RAISE All-Hands Meeting in Iceland
Task 4.1:
Event reconstruction and classification at the CERN HL-LHC

August 29th, 2023

Eric Wulff [1], Maria Girone[1], **David Southwick[1]**, Juan Pablo García Amboage[1], Arnis Lektauers[2]

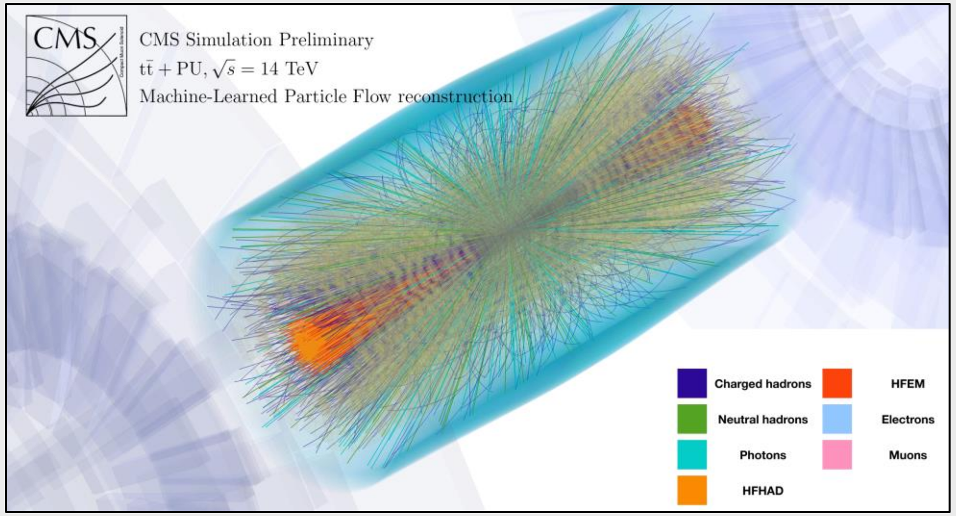
[1]CERN

[2]RTU

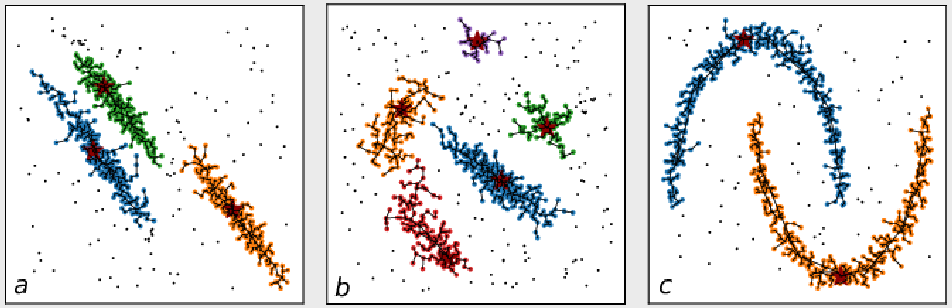
Introduction

- Task 4.1: Event reconstruction and classification at the CERN HL-LHC

- We work on four different but related efforts
  - AI–based particle flow reconstruction algorithm, *MLPF (Machine-learned Particle-Flow)* [1] (in collaboration with CMS)
  - Traditional GPU-accelerated clustering algorithm, *CLUE (CLUstering of Energy)* (in collaboration with CMS)
  - HW benchmarking based on AI-models for HEP
  - Data challenges for the Exascale era



Using MLPF in CMSSW.



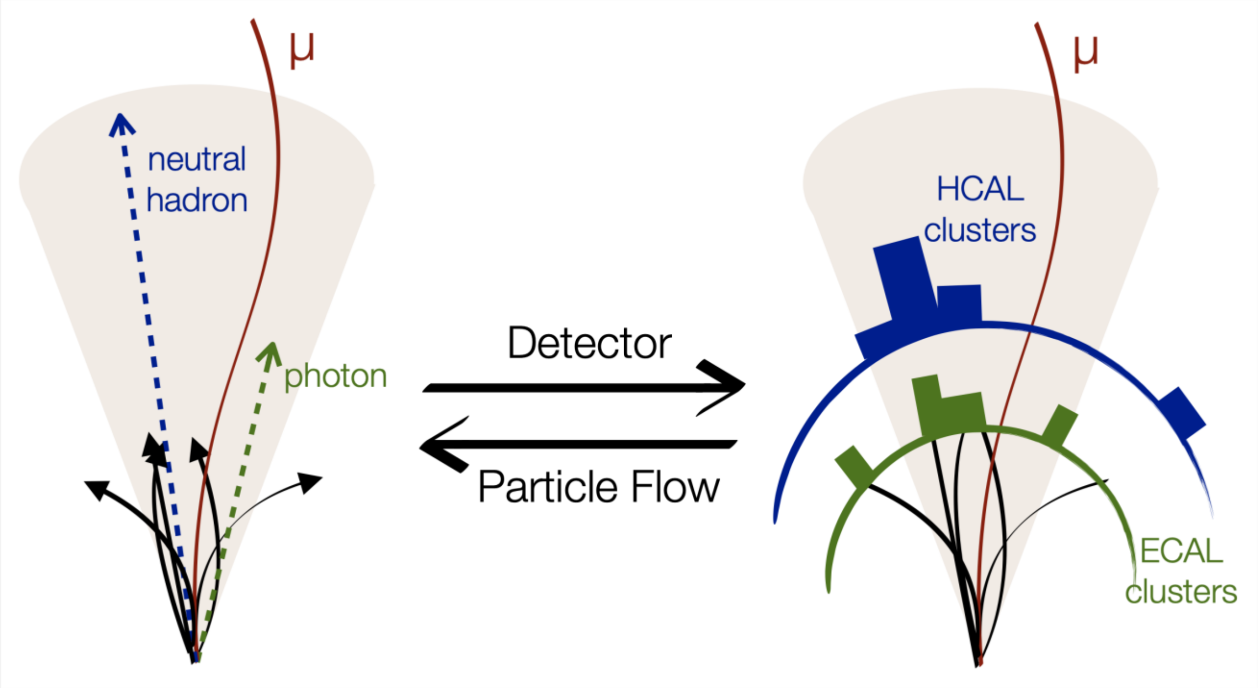
CLUE clustering on synthetic data. [2]

[1] J. Pata et al. Eur. Phys. J. C (2021) 81: 381 (https://link.springer.com/article/10.1140/epjc/s10052-021-09158-w)
[2] https://doi.org/10.3389/fdata.2020.591315

Event reconstruction using MLPF

What is event reconstruction?

