotonic response where desired.

V.Gligorov, MW, JINST 8 (2012) P02013.

T.Likhomanenko et al [1510.00572]

Real-Time Particle ID

Charged PID: determining whether a track originates from an e, μ, π, K, p, or fake.

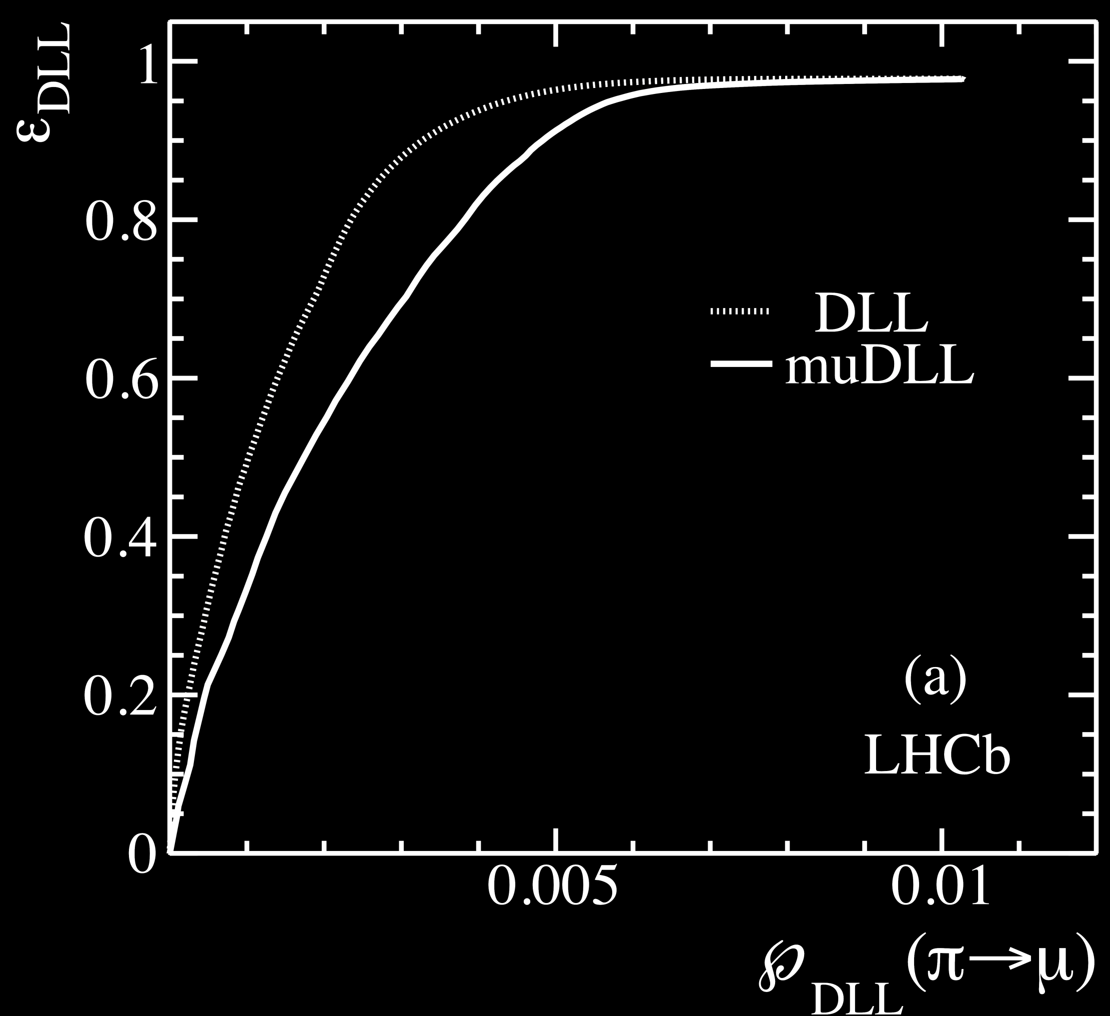


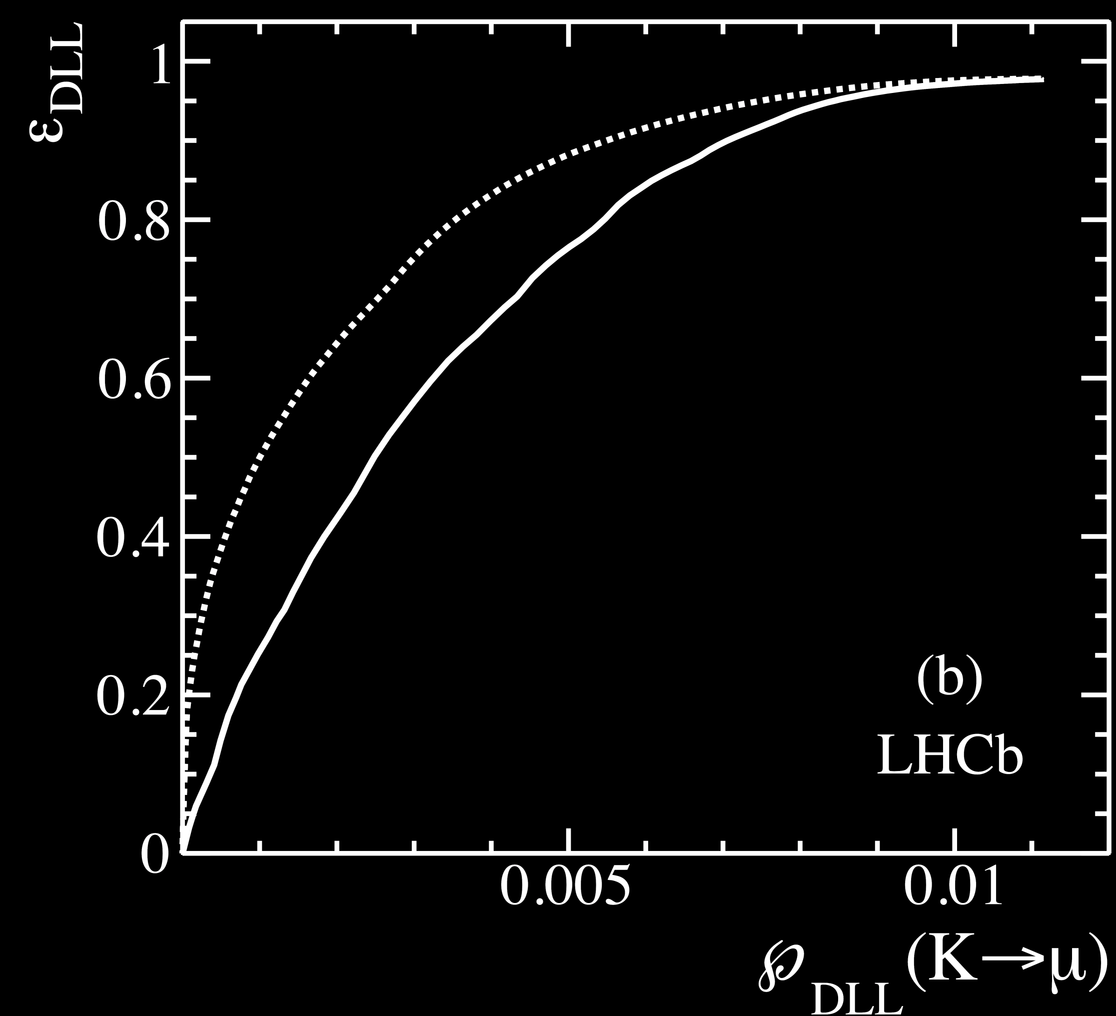
Info from the tracking, calorimeter, RICH, and muon systems all play an important role here.

(Classical) Likelihood Approach

By combining the likelihoods from the RICHs, calorimeter system, and the muon system, LHCb obtains better PID performance than using any individual system.

