Computing and IA in HEP experiments

ASFAE/2022/006
ASFAE/2022/010

ASFAE/2022/030

F. Martinez Vidal[1]

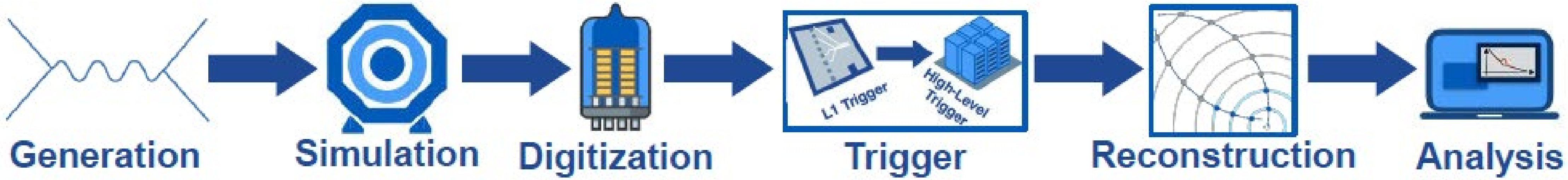
[1]IFIC, Universitat de València-CSIC

Marzo 2024, Alicante

•The ASFAE's research projects acknowledge the financial support from the MCIU with funding from the European Union NextGenerationEU and Generalitat Valenciana.

Introduction

- Main objectives of today's Particle Physics program are:
  - Probing the SM with increasing precision $\Rightarrow$ search for anomalies as evidence for BSM physics
  - Searching directly for BSM physics
- Require the processing, identification, storage and analysis of rare and/or complex signals hidden in immense amount of data (background)
  - Eg. at the LHC, ~99.999% of the data has no interest
- ML used since the 80's & 90's for (offline) event and particle identification, energy estimation, flavor tagging
- Since then, hardware and software technology progress lead to extensive HEP R&D adaptations and applications…



Generation: Integration, Sampling, Showering/hadronization, Parameter tuning, Syst. uncertainties

Simulation: Generative (GANs, VAEs,…), Refinement, Tuning, End-to-end, …

Digitization: Pileup, Electronics modeling, …

Trigger: Filtering & classification, Low-latency, Anomaly detection, On-detector, …

Reconstruction: Calibration, Tracking, Clustering, …

Analysis: Classification, Regression, Anomaly detection, Simulation-based inference, …

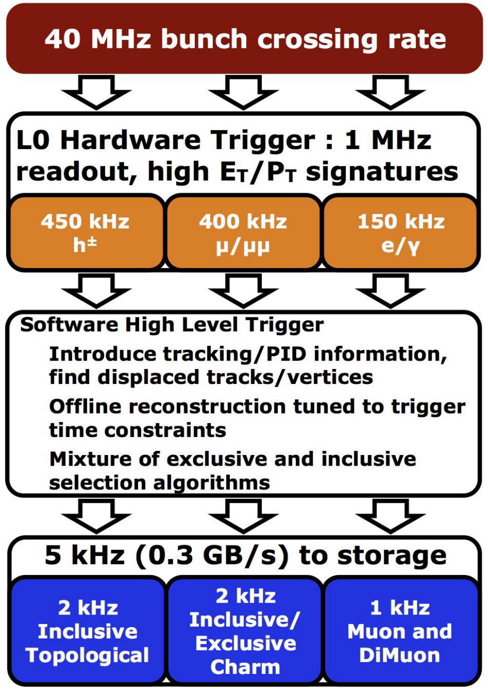
…everywhere

Case example: LHCb trigger evolution

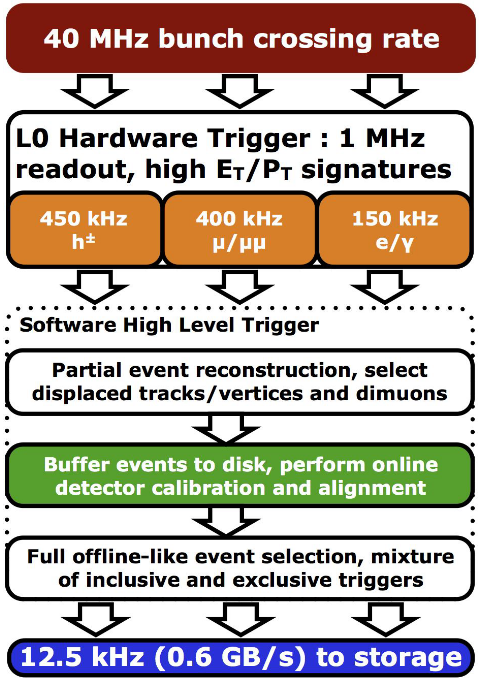
Increased Lumi by x 5
~ 6 int. per bxing

Run 1 & Run 2 (2011-2018)

LHCb 2012 Trigger Diagram

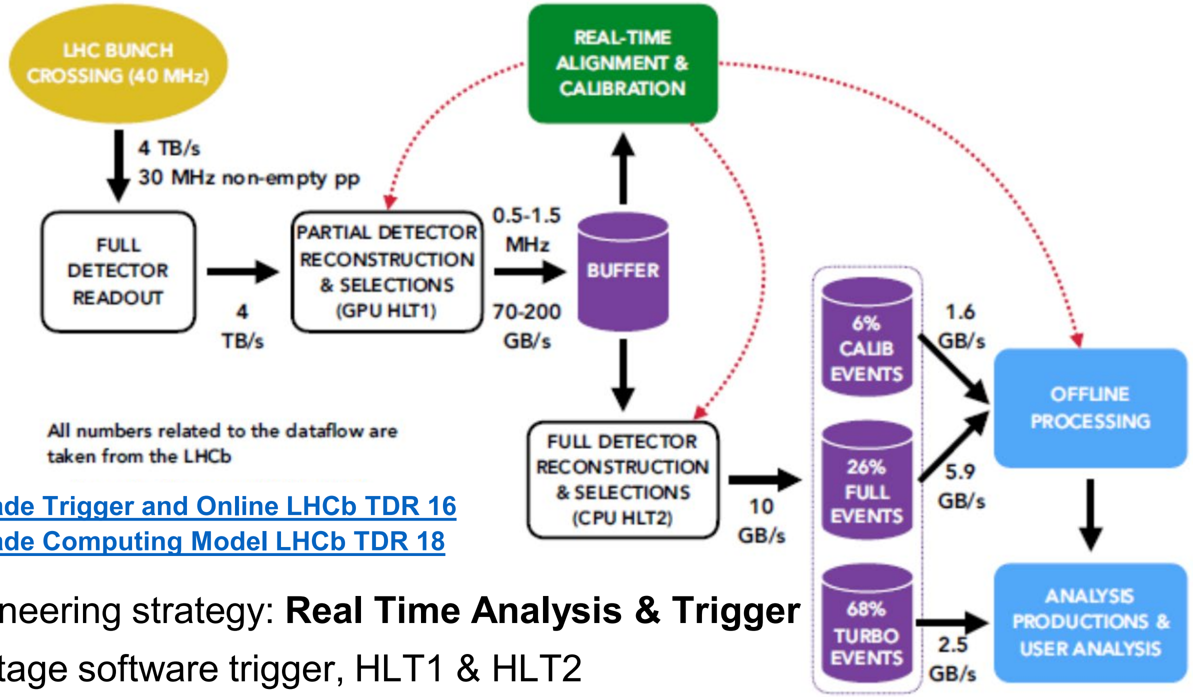


LHCb 2015 Trigger Diagram



- Traditional approach
- Hardware (L0) and software (HLT1, HLT2) on CPUs
- Exclusive and inclusive algorithms
- Full reconstruction offline
- Mostly an expert system combining different subdetectors

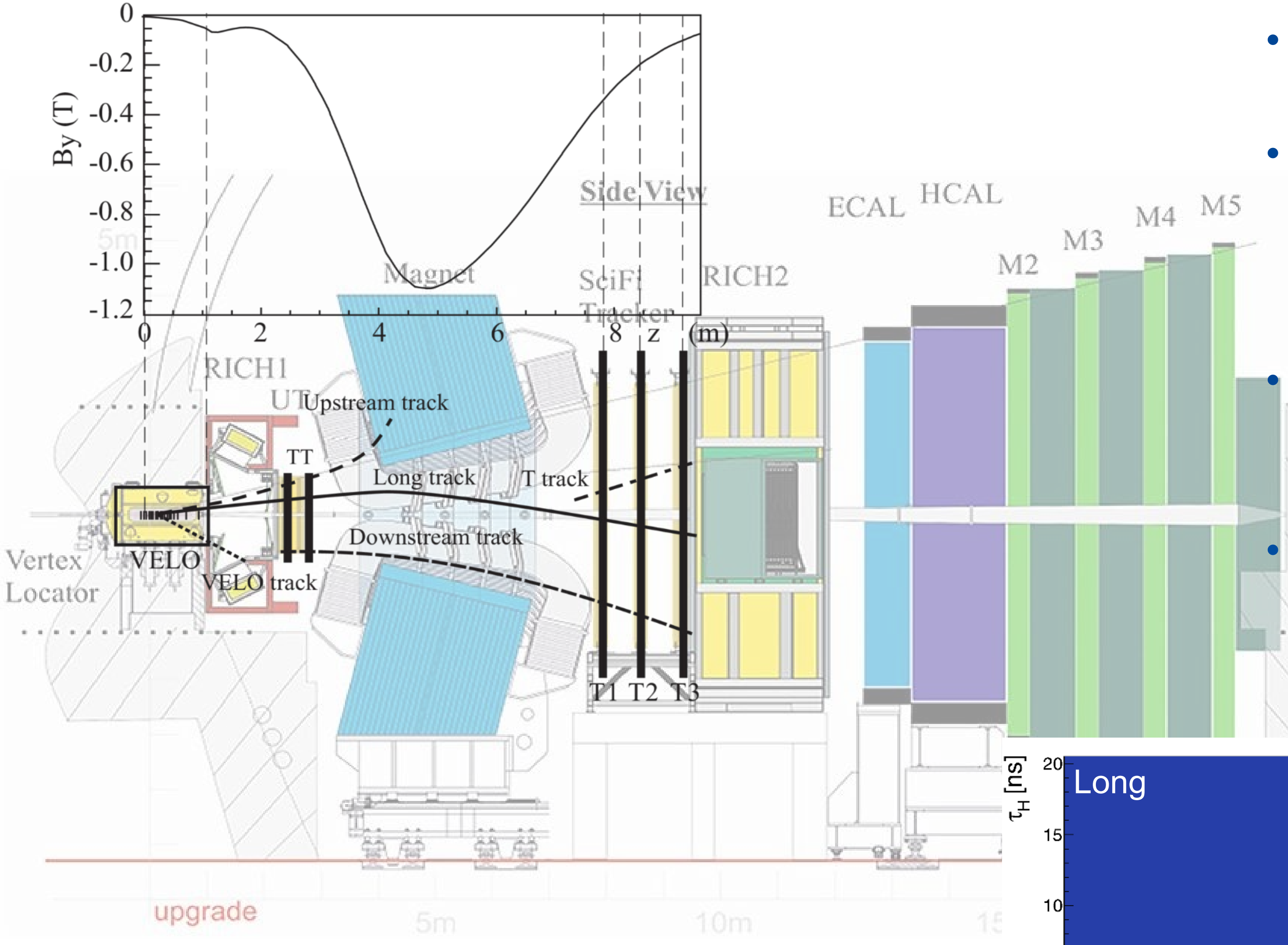
Run 3 (2022-2025) & Run 4 (2029-2032)