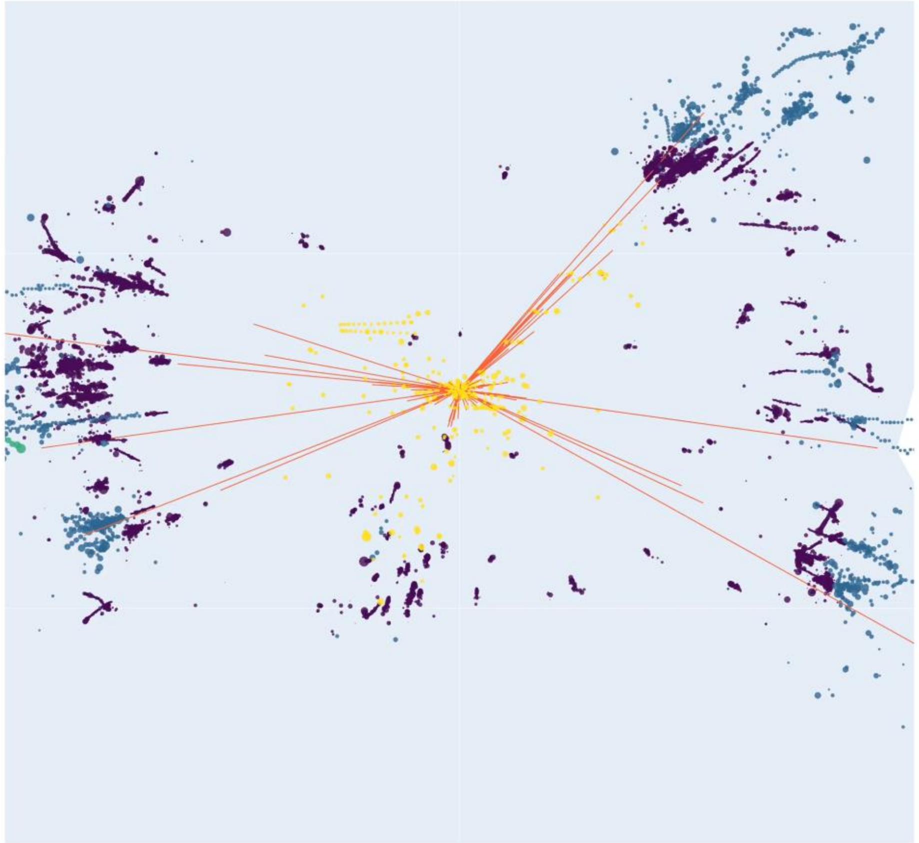
- Event reconstruction attempts to solve the inverse problem of particle-detector interactions, i.e., going from detector signals back to the particles that gave rise to them
- Particle-flow (PF) reconstruction takes tracks and clusters of energy deposits as input and gives particle types and momenta as output



New CoE RAISE Open Data

- https://www.coe-raise.eu/od-pfr
- An extensive open dataset of physics events with full GEANT4 [1] simulation, suitable for PF reconstruction, available in the EDM4HEP [2] format
- ~2.5 TB before pre-processing
- The dataset contains
  - Reconstructed tracks, calorimeter hits and clusters
    - We use these as inputs
  - All generator particles
    - We use these as targets
  - Reconstructed particles by the Pandora algorithm [3,4,5]
    - We use these as a baseline for comparison
- A mixture of $t\bar{t}$, $q\bar{q}$, ZH and WW events

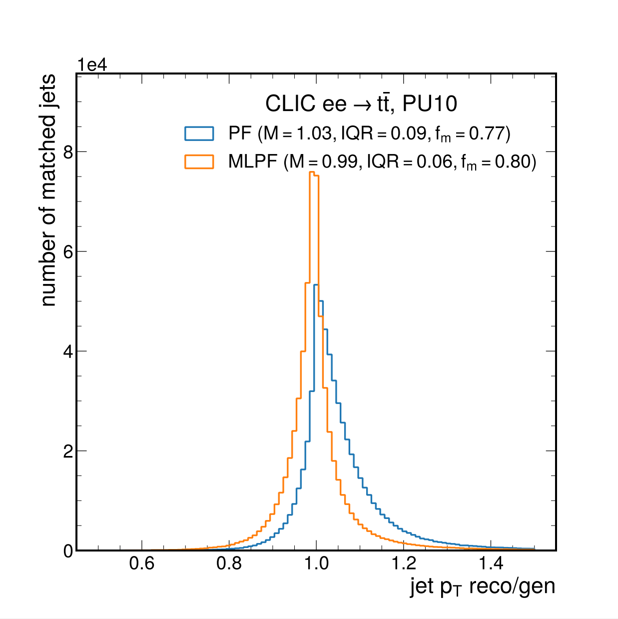
3D visualization of a single event

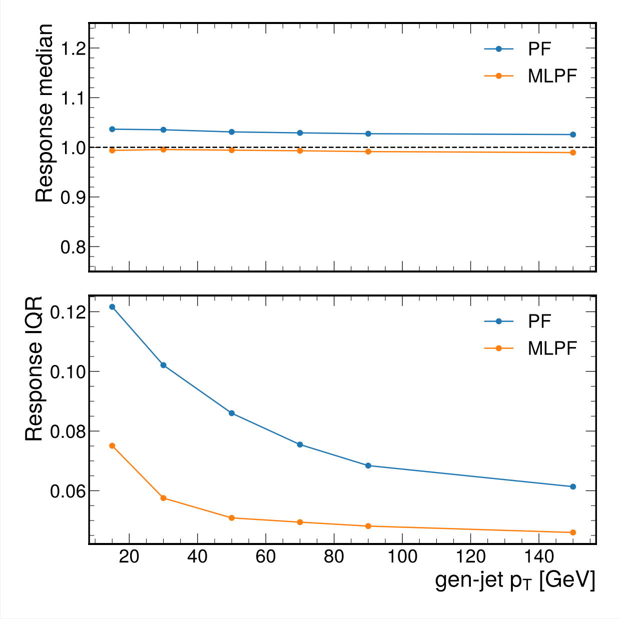


Jet and MET in ttbar + PU10 test data

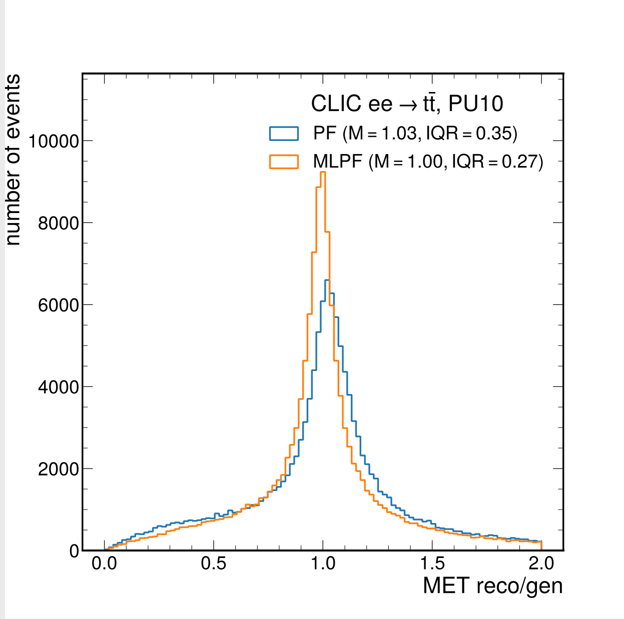
- For all test samples MLPF outperforms PF in Jet and MET reconstruction in terms of response width (quantified by median and interquartile range (IQR))
- MLPF also outperforms PF in terms of fraction of reconstructed jets $\left({}^{n_{reco\ jets}}/_{n_{ground-truth\ jets}}\right)$
- Very similar results are seen in ZH and WW events