LHCb-DP-2014-002



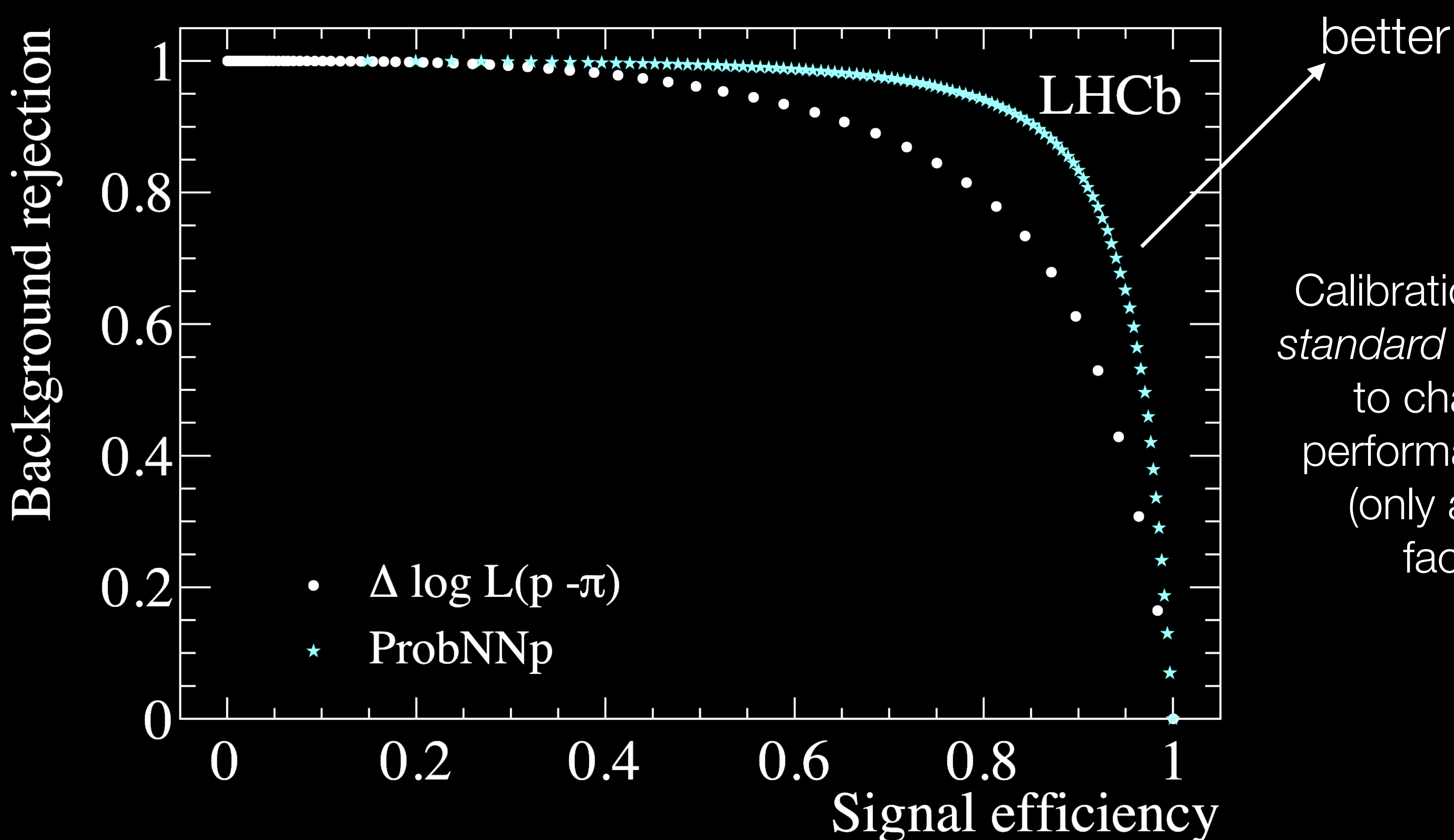


Consider the common case of K→μ decay in flight. If it was still a kaon when it passed through the RICH, then the RICH likelihood will show this. E.g., CombDLL reduces the B→hh misID rate by a factor of 6 for a loss of only 3% of B_s→μμ signal.

ML Approach

Use ML instead to identify particle types: LHCb used NNs trained on 32 features from all subsystems, each of which is trained to identify a specific particle type.