[Upgrade Trigger and Online LHCb TDR 16]
[Upgrade Computing Model LHCb TDR 18]

- Pioneering strategy: **Real Time Analysis & Trigger**
- 2-stage software trigger, HLT1 & HLT2
- All raw data (4 TB/s) at 30 MHz as input
- At HLT1, simplified reconstruction to reduce input rate by a factor of 20, using O(500) A50000 GPU cards
- At HLT2, full offline quality reconstruction & physics selection, reduce by another factor of 20 to 10 GB/s
- An expert system relying strongly on ML to exploit internal symmetries & correlations within and between subdetectors…

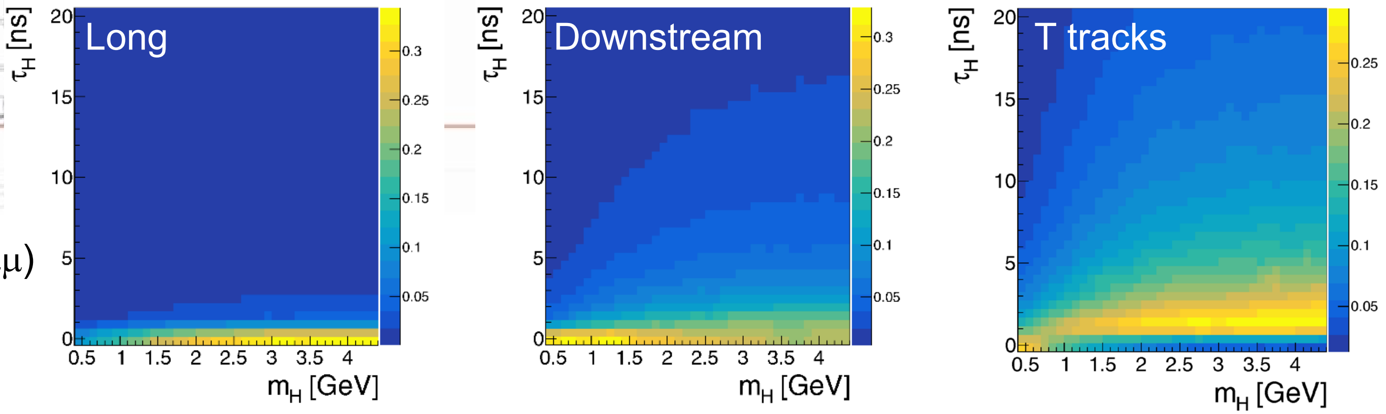
Track types and LLPs @ LHCb



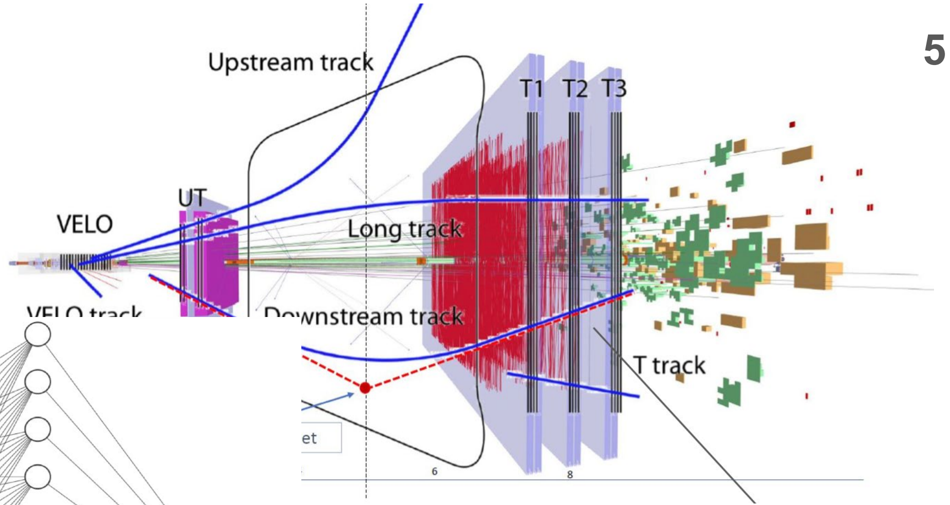
- Tracks reconstructed from segments in the different trackers
- Downstream and T tracks can be used to reconstruct LLPs with τ > 100 ps
  - Downstream for particles decaying 1-2 m from IP
  - T tracks used up to 7.5 m from IP
- The flexibility of the new trigger can be exploited to trigger on these signatures
  - Not previously triggered
- Expands physics program for decays with strange hadrons (~80% of these decay after VELO) and lifetime reach of BSM LLPs (Hidden Sector)

e.g. $B^+ \to K^+ H (\to \mu\mu)$

Acceptance

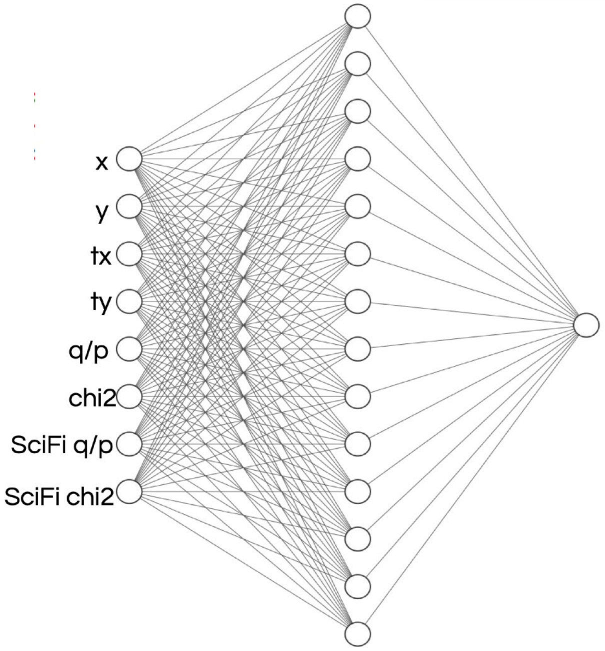


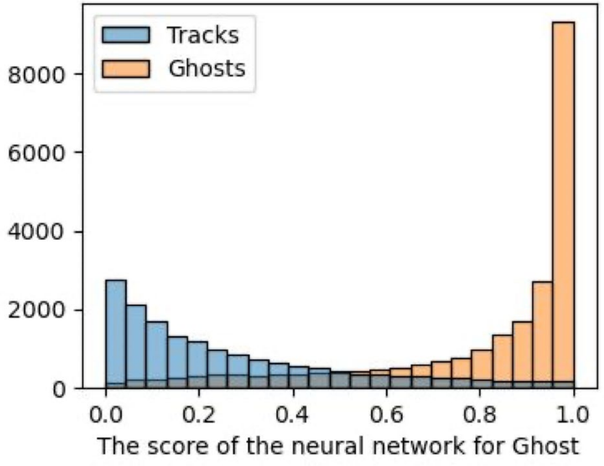
HLT1 Downstream

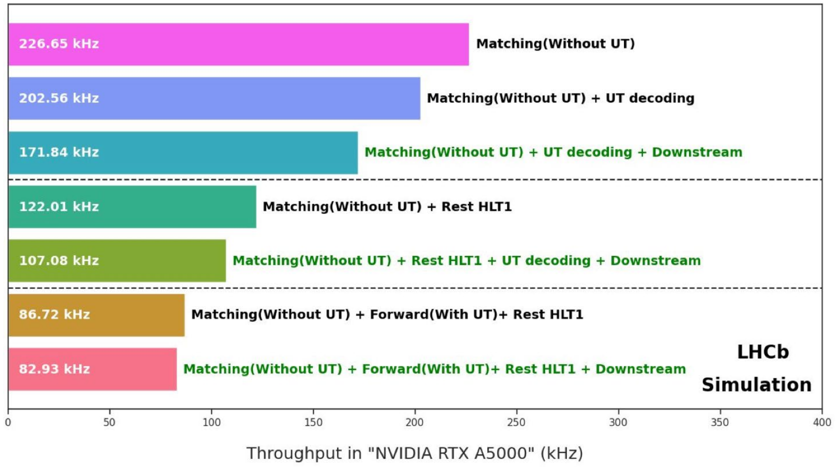


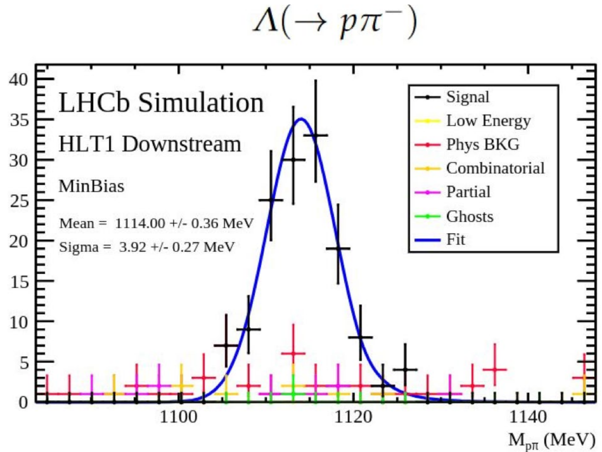
- Algorithm based on
  - Extrapolation of SciFi seeds to UT, including the effect of the magnetic field in the *x* coordinate
  - Search hits in windows of UT that are compatible with tracks coming from SciFi, and that are not used for Long tracks
- Ghost killer
  - Removal of fake tracks originating from spurious hits in the detector
  - NN with single hidden layer (14 nodes) & 8 features, model trained with RoE $B^0_s \to \phi\phi$ MC sample
- Performance from $\Lambda_b \to \Lambda\gamma$ decays
  - Efficiency ~75%
  - Ghost rate ~20%
  - Momentum resolution ~4%, ~ constant for $p < 50$ GeV
  - Global impact on throughput ~3%