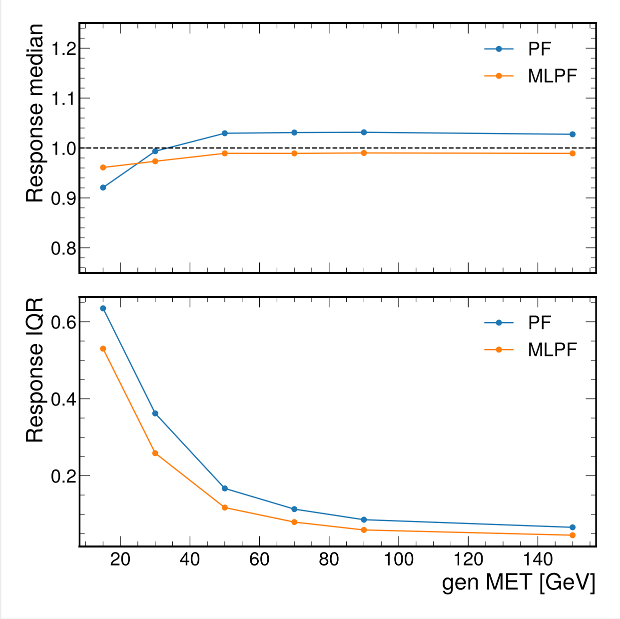
Jets resolution





MET resolution

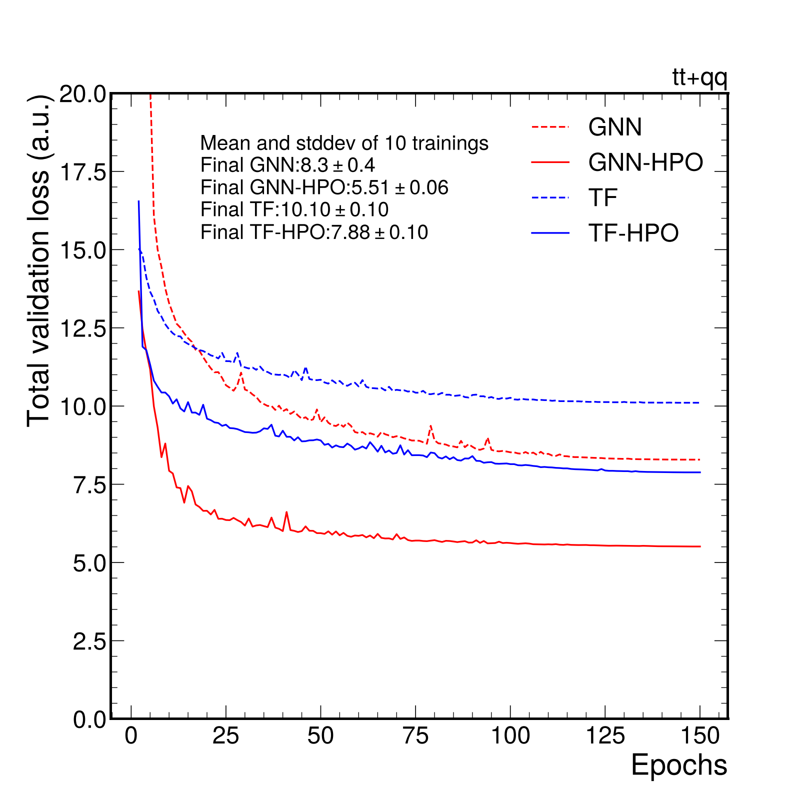




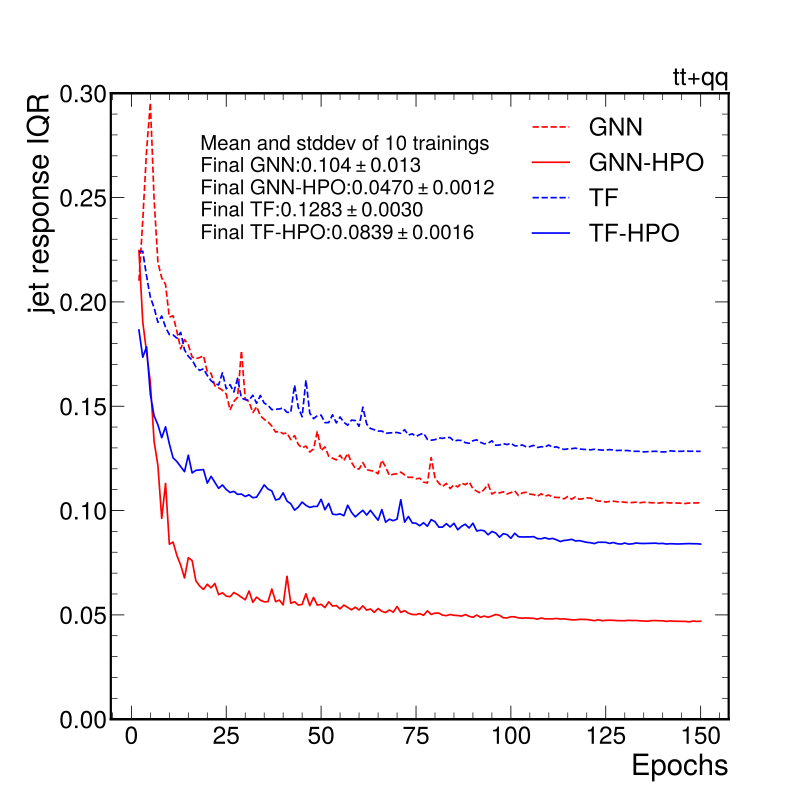
Improvement in training from HPO

- HPO significantly improved model performance for both the GNN-based and the transformer-based MLPF models
- GNN outperforms transformer