LHCb-DP-2014-002



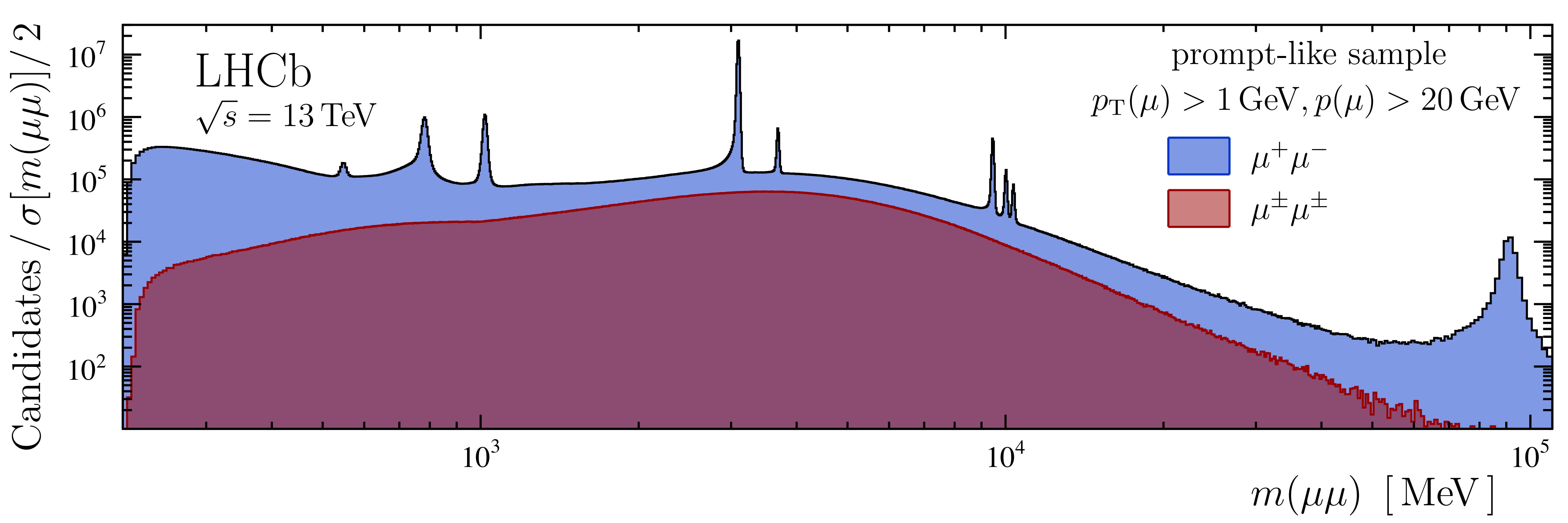
Calibration samples from *standard candles* are used to characterize the performance of the NNs (only assumption is factorization).

Typically get ~3x less misID background per particle. Currently exploring more advanced algorithms for Run 3, which can further reduce the BKGDs.

ML Approach in Real Time

NN-based PID used in real-time for trigger selections, including very hard cuts on this used to obtain a prompt dimuon sample that provided world-leading sensitivity to dark photons.

LHCb-PAPER-2017-038



The dipion mis-ID rate is a few per 10M at about 50% dimuon efficiency, which enables selecting a high-purity prompt dimuon sample even at low mass where the prompt dipion rate is huge. (Without the NN-based PID, triggering on this reaction would not fit into the available bandwidth.)

Summary

- LHCb successfully managed to calibrate — and fully reconstruct — all data in real time in Run 2.

- Since 2011, we have used ML-based selections in our primary trigger algorithms. Roughly 400 LHCb papers thus far are based on ML-selected data. These were based on a discretization method; a novel NN architecture has been developed for Run 3.

- Since 2015, fake tracks and clusters have been rejected in real time using NNs.

- Since 2016, NN-based particle ID selections have been used in the real-time selections. Several high-profile results published in Run 2 were only possible because of the performance increases provided by ML.