CTD2023



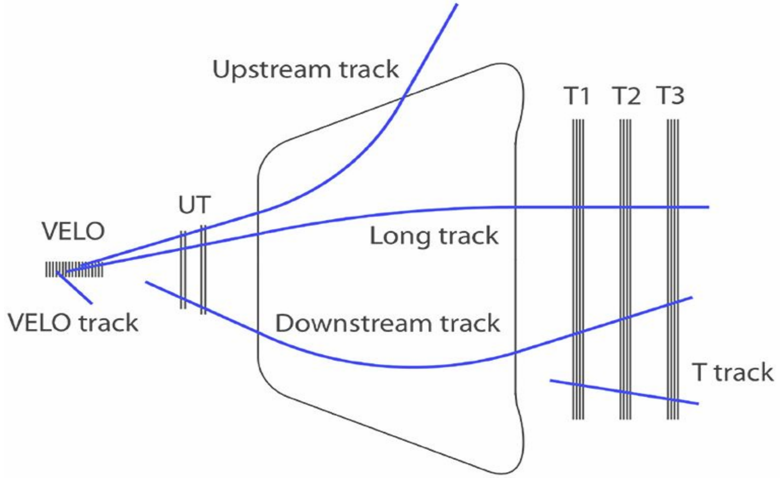


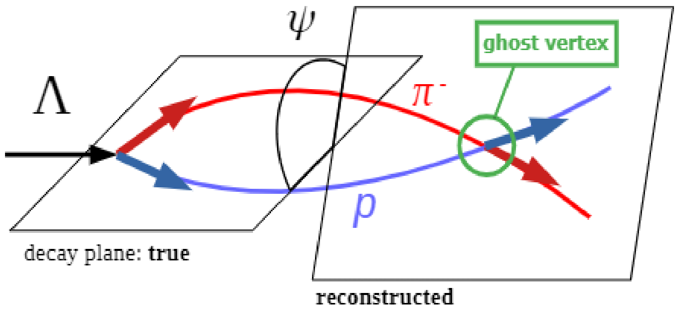




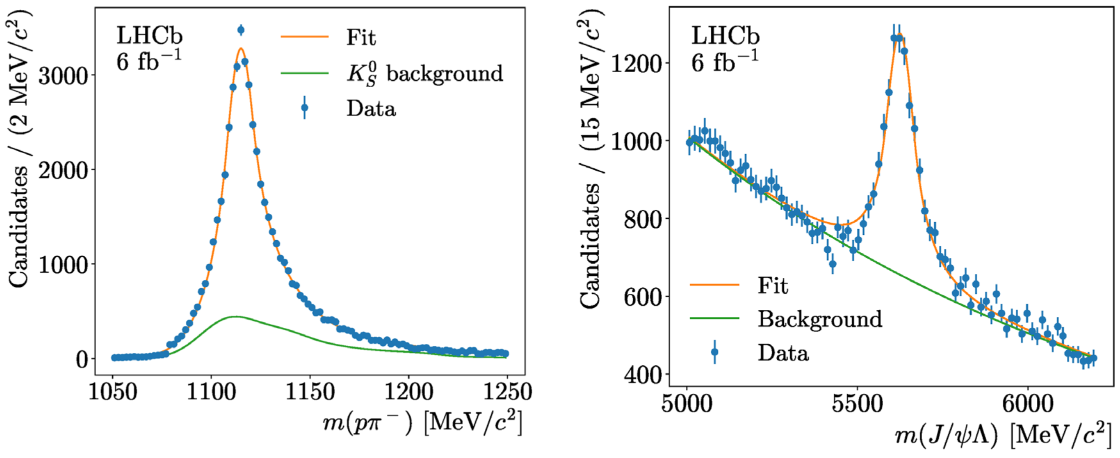
T tracks

- Unmatched SciFi segments to Long nor Downstream tracks historically unused
- Challenging
  - Short lever arm & weak B field
  - Large extrapolation through strong and inhomogeneous B field
  - Poor momentum resolution ~ 20%
  - Large combinatorics, ~1500 2-track combinations/event @ 10 MHz
  - “Ghost vertices” due to closing-track topologies
- Feasible selection and offline reconstruction for physics
- Clear benefits for physics with strangeness (~40% of the decays) and BSM LLPs with τ > 100 ps
- Could be triggered?



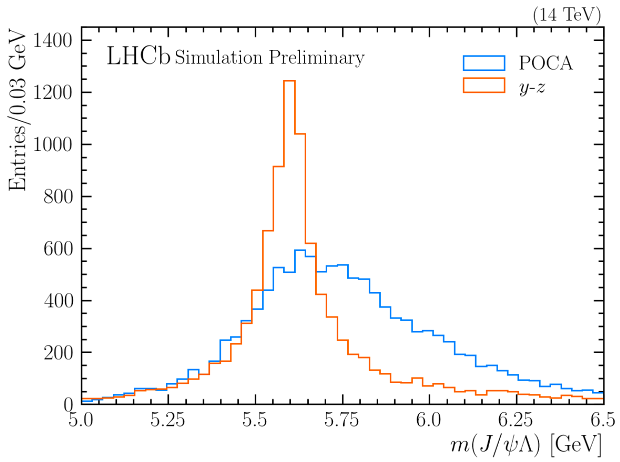


$\Lambda_b \to \Lambda\, J/\psi$ (Run 1 & 2) arXiv:2211.10920 [hep-ex]



HLT2 T tracks

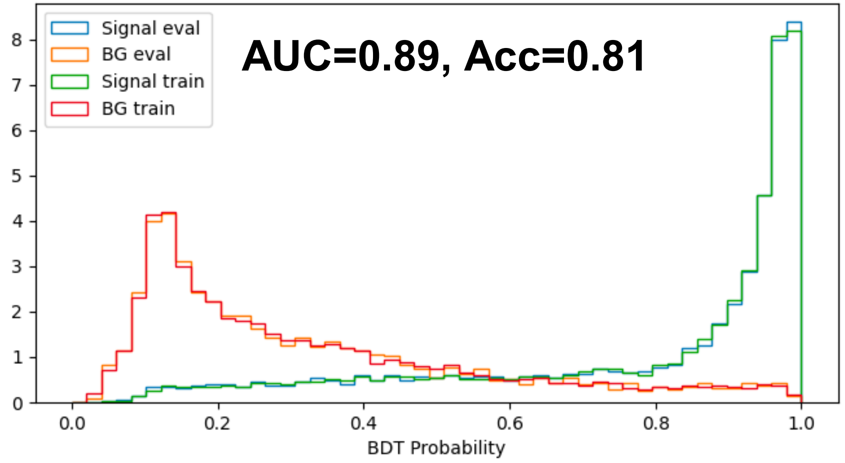
- Require clear signature in signal decay to control throughput, eg. $J/\psi \to \mu\mu$
  - Exploit small opening angle of decay products & linearity of tracks in yz place
- Vertex finding & fitting
  - Expensive, large combinatorics and extrapolation through strong/inhomogeneous field
  - Seeding with yz intersection instead of default POCA
  - 5^{th}-order Runge-Kutta extrapolation in 1^{st} iteration, polynomial after (for now)
- For processes w/o clear signatures



Binary CatBoost BDT to filter single tracks (20 iterations/trees)

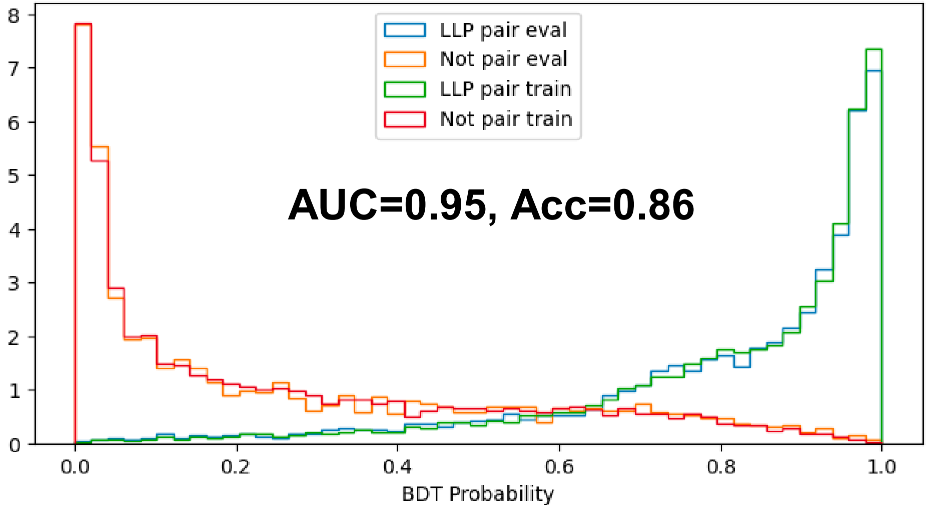
– Use "unfitted" tracks, proton or pion from Λ or K^0_s as signal

– 5 features: p_T, p_z, radial distance to beam pipe, η, y



Binary CatBoost BDT to filter pairs (100 iterations/trees)

– 7 features: y_{yz}, z_{yz}, Δy, Δr, Δt_y, Δt_x, $\mathrm{sign}(t_y^1.t_y^2)$



- Reduces impact on throughput and combinations by 70%-85%, with ~75% signal efficiency
- Factory of SM and BSM signal signatures deployed/ongoing. PID also integrated

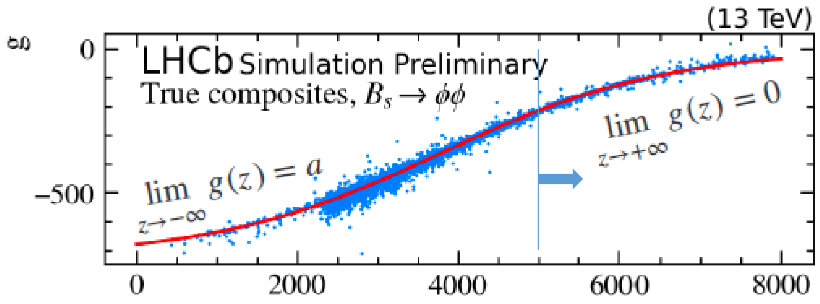
CTD2023

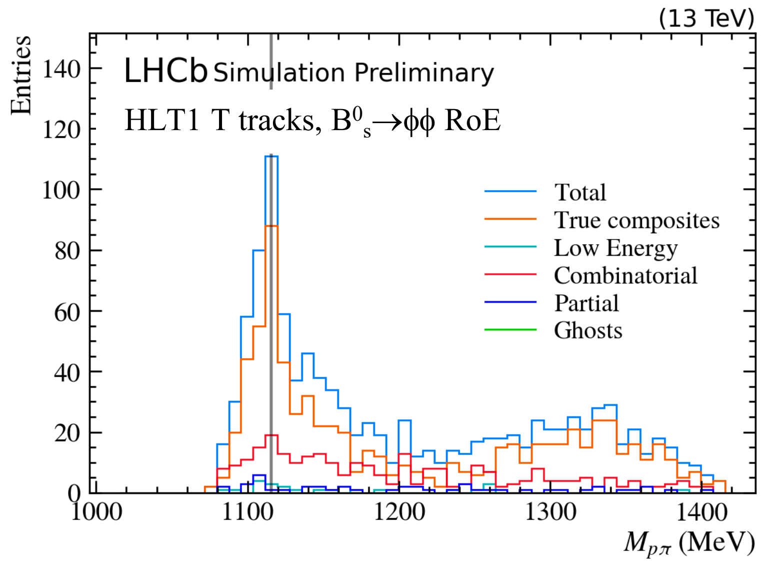
HLT1 T tracks

- No time for anything complicated
- Use simple extrapolation model

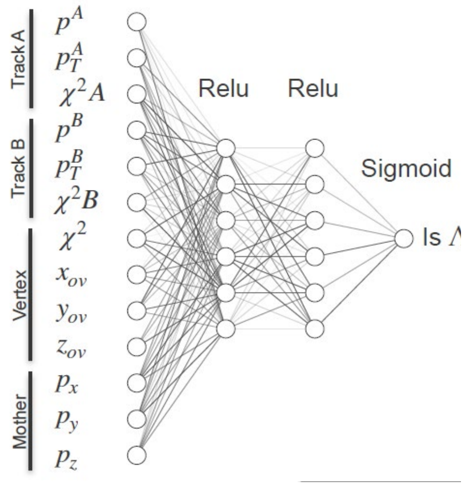
$$\begin{cases} x(z) = x_0 + tx(z - z_0) + \boxed{\dfrac{q}{p}(z - z_0)g(?)} \\ y(z) = y_0 + ty(z - z_0) \end{cases}$$

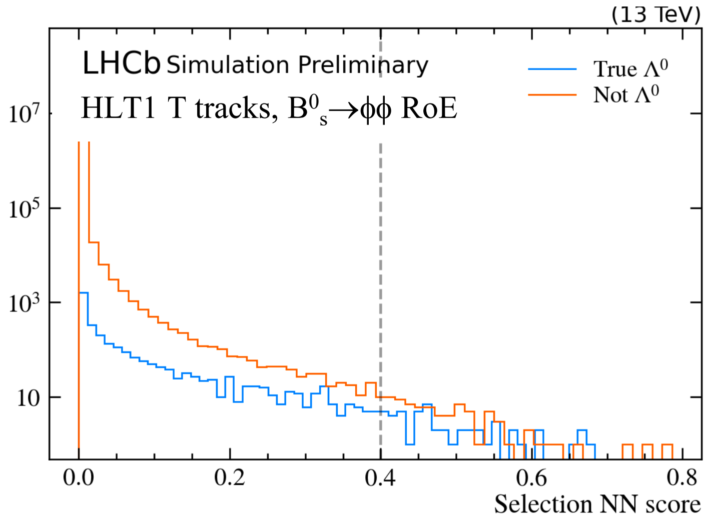
Unknown function,
Approximated with sigmoid