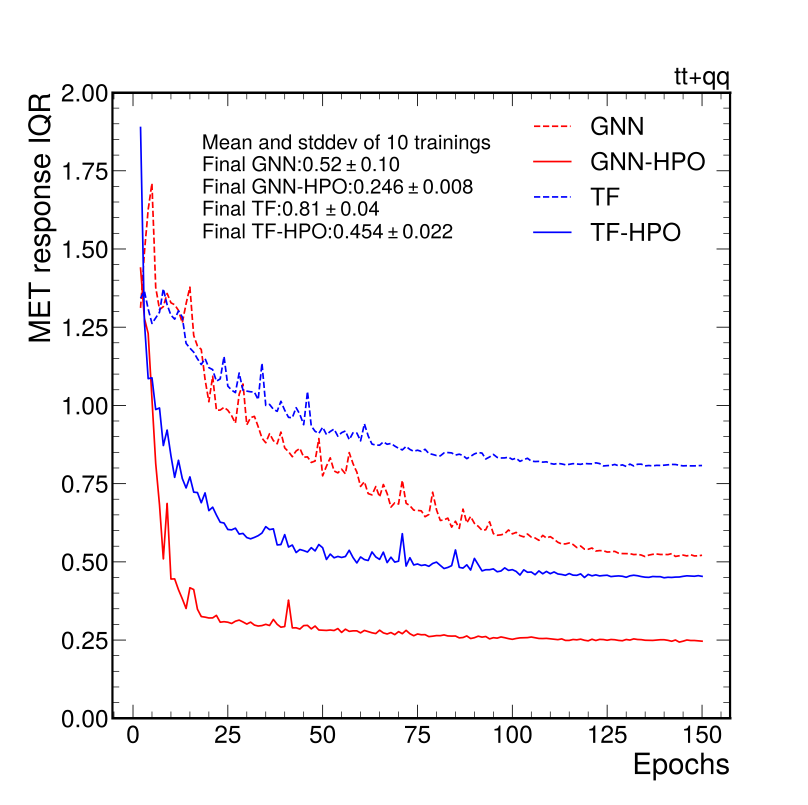
Validation loss



Jet resolution



MET resolution

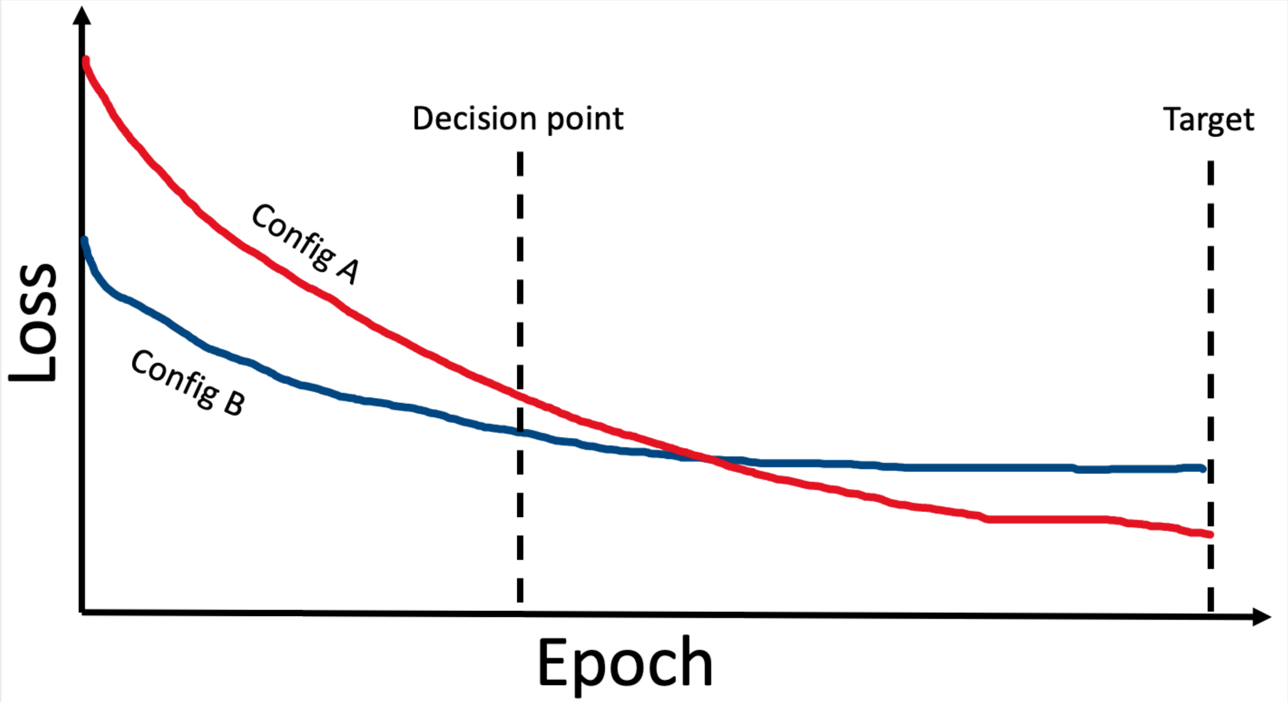


QSVR for model performance prediction

Hypertuning workflow

- Current STOTA hypertuning algorithms rely on early stopping
- Stopping criterion: ranking according to a single metric (e.g., validation loss)
- Potential problem: loss curves are not linear
- Idea 1: Use a non-linear stopping criterion
  - For instance, an SVR model, inspired by [1]
- Idea 2: Use quantum computing to fit Quantum-SVR (QSVR)

Sample random configs → Train n epochs → Terminate worst x% at last epoch



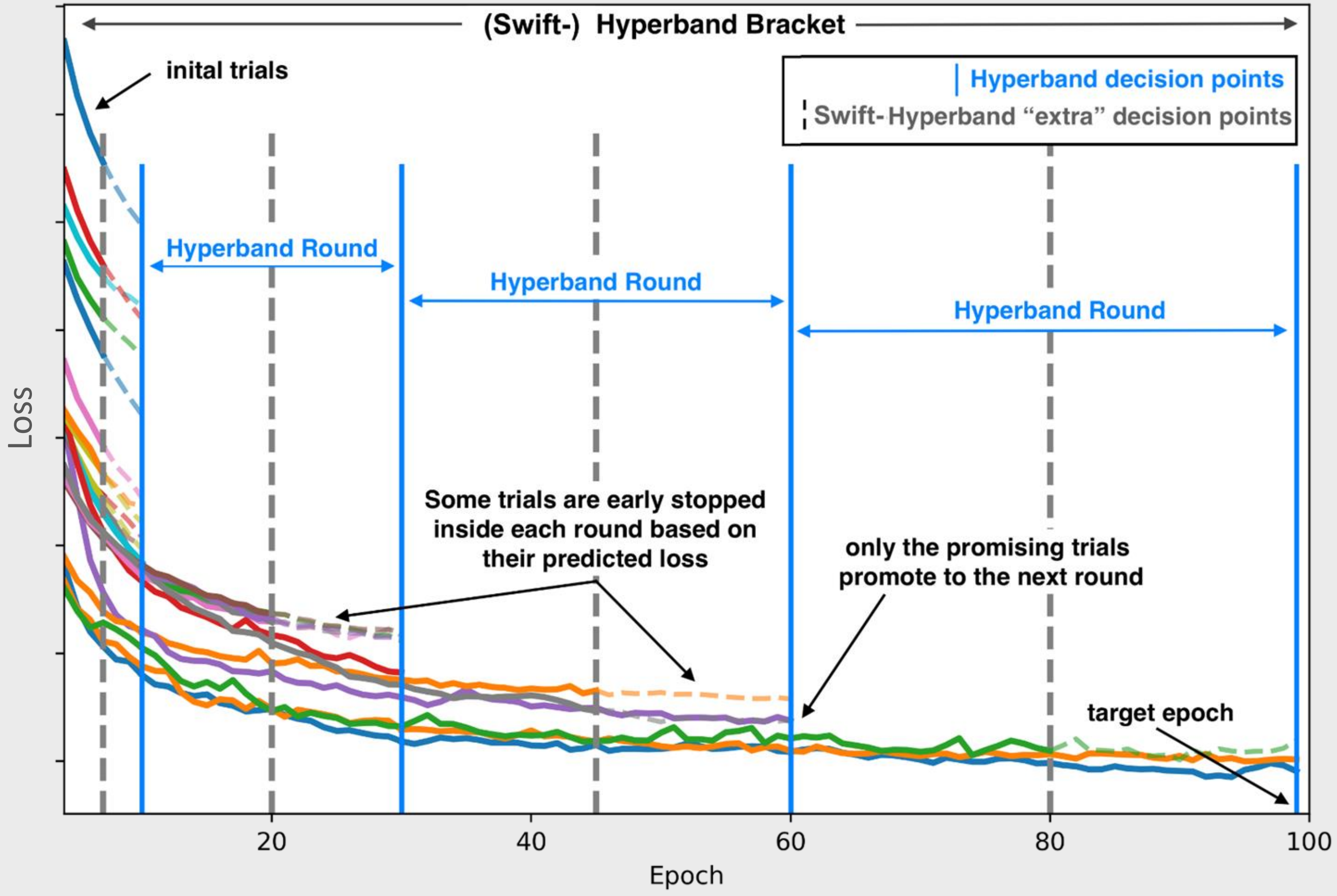
Swift-Hyperband

- Fast-Hyperband is not suitable for use with Quantum-SVRs
- Introducing: *Swift-Hyperband* – a new approach to combine performance prediction with Hyperband

Fast-Hyperband		Swift-Hyperband	
Multiple decision points inside each round	Trains **many** SVRs	Only 1 decision point inside each round	Trains **few** SVRs
Estimates σ for every SVR	↓	No need to estimate σ	↓
Sequential	**Not suitable for Q-SVRs**	**Easily parallelizable**	**Suitable for Q-SVRs**