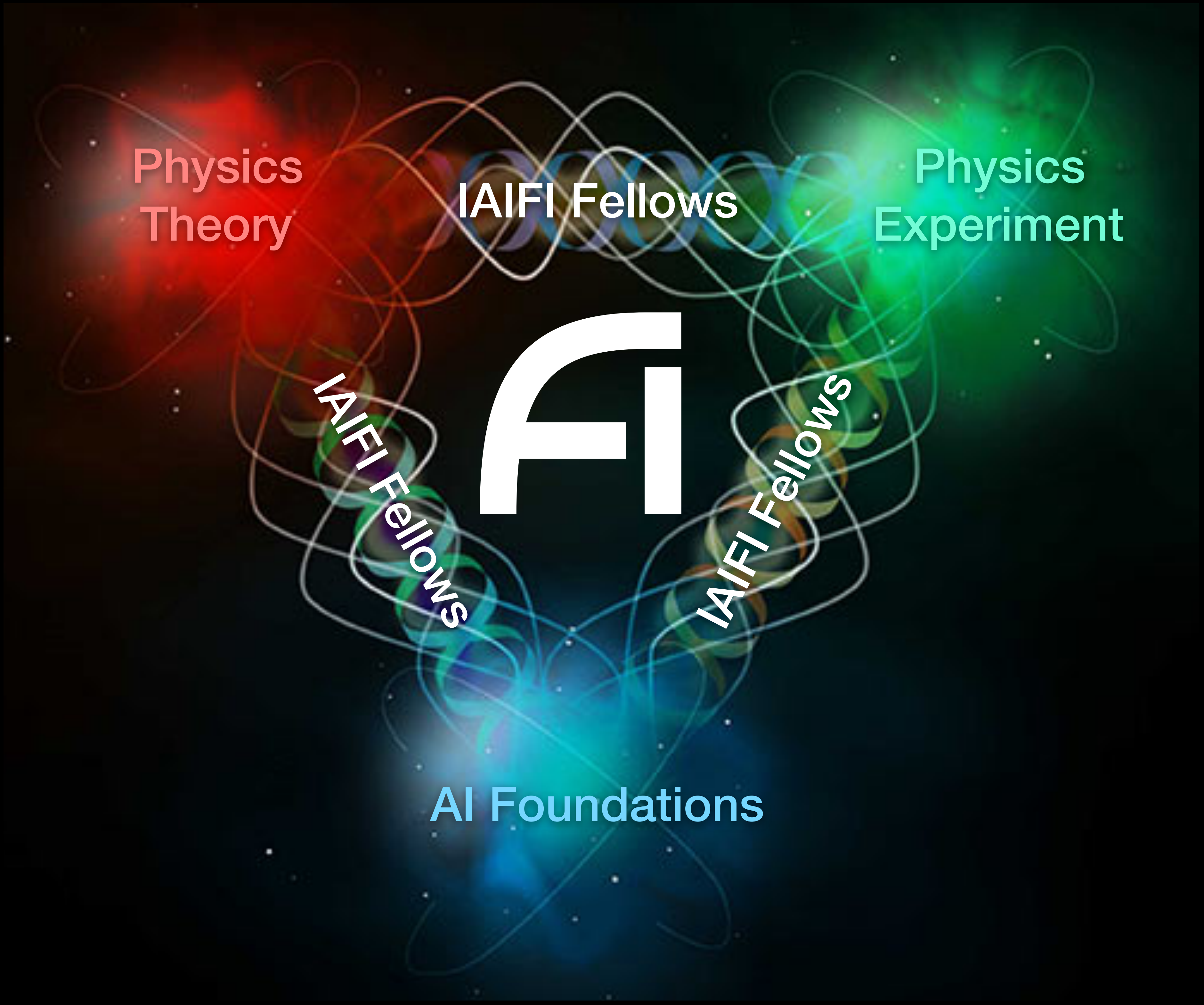
- In Run 3, we are removing our hardware trigger and will process every event in a GPU-based application that will track all particles with a very low pT threshold (and possibly do much more).

- Many other studies are underway to expand the use of real-time AI in Run 3 and beyond!

Institute for Artificial Intelligence and Fundamental Interactions *(IAIFI /ai fai/ https://iaifi.org)*

The IAIFI is advancing physics knowledge — from the smallest building blocks of nature to the largest structures in the universe — and galvanizing AI research innovation.

IAIFI Fellowship Competition



IAIFI Fellows are the gluons of the institute. They are given complete freedom of choosing their research projects/directions within IAIFI: https://iaifi.org/fellows.html

Thanks!
Questions?

LHCb if painted by Van Gogh according to a Deep Neural Network.
https://github.com/jcjohnson/neural-style