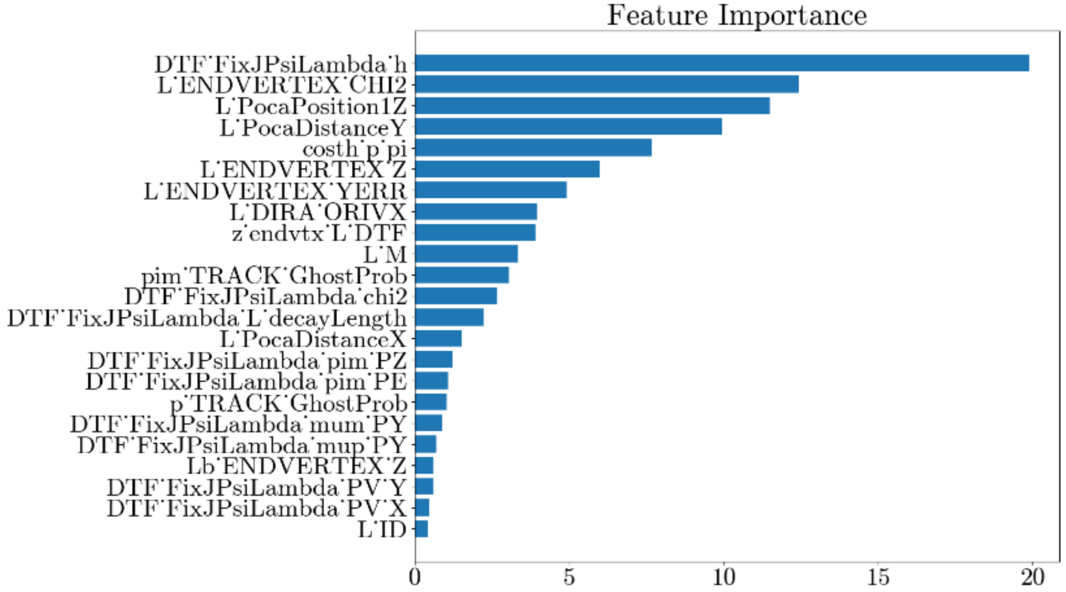
- For each track pair, vertex hypothesis created at POCA
  - Allows simple improvement of track momentum and slope
- Selection of two-track candidates
  - Requires opposite q/p signs, 1 < POCA < 8 m, DOCA < 0.1 m
  - NN with 2 hidden layers, with 6 nodes each & 13 input features
- First CUDA implementation in place, optimization ongoing

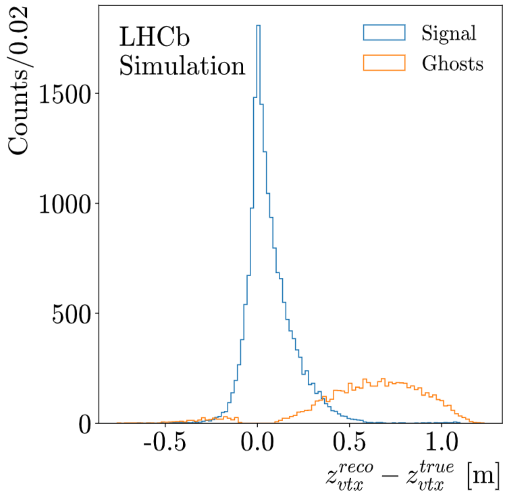


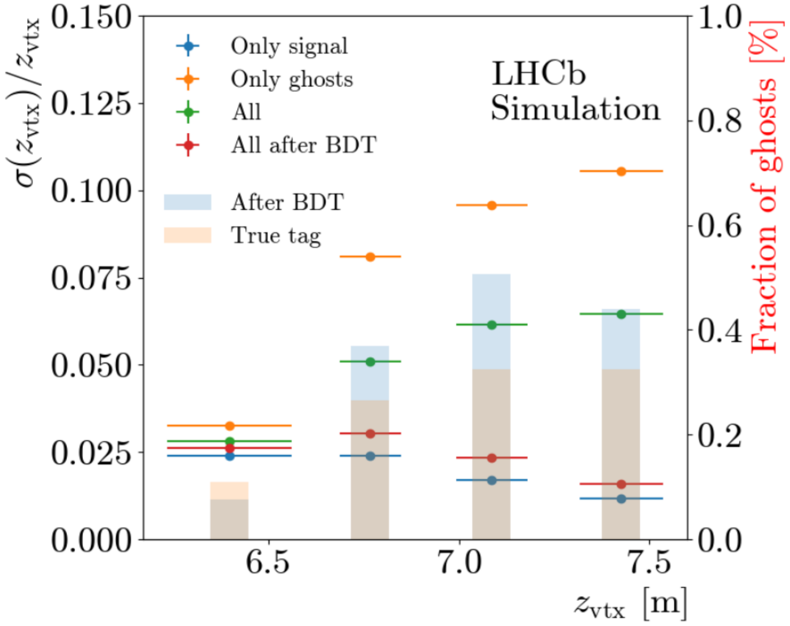


Offline ghost vertex reduction

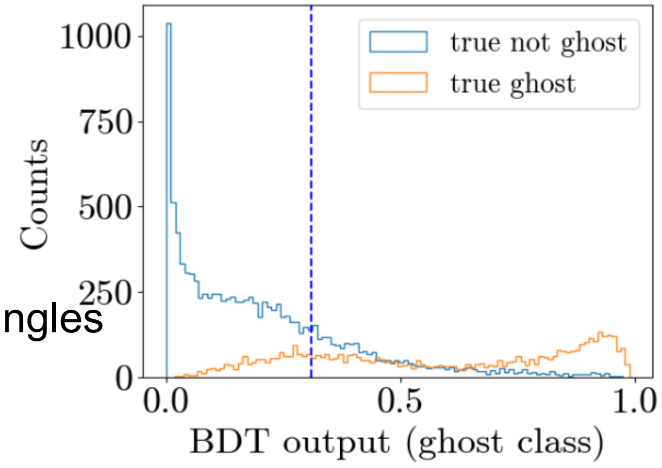
- Closing-track events in very displaced vertex finding are an issue, mainly for low-Q decays
  - For $\Lambda \to p\pi^-$, ~30% of events affected, poor vertex resolution and large bias
- Binary CatBoost BDT classifier
  - Trained with topological variables

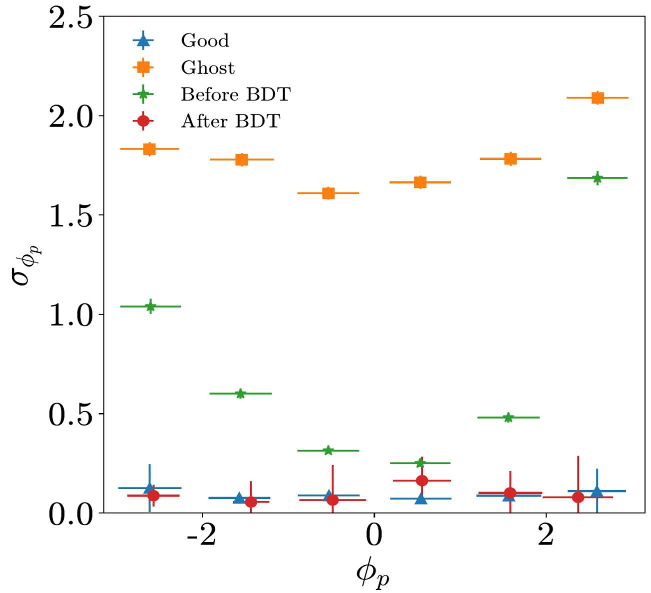






AUC = 0.84



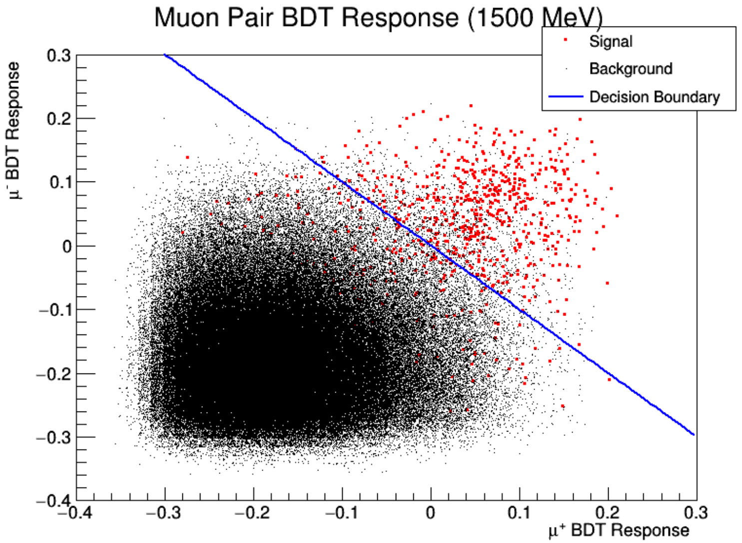


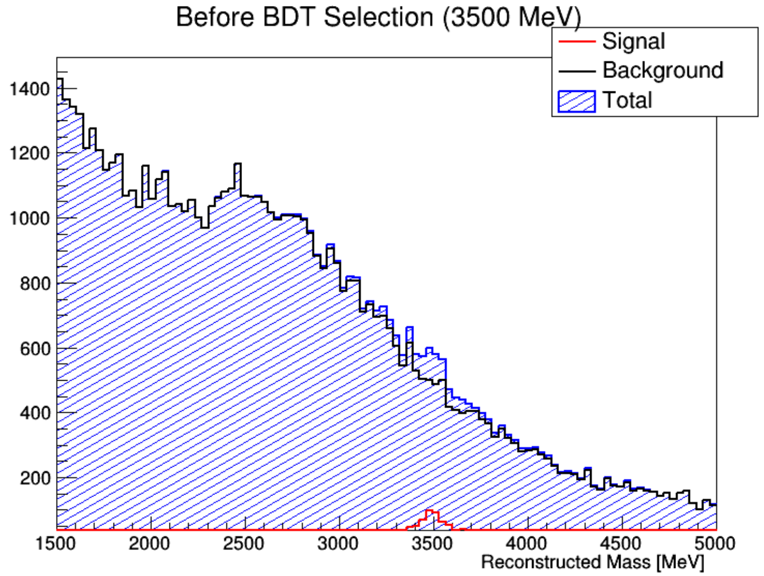
- Large impact on vertex and proton helicity angles resolutions, negligible on mass resolution