Credit to Juan Pablo García Amboage, CERN Technical Student

Swift-Hyperband

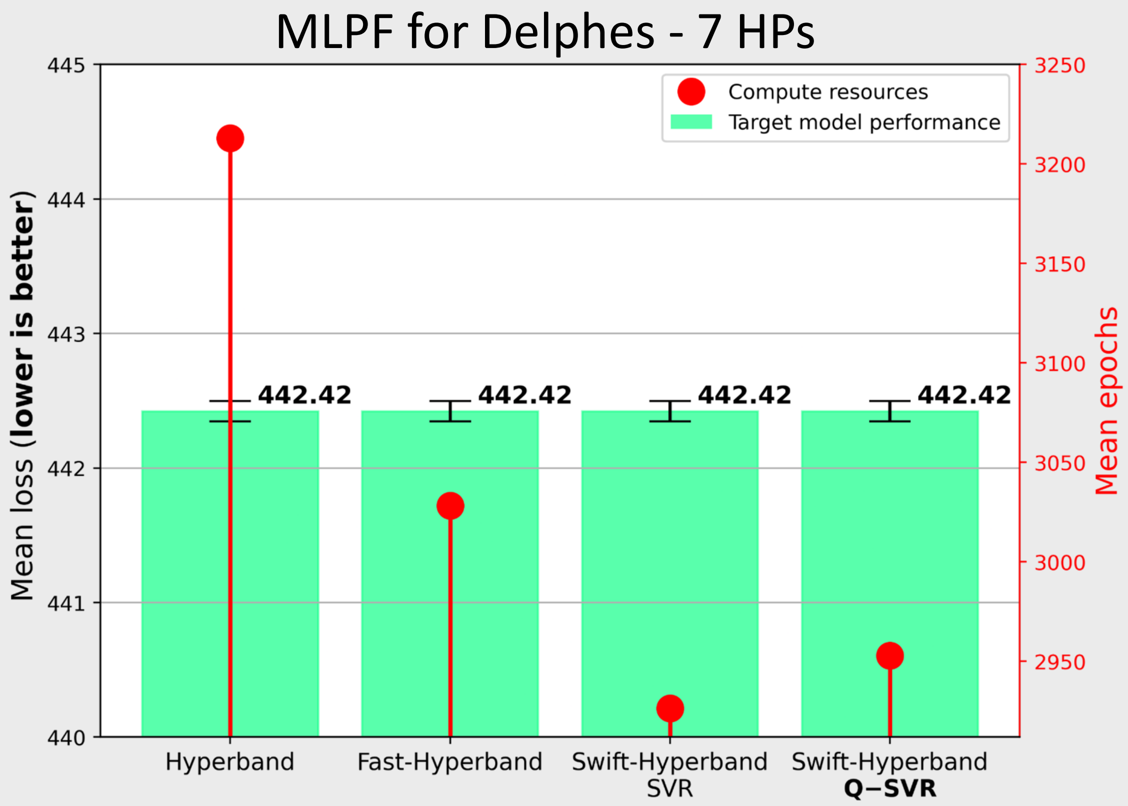


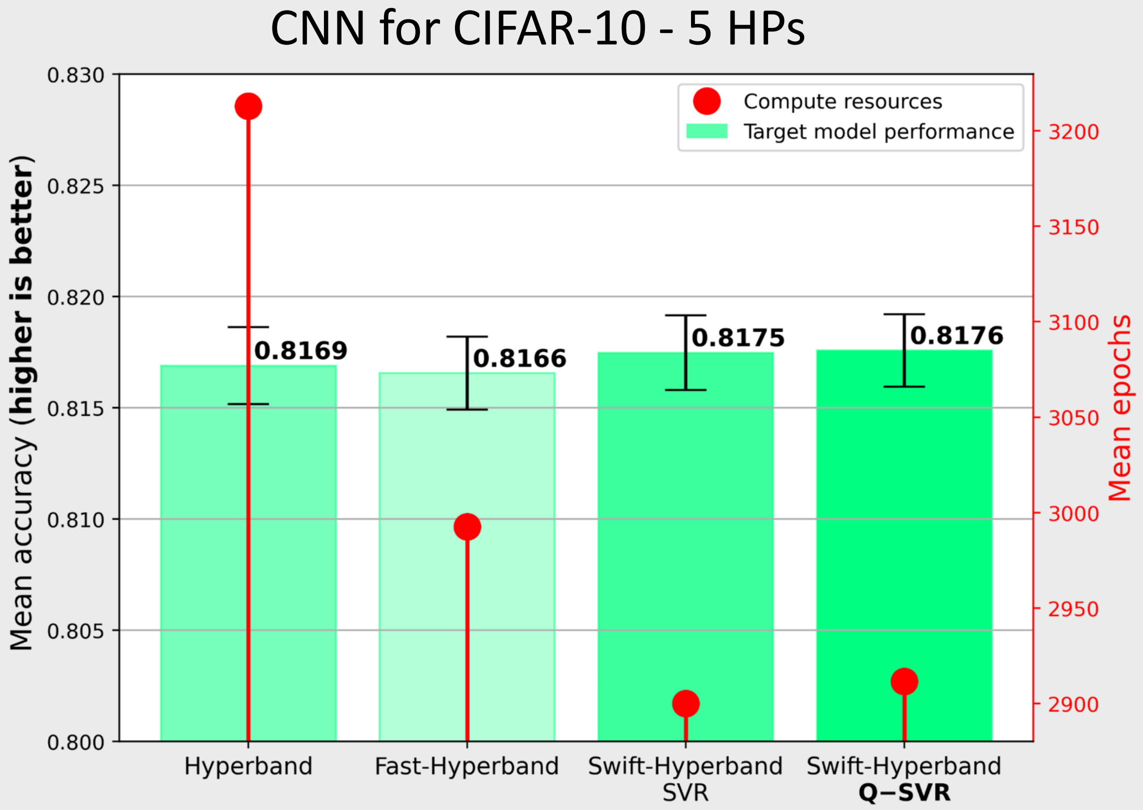
- One extra decision in each round (dashed vertical lines)
- First, some trials are fully trained to define a threshold and to fit (Q)-SVR performance predictors
- Other trials are then partially trained
- If their predicted performance is worse than the threshold, they are stopped immediately

Algorithm Comparison

RAISE Center of Excellence

- Simulated results using datasets of existing learning curves

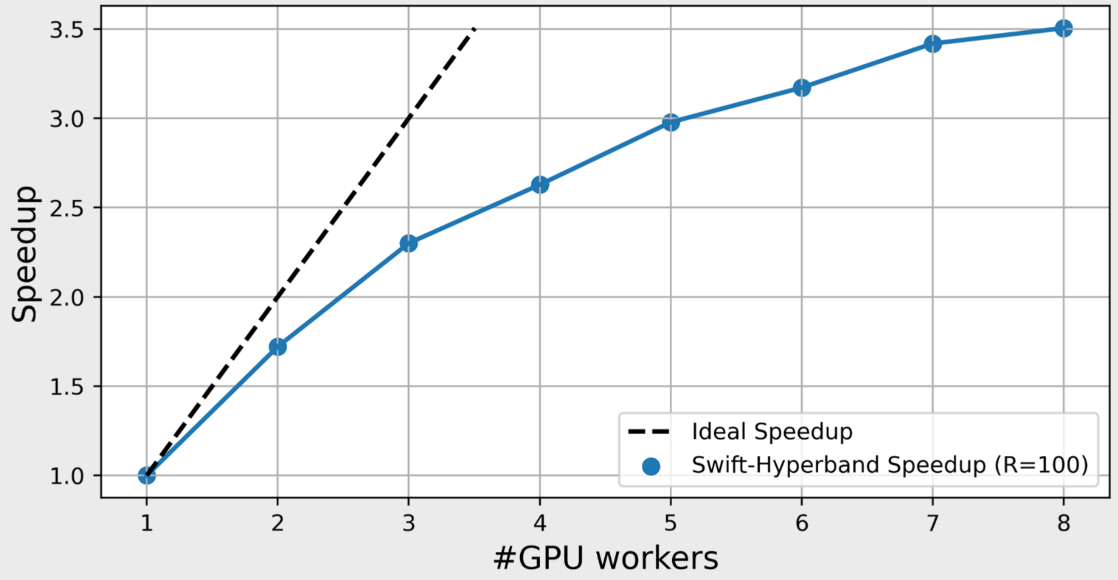




Credit to Juan Pablo García Amboage, CERN Technical Student

Distributed Swift-Hyperband

- Run on the DEEP-EST Extreme Scale Booster at FZJ
- Head node coordinates the workflow (1 CPU)
- Multiple GPU worker nodes for training trials
- Quantum Annealer for training the performance predictors
- Implemented using MPI and dwave-ocean-sdk

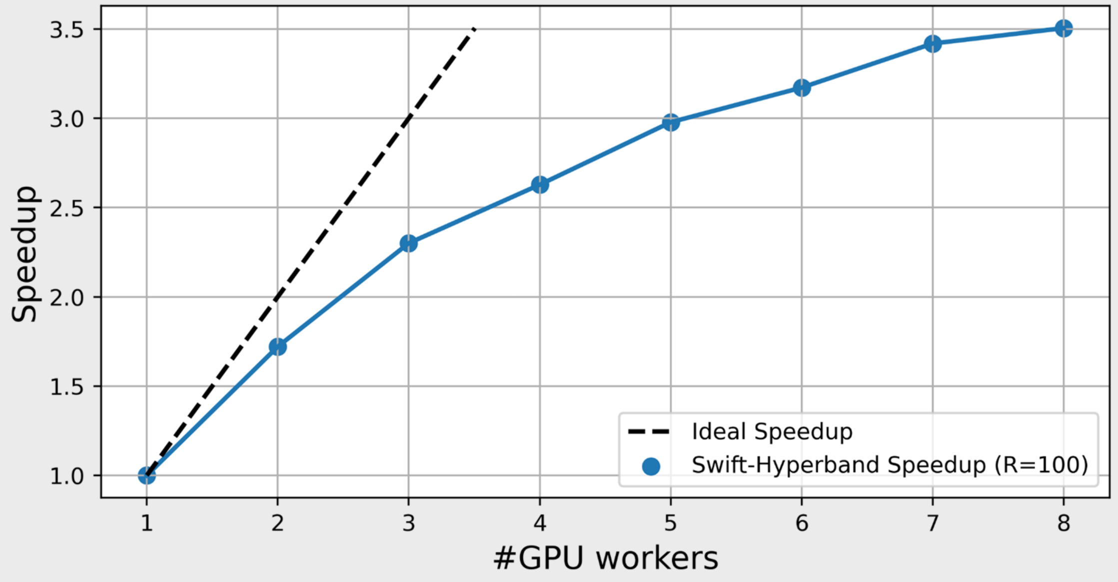


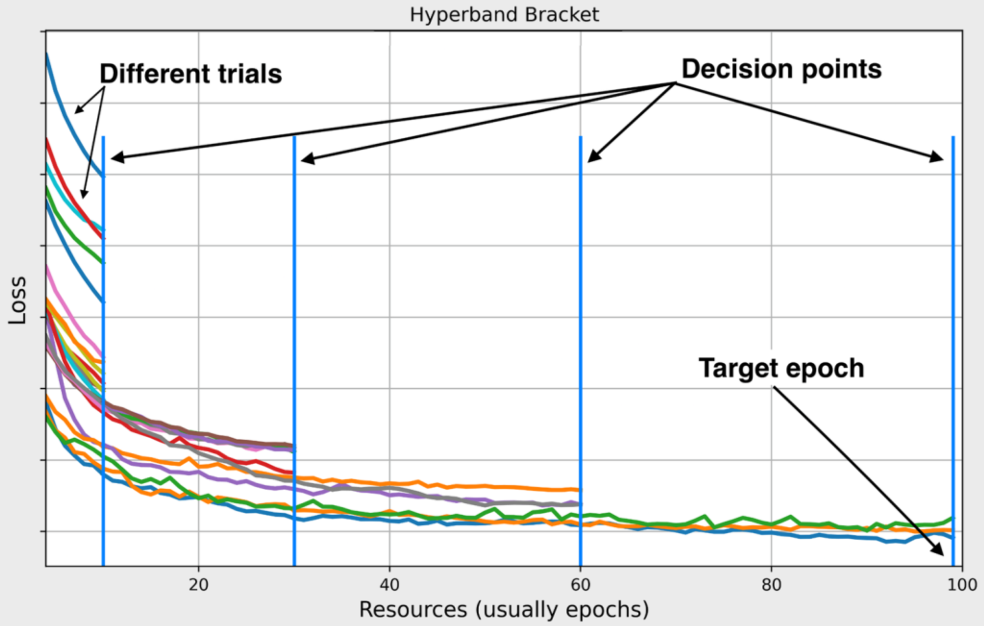
Credit to Juan Pablo García Amboage, CERN Technical Student

Distributed Swift-Hyperband

RAISE Center of Excellence

- Run on the DEEP-EST Extreme Scale Booster at FZJ
- Head node coordinates the workflow (1 CPU)
- Multiple GPU worker nodes for training trials
- Quantum Annealer for training the performance predictors
- Implemented using MPI and dwave-ocean-sdk