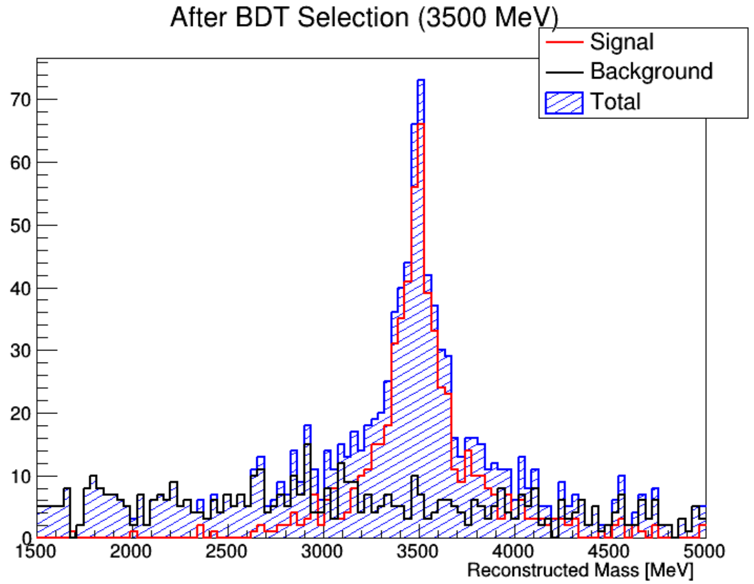
Offline/HLT2 combinatorial background suppression

- Suppression of combinatorial background in largely displaced vertex searches is a challenge due to large combinatorics and poorer resolutions

- Example: search for dark Higgs in $B^+ \to K^+ H(\to \mu^+\mu^-)$ decays
  – BDT trained on single muons
  – Features: p_T, p_z, IP, χ^2_{IP}, $\Delta LL(p\text{-}\pi)$, $\Delta LL(K\text{-}\pi)$, $\Delta LL(\mu\text{-}\pi)^{CALO}$, t_x, t_y
  – Applied to combinations of tracks to account for correlations





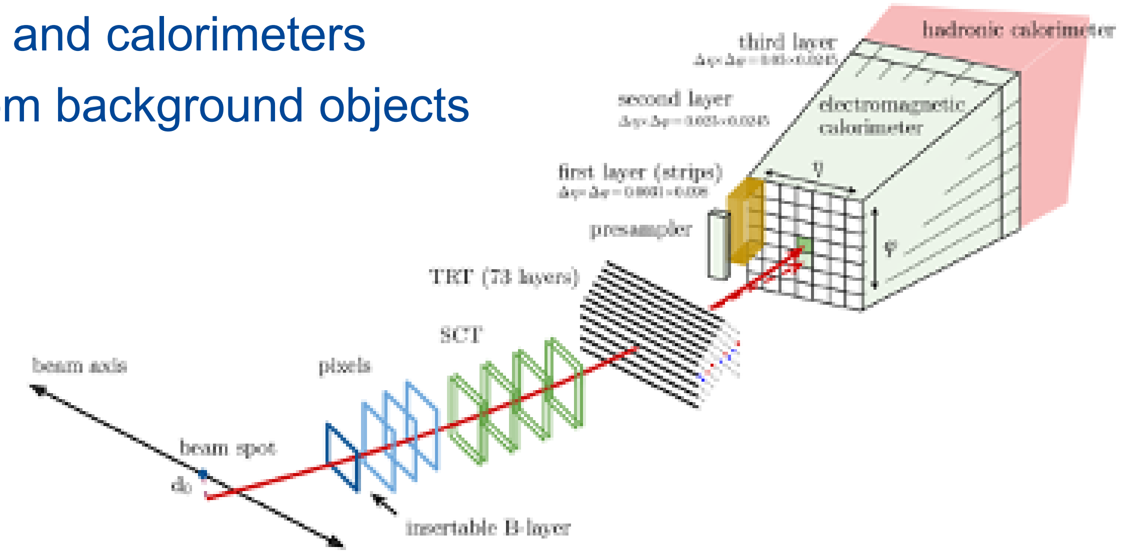


– Suppression of more than 99% of background
– Signal efficiency > 50%

- Other BSM modes under investigation & development

ATLAS electron identification

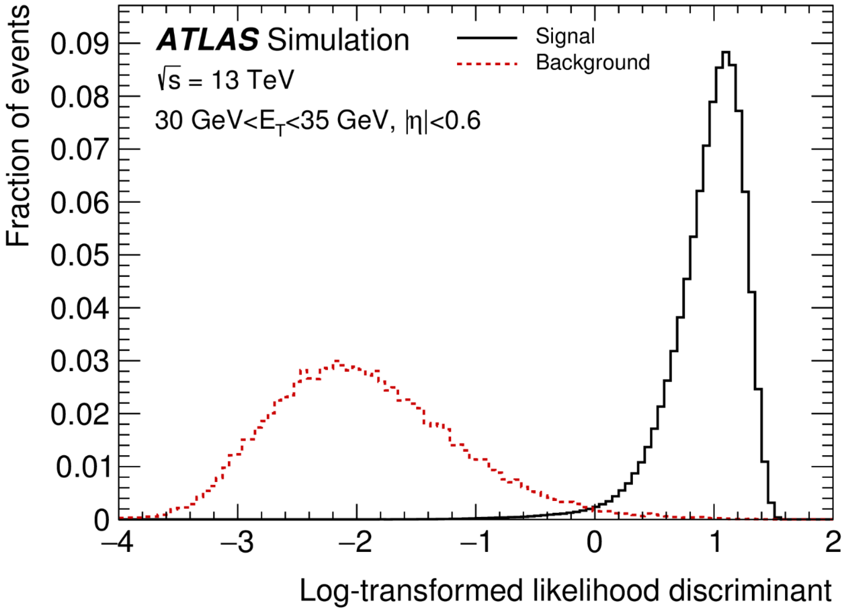
- Electrons are reconstructed combining information of tracker and calorimeters
- Likelihood (LH) discriminant to separate prompt electrons from background objects (hadrons, non-prompt electrons from b-quark decays, etc.)
- LH built using variables (PDFs) related to
  - properties of the track
  - shower development (energy in each calorimeter layer, depth/width of the shower)
  - track-cluster matching



$$L_{\mathrm{S(B)}}(\mathbf{x}) = \prod_i P_{\mathrm{S(B)},i}(x_i)$$

$$d_{\mathrm{L}} = \frac{L_{\mathrm{S}}}{L_{\mathrm{S}} + L_{\mathrm{B}}}$$

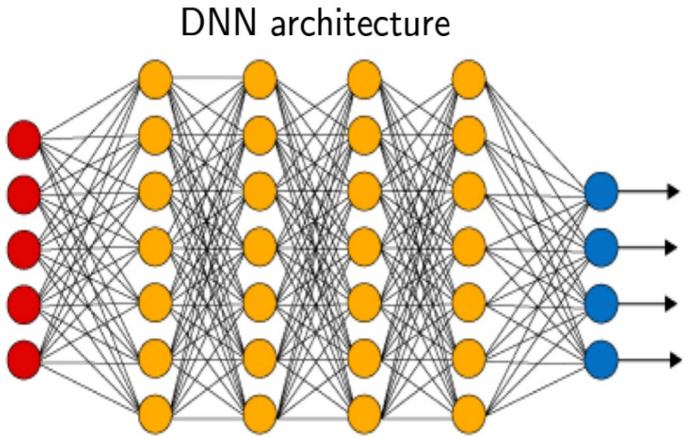
Eur. Phys. J. C 79 (2019) 639



- Better background rejection than a simple cut-based algorithm (criteria applied on each variable individually)
- Drawbacks:
  - Doesn't take into account correlations between input variables
  - Binomial classification (signal vs. background)

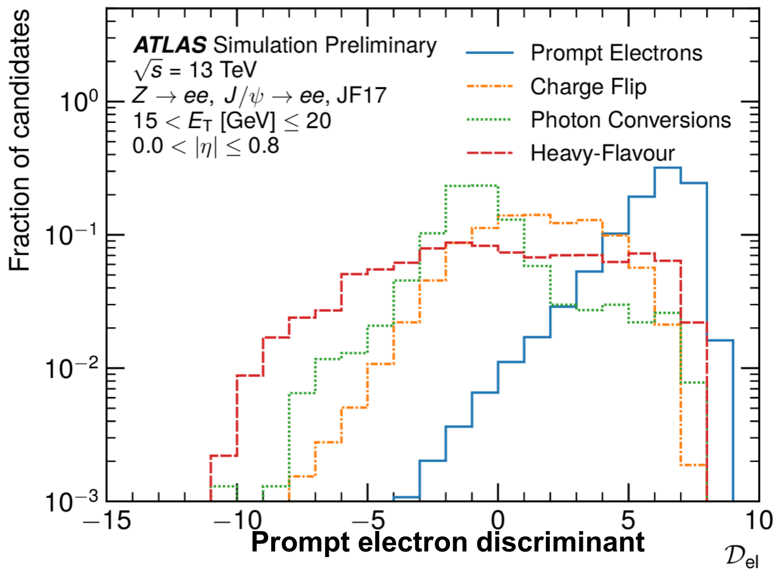
ATLAS electron identification using a DNN

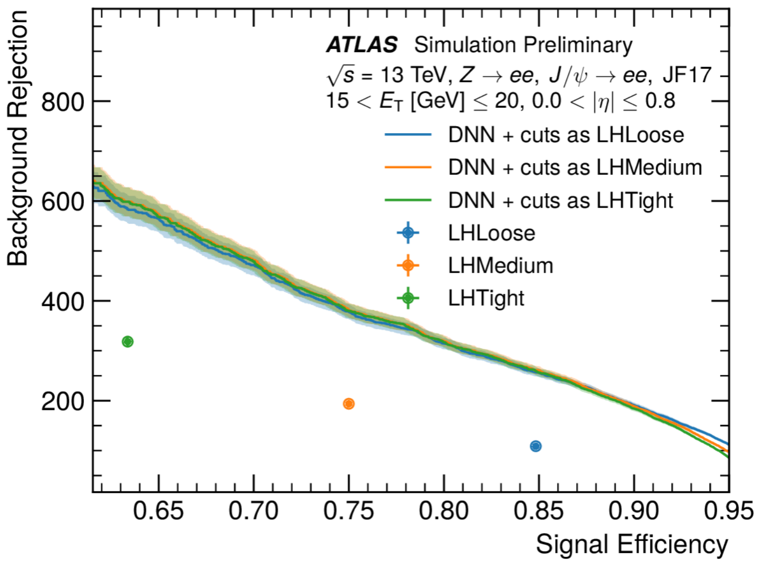
- Architecture and configuration detailed in **ATL-PHYS-PUB-2022-022**
- Neural Networks are powerful in signal to background discrimination → Replacing previous electron identification based on a Likelihood-based approach (LH), while using **similar high-level input variables** as LH



Class	Description
Prompt Electrons (El)	Prompt isolated electrons coming from Z, W and J/psi
Charge Flip (CF)	Prompt with incorrectly reconstructed charge
Photon Conversion (PC)	Electrons coming from prompt photons
Heavy-Flavour (HF)	Electrons coming from a b- or c- hadron decay
Light-Flavour e/gamma	Electron coming from a u-, d- or s- hadron
Light-Flavour Hadrons	Undecayed hadrons