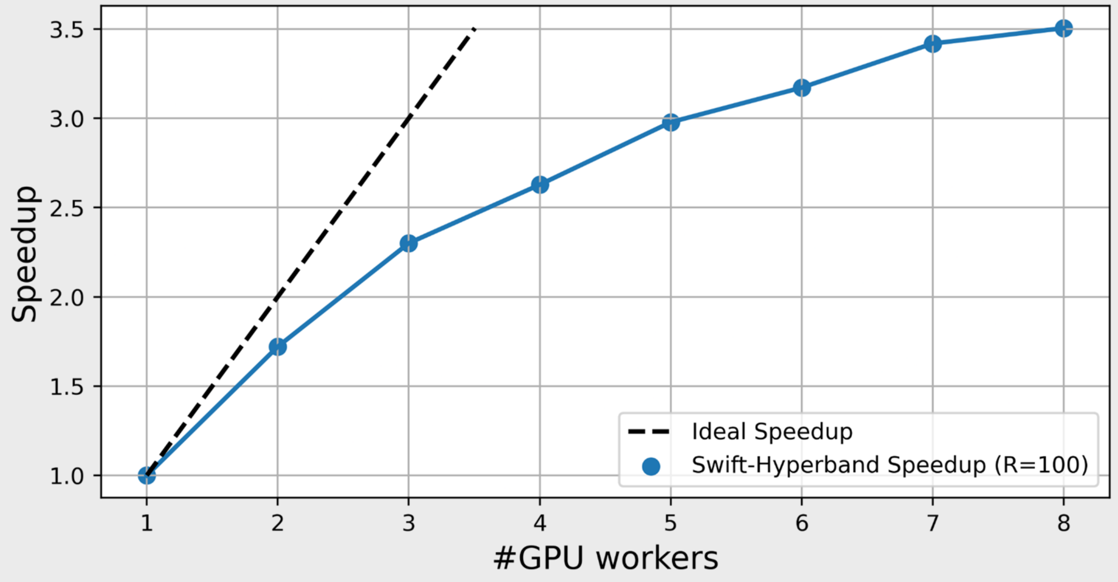


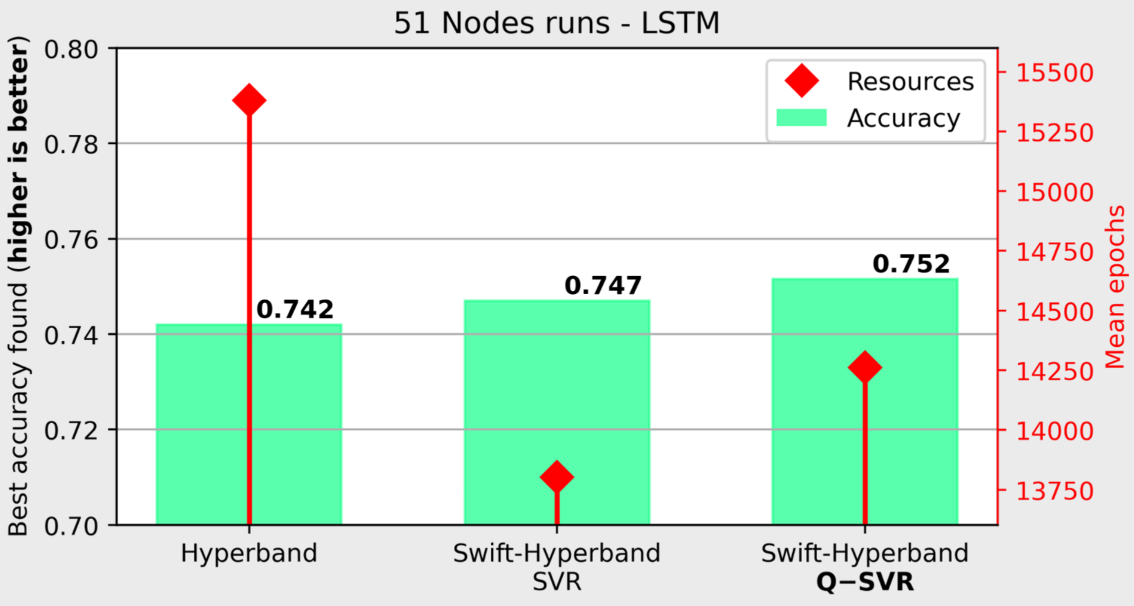


Credit to Juan Pablo García Amboage, CERN Technical Student

Distributed Swift-Hyperband

- Run on the DEEP-EST Extreme Scale Booster at FZJ
- Head node coordinates the workflow (1 CPU)
- Multiple GPU worker nodes for training trials
- Quantum Annealer for training the performance predictors
- Implemented using MPI and dwave-ocean-sdk



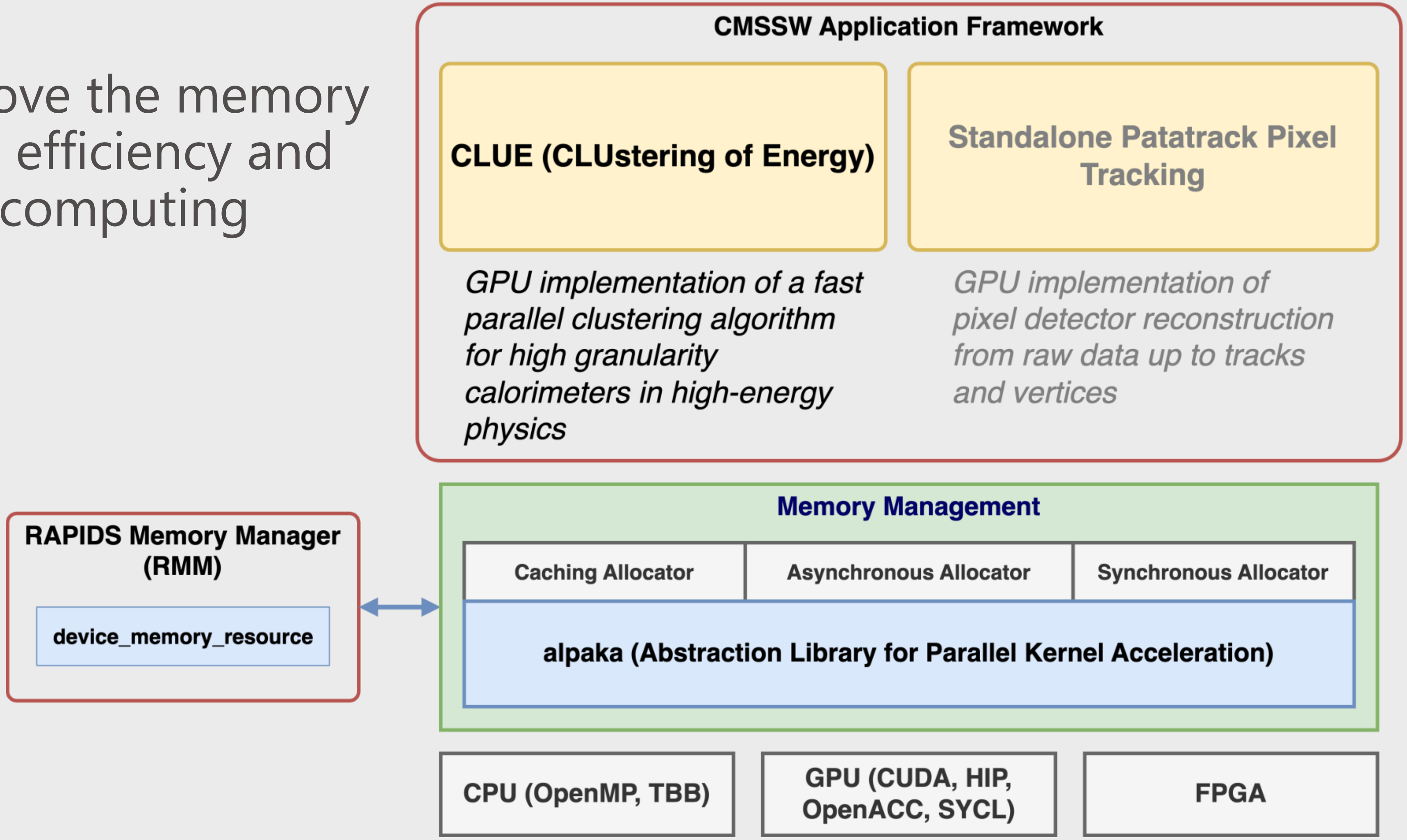


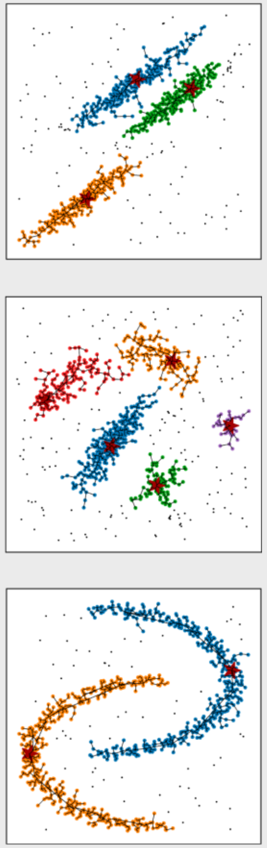
Credit to Juan Pablo García Amboage, CERN Technical Student

Performance tuning and generalization of CLUE

Heterogeneous CLUE Algorithm

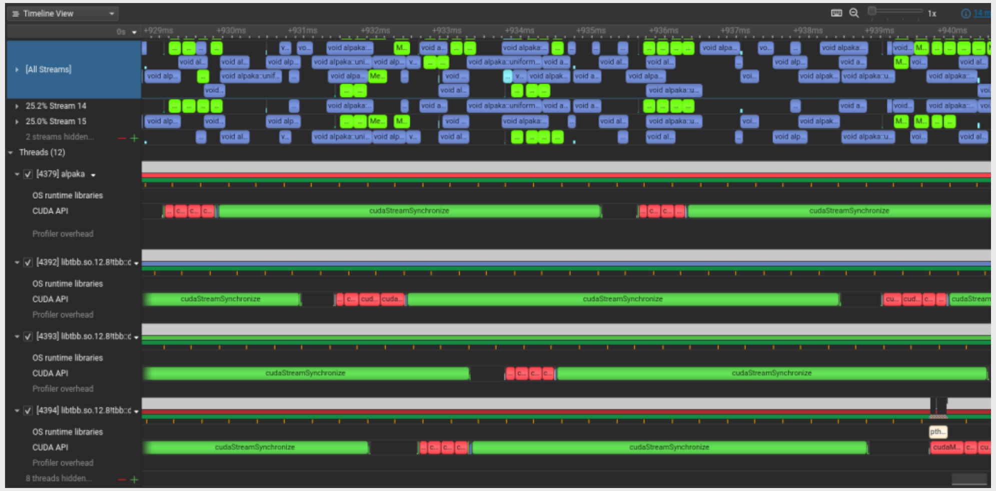
- Task: to improve the memory management efficiency and multi-thread computing performance

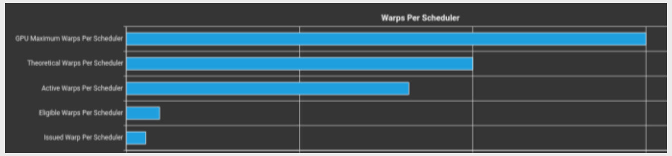




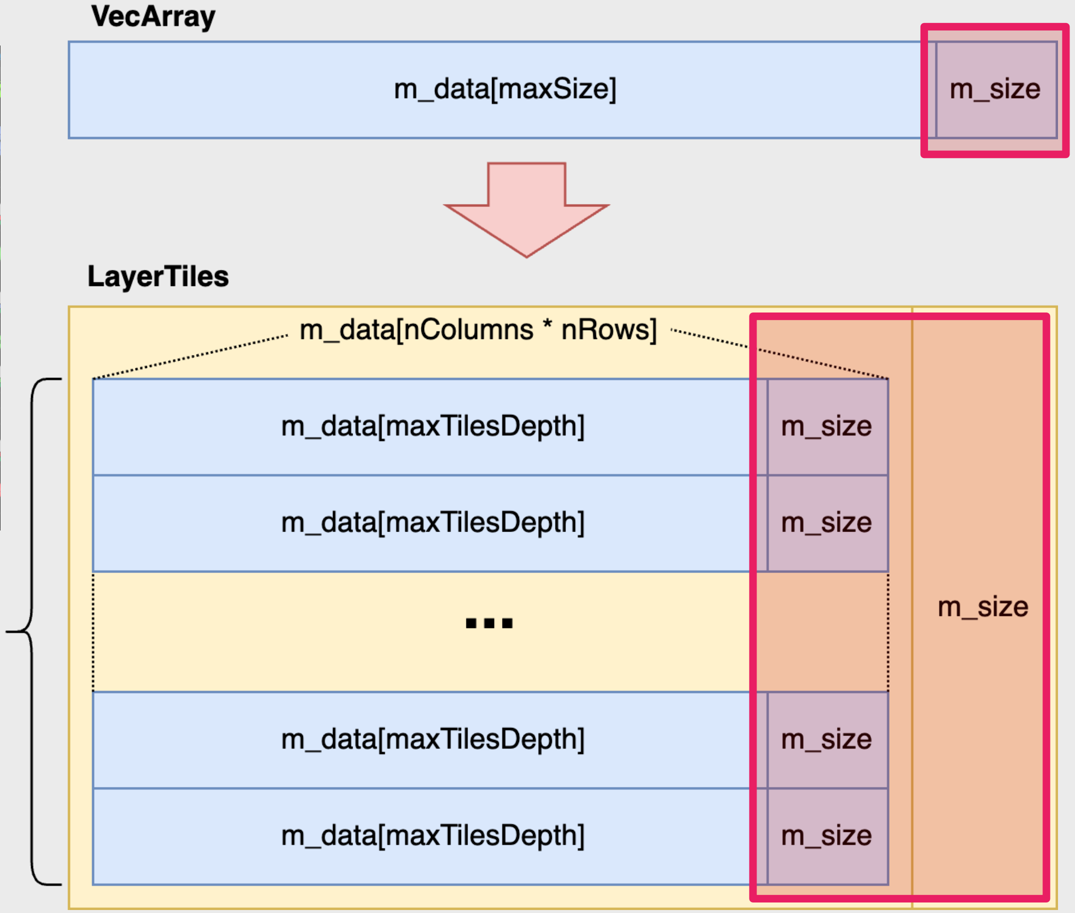
CLUE Performance Optimization

Application profiling with *Nsight Compute*





CLUE Data Structures

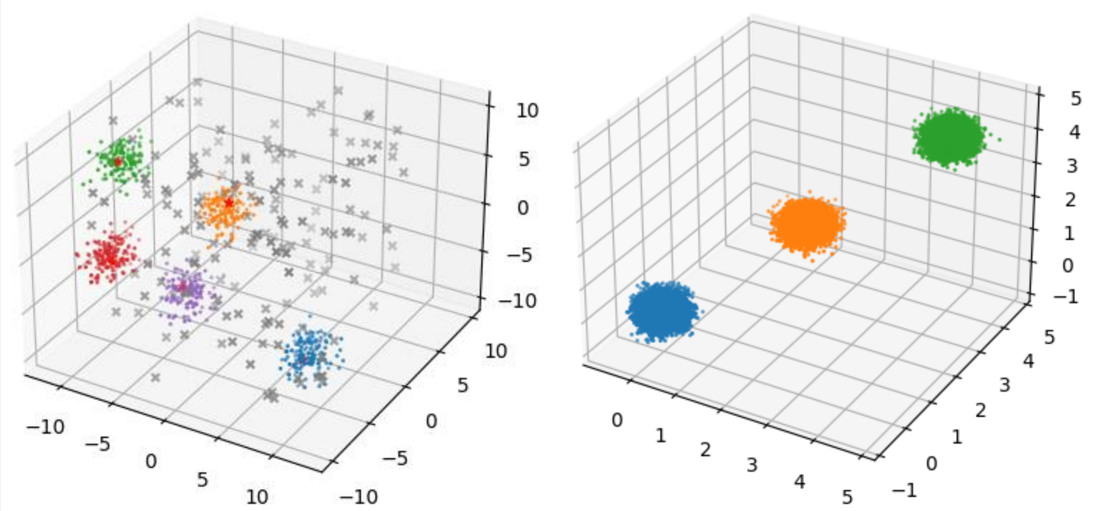


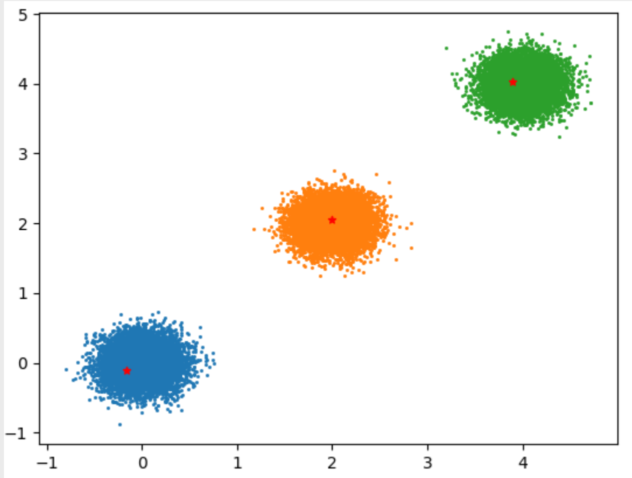
Memory coalescing and threadlocking issues

Solution: use <u>Structure of Arrays (SoA)</u> instead of *Array of Structures* (AoS)

N-Dimensional CLUE

- Generalization of the original algorithm, making it N-dimensional
  - The original algorithm was designed to work in 2 dimensions, with the data distributed in parallel layers
  - CLUE takes the coordinates of the points and their weight, which represents their energy, and calculates the energy density of each point
- Development of Python API





Summary

Conclusions and future work

- An extensive open dataset of physics events has been released on the CoE RAISE Open Data website
- MLPF outperforms PF in both particle level and event-level physics performance metrics
- Swift-Hyperband is a new HPO algorithm that integrates performance prediction with Hyperband
- Swift-Hyperband can run in a hybrid Quantum-Classical workflow manner
- Next steps:
  - Focus MLPF efforts on CMS datasets again
  - Together with WP2, submitted abstract to the *Quantum Technologies in Machine Learning* conference (QTML) which will take place at CERN 19th to 24th of November
  - Implement a Swift-ASHA with the aim of solving the straggler issues inherent to Hyperband

Bonus slide