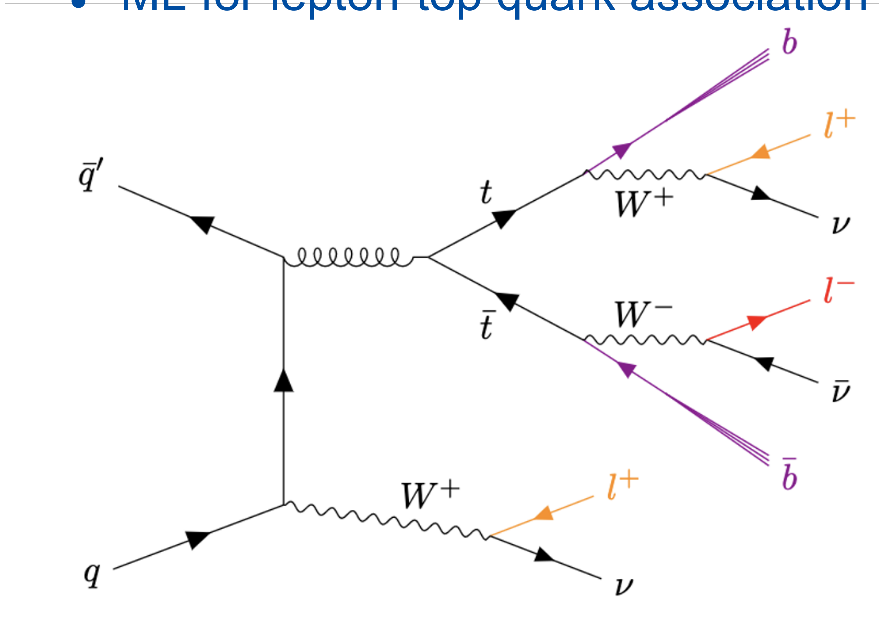
- **DNN architecture**
  - 5 hidden layers with 256 nodes each, activation function: leaky ReLU, batch normalization
- **Output layer**
  - Six outputs (signal + 5 background classes) with softmax activation for multiclass classification
- Outstanding discrimination, flexible discriminants built out of the DNN scores
- Discriminating performance comparing signal efficiency (ε) and background rejection ($1/\varepsilon$)
- At ~75% of signal efficiency, DNN outperforms LLH by a factor >2 background rejection

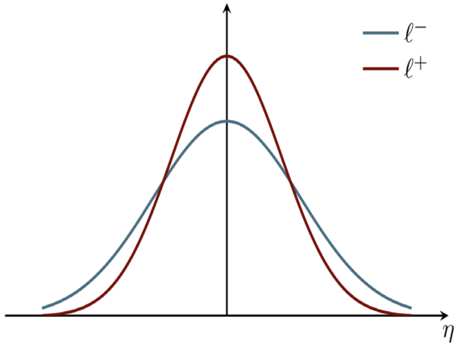




ttbarW leptonic charge asymmetry

- ML for lepton-top quark association





JHEP **07** (2023) 033

Phys. Lett. B **736** (2014) 252

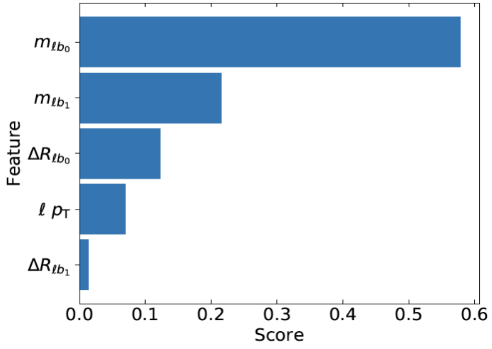
$$A_c^\ell = -13.2 \pm 0.1 \ \text{(theory)}\ \%$$

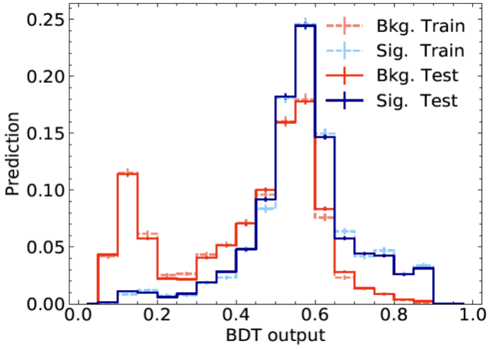
inclusive NLO+PS

$$A_c^\ell = \frac{N(\Delta|\eta_\ell| > 0) - N(\Delta|\eta_\ell| < 0)}{N(\Delta|\eta_\ell| > 0) + N(\Delta|\eta_\ell| < 0)} \quad \text{with} \quad \Delta|\eta_\ell| = |\eta_{\ell^+}| - |\eta_{\ell^-}|$$

- Charge asymmetry between the leptons coming from top and antitop quarks: enhanced in ttbarW events compared to ttbar
- Experimental challenge in ttbarW 3l final state:
  - Identify the charged leptons coming from top and antitop quarks
  - The correct even lepton is selected using GBDT

Odd lepton: always from (anti)top quark
Even lepton: need to select the correct one

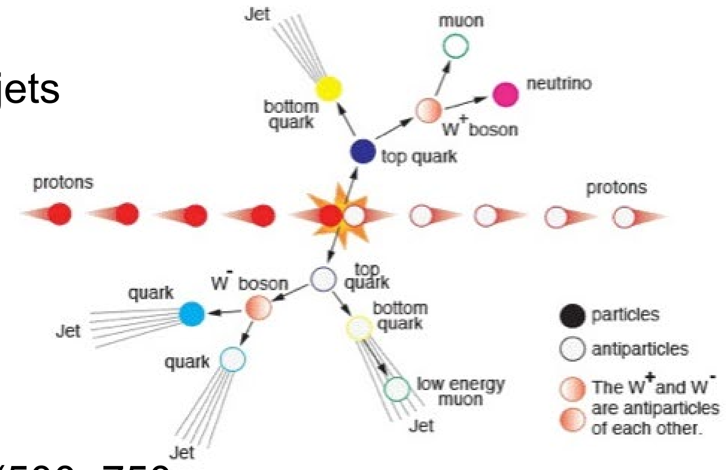




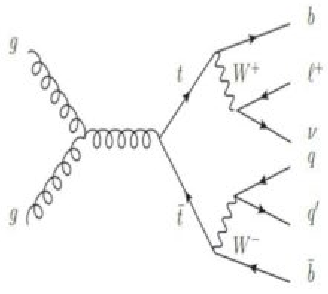
- For each event, trained even leptons (object level MVA, per lepton)
- The accuracy of the BDT for selecting the correct lepton is 71%

Searches for ttbar resonances

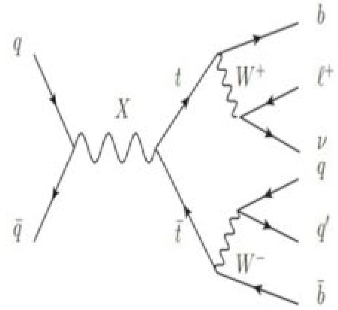
- Application of ML methods (DT & NN) to signal/background classification for searches for ttbar resonances
- Features
  - p_T, η, ϕ of the lepton (μ or e)
  - p_T, η, ϕ, b-tagging of the 4 more energetic jets
  - p_T, ϕ of missing E_T
  - both W invariant masses



SM

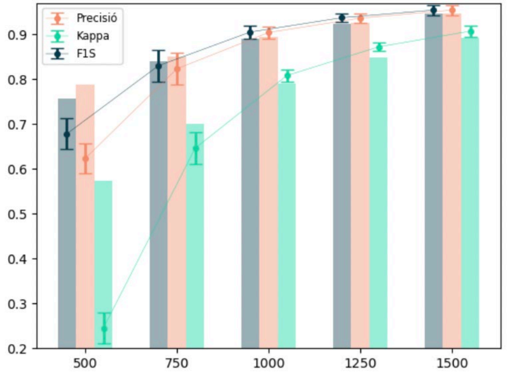


BSM

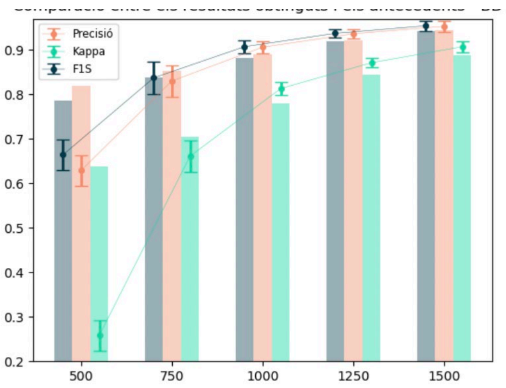


- Training with simulated events
  - 1 background dataset
  - 5 signal datasets for 5 resonance masses (500, 750, 1000, 1250, 1500 GeV)
- Comparison of Random Forest (RFs), BDTs and NNs
- Syst. Uncertainties varying hyperparameters
  - NN studies for syst. errors in progress

arXiv:2002.12220 [hep-ph]

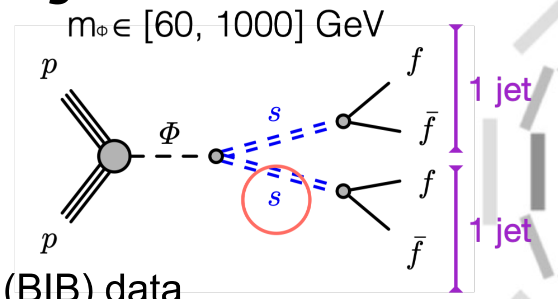


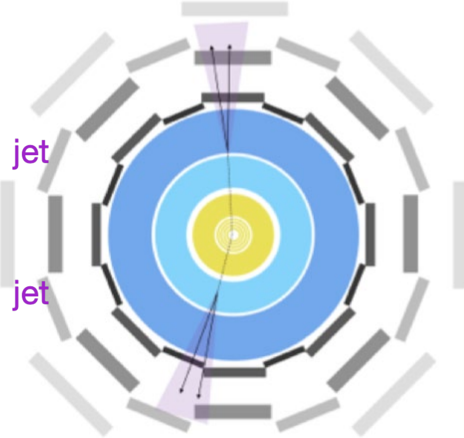
RF Signal mass (GeV)



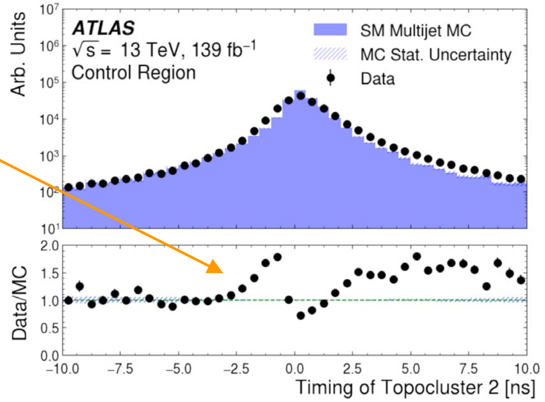
BDT Signal mass (GeV)

LLPs decaying into displaced hadronic jets





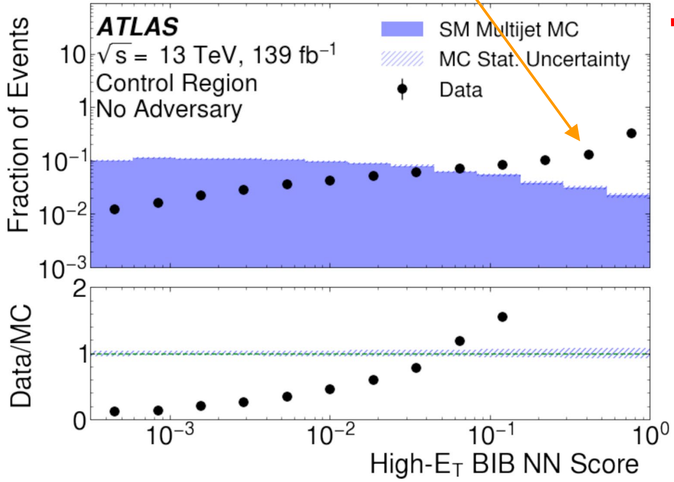
- Hidden Sector with a heavy boson decaying to long-lived scalars
- Signature: **2 displaced jets** in the ATLAS calorimeter
- Adversarial Neural Network separates signal from two types of bkgs
  - training on a set of signal MC, SM background MC and beam-induced background (BIB) data
  - mismodelling in input variables (eg. jet timing) had a big impact in the data/MC behaviour of the NN scores
  - Adversarial NN running in a control region avoids the NN to use data/simulation differences



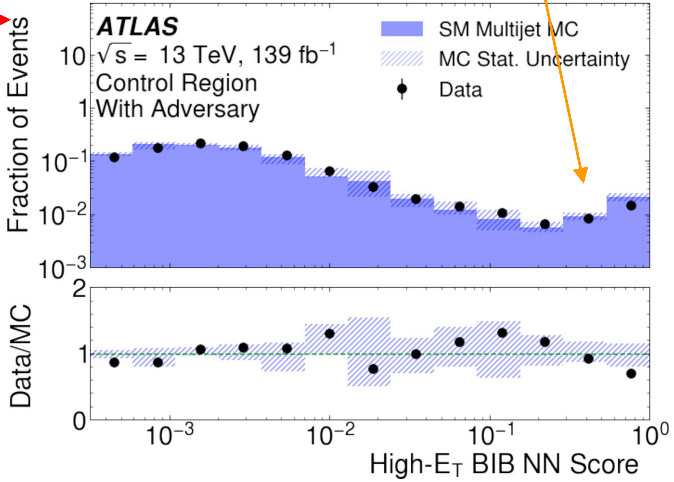
mismodelling in jet timing

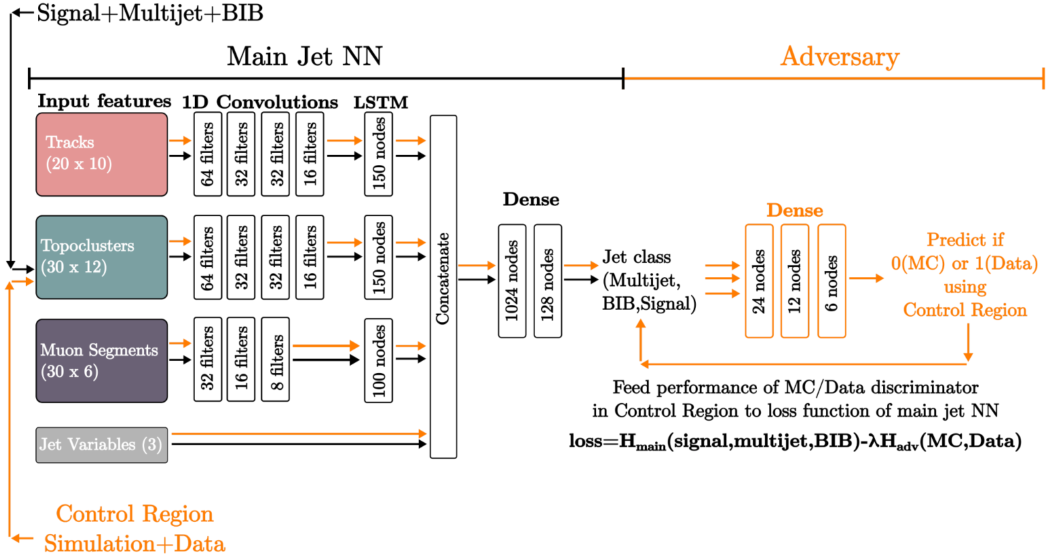
large impact in the BIB NN score

impact in the BIB NN score becomes tiny



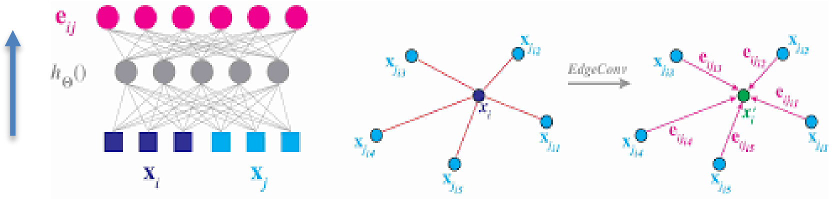
adding Adversarial NN





Anomaly detection

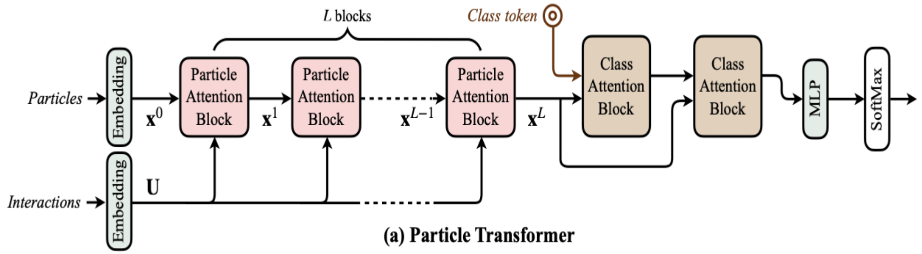
- Much progress in recent years developing powerful architectures for supervised DL tasks, such as
  - **ParticleNet (PNet)** Graphs that enhance the correlations among closest neighbours (pairwise features)



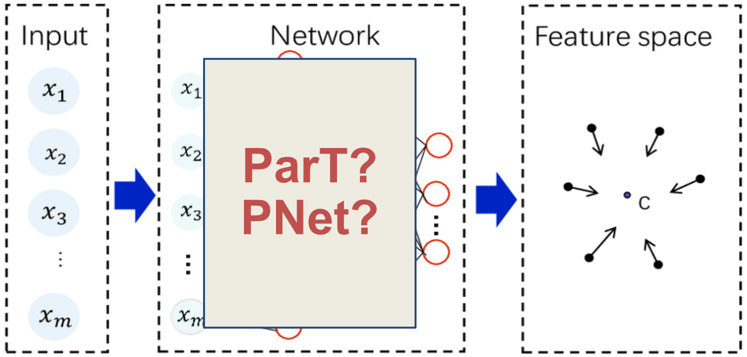
  - **Particle Transformer (ParT)** The self-attention mechanism allows to focus on the relevant correlations among the different objects of the event

Four-momentum

Higher-level features m_{ij}^2, ΔR



- Support Vector Data Description technique (SVDD) can be used to adapt any classifier into an anomaly detector
  - Add an output layer (output space)
  - Take Loss function as distance to a center in the output space
  - Background events get closer to the center



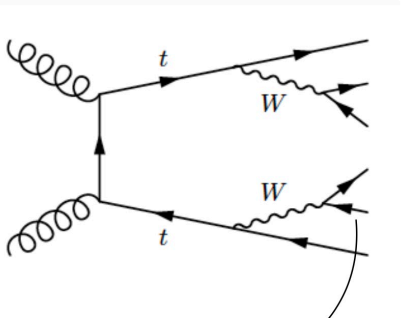
- Are the best-performing classifiers also the best anomaly/outliers detectors?
  - Trained only with "background" events (unsupervised)
- ***DarkMachines collab.*** already explored anomaly detection in the HEP context
  - PNet and ParT not explored yet for exotic searches
- Use pp collisions at 13 TeV
  - Detector simulation performed with Delphes 3 using a simplified ATLAS detector card
  - **Input variables**: four-momentum and type of objects, and the missing transverse momentum of the event

Two other techniques are also being explored to adapt any classifier, known as *DROCC* and *Smearing*

Work still in progress but results look promising

Generative Models for simulation

- Classical simulation of proton-proton collisions implies:
  - PS generation
  - Hadronization/fragmentation
  - Pass the particles through the detector
- Time consuming and expensive
  - Last above step is particularly expensive (eg. dense materials)
  - Billions of events
- Alternative are high fidelity fast generative models, eg. GANs, VAEs & NFs
  - Able to sample high dimensional feature distributions by learning from existing samples, eg. classical simulation
  - Generate SM background and BSM physics scenario and process the data in a easy-format (sequence of 4 – vectors)
  - Metrics to asses performance & syst. errors to be defined
- Use case:
  - pp→ttbar with 6 jets in the final state



Datasets used in this work have been taken from a uptodate repository; the ones generated by DarkMachines community. LHCsimulationProject, Feb 2020, doi:10.5281/zenodo.3685861. Available here

Loss Function

Variational AutoENcoder (VAE)

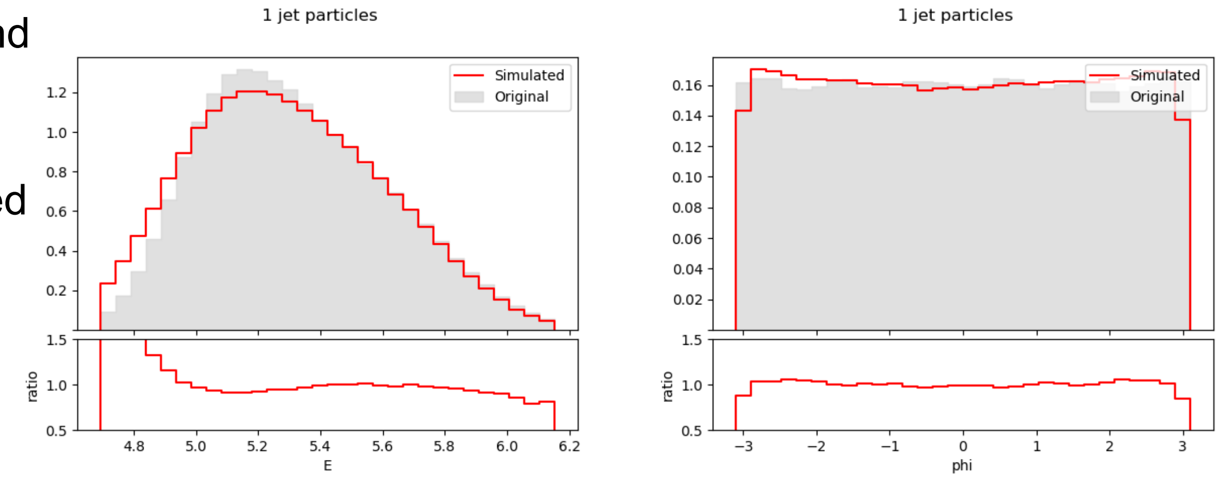
$$L_{VAE} = (1-\beta)MSE + \beta KL$$

MSE: Mean Squared Error. Reconstruction term on the Final layer, which tends to improve The performance of the encoding-decoding schema

a regularisation term on the latent layer, that is proportional to the Kullback-Leibler (KL) divergence and tends to regularise the organisation of the latent space by making the distributions returned by the encoder close to a standard normal distribution with zero mean and unit variance

To avoid Overfitting

Results with β-VAE

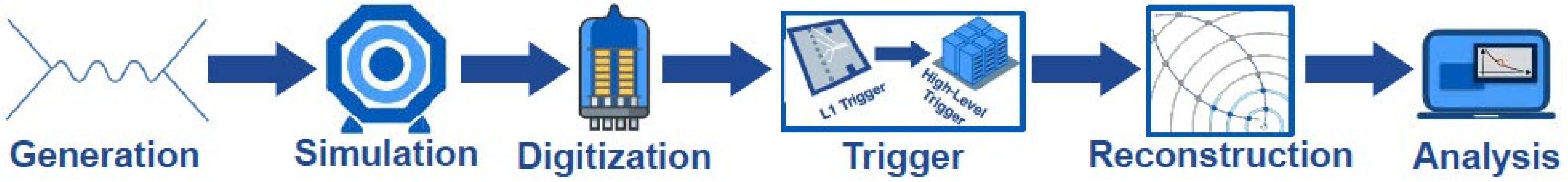


CHEP2023

E, ϕ of first jet using β-VAE with β=0.001

Conclusion

- ML/DL is a fundamental tool **everywhere** in today's (and future) HEP experiments
- Discussed a wide range of applications at ATLAS and LHCb experiments within the framework of the ASFAE projects



Generation:
- Integration
- Sampling
- Showering/ hadronization
- Parameter tuning
- Syst. uncertainties

Simulation:
- Generative (GANs, VAEs,…)
- Refinement
- Tuning
- End-to-end
- …

Digitization:
- Pileup
- Electronics modeling
- …

Trigger:
- Filtering & classification
- Low-latency
- Anomaly detection
- On-detector
- …

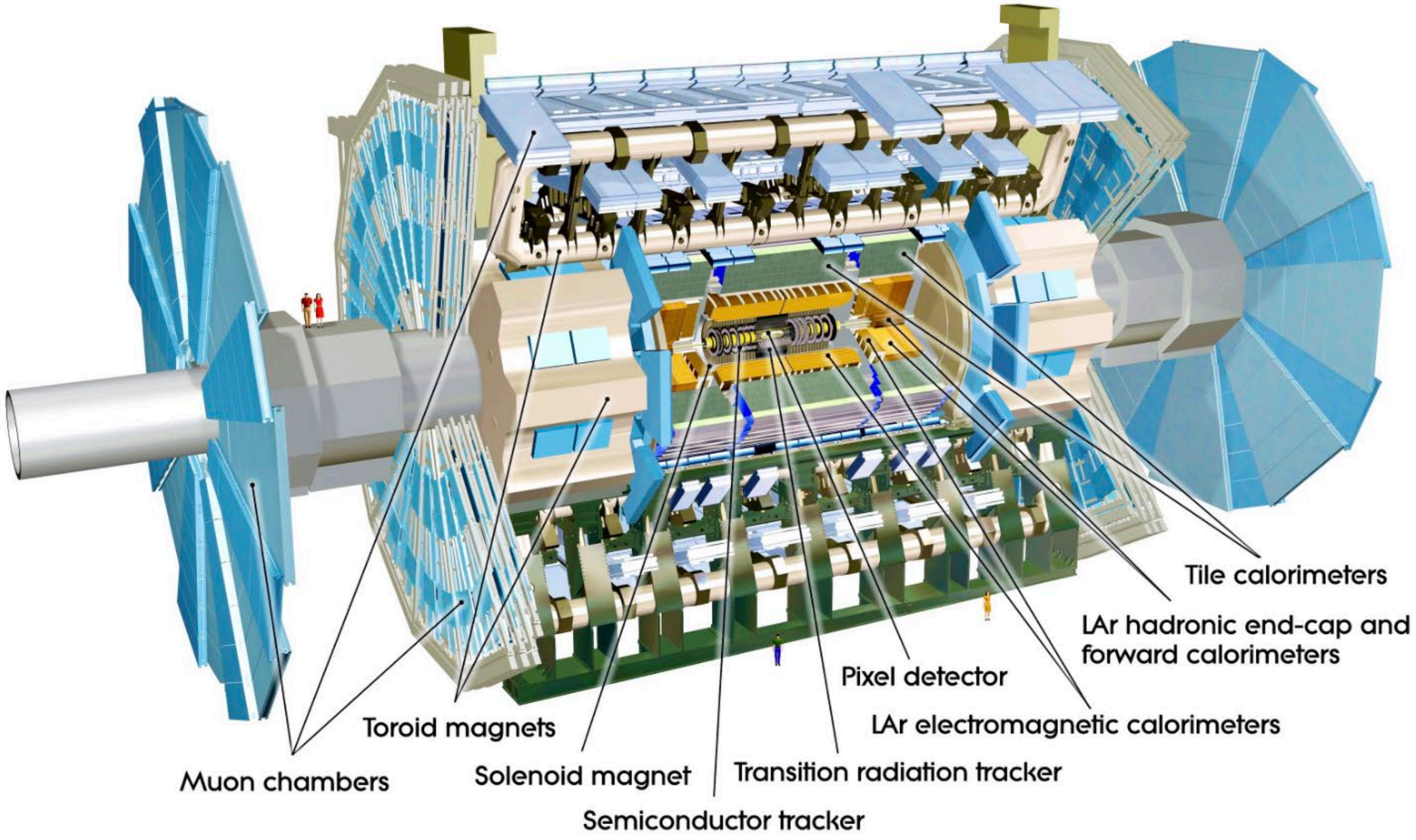
Reconstruction:
- Calibration
- Tracking
- Clustering
- …

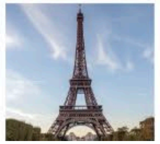
Analysis:
- Classification
- Regression
- Anomaly detection
- Simulation-based inference
- …

- Of critical importance for fully exploiting the physics potential of LHC during Run 3 and beyond
  - Both for probing with increasing precision the SM and searching for BSM physics
  - ATLAS @ IFIC: ttbarW production & ch. asym., mono-top, ttbar resonances, search for BSM LLPs, $H\rightarrow\tau\tau$, di$H\rightarrow bb\gamma\gamma$
  - LHCb @ IFIC: Radiative b-hadron decays, Λ MDM/EDM, Charm baryon decays, BSM LLPs, Penta & Sexaquarks

Backup

ATLAS experiment



25 m diameter, 46 m long, 7000 tons weight ~ 

Axial field provided by **solenoid** (2 T) in central region (momentum measurement)

High resolution silicon **detectors**:

- 100 Mio. channels (50 μm x 250 μm)
- 6 Mio. channels (80 μm x 12 cm)

spatial resolution ~15 μm (in azimuthal direction)

Energy measurement down to 1º to the beam line with a **calorimeter system**

Independent **muon spectrometer** (superconducting toroid system)

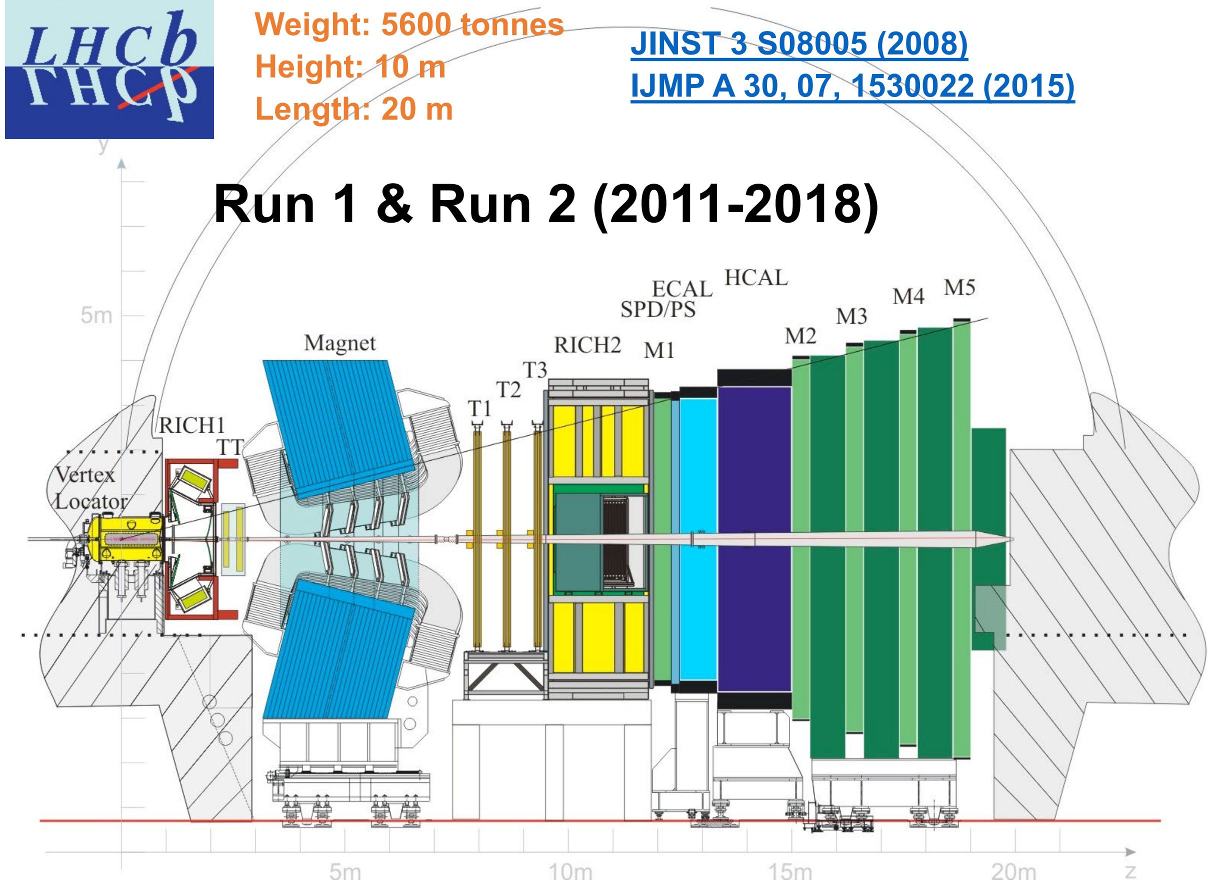
Ultra-fast custom electronics filters the collisions: only 1 out of 30,000 bunch collisions is kept!

LHCb experiment

Weight: 5600 tonnes
Height: 10 m
Length: 20 m

JINST 3 S08005 (2008)
IJMP A 30, 07, 1530022 (2015)

Run 1 & Run 2 (2011-2018)



- Forward spectrometer, optimized for the study of of b and c hadron decays
  - 2 < η < 5 acceptance
  - Huge b and c hadron production

$$\sigma(pp \to b\bar{b}X)_{2<\eta<5} \approx 144\ \mu\text{b}$$

$$\sigma(pp \to c\bar{c}X)_{p_T<8\ \text{GeV},2.0<y<4.5} \approx 2400\ \mu\text{b}$$

  - Integrated 10 fb^{-1}
  - Excellent vertexing, tracking, momentum resolution < 1%, and PID (K/π/p/μ/e/γ)

- Wide physics program
  - Mixing and CP violation in B decays
  - Rare B/D/K decays
  - Charm decays
  - Semileptonic B decays
  - Spectroscopy and exotic hadrons
  - Hadron production
  - Heavy ion physics, fixed target with SMOG
  - Electroweak physics, QCD
  - Exotics (dark matter, long-lived particles)

LHCb Upgrades timeline

LHCb —— LHCb Upgrade I —— LHCb Upgrade II

Run 1 - Run 2		Run 3		Run 4		Run 5		Run 6
L_{int} = 10 fb^{-1} L = 4 x 10^{32} cm^{-2} s^{-1}	LS2 Injector Upgrades LHCb Upgrade I	L = 2 x 10^{33} cm^{-2} s^{-1}	LS3 HL-LHC ATLAS/CMS Phase 2 Upgrades LHCb Upgrade Ib (Consolidation)	L_{int} ~ 50 fb^{-1}	LS4 LHCb Upgrade II	L = 1-2 x 10^{34} cm^{-2} s^{-1}	LS5	L_{int} ~ 300 fb^{-1}
2010 - 2018	2019 - 2022	2022 - 2025	2026 - 2028	2029 - 2032	2033 - 2034	2035 - 2038	2039	2040 – 20XX

Increased Lumi by x 5
Limited to ~ 6 int. per bxing

Limited to ~ 6 int. per bxing

~ 40 int. per bxing

~ 40 int. per bxing

HL-LHC

- Upgrade I – Run 3
  - Increase the luminosity from 4x10^{32} $cm^{-2}s^{-1}$ to 2x10^{33} $cm^{-2}s^{-1}$
  - Detectors and electronics upgrades needed
  - Trigger and DAQ redefined
- Consolidation/enhancement phase in LS3
  - First stage of Upgrade II “Upgrade Ib”
  - No luminosity change (baseline)
- Main installation phase in LS4
  - Full Upgrade II (luminosity increase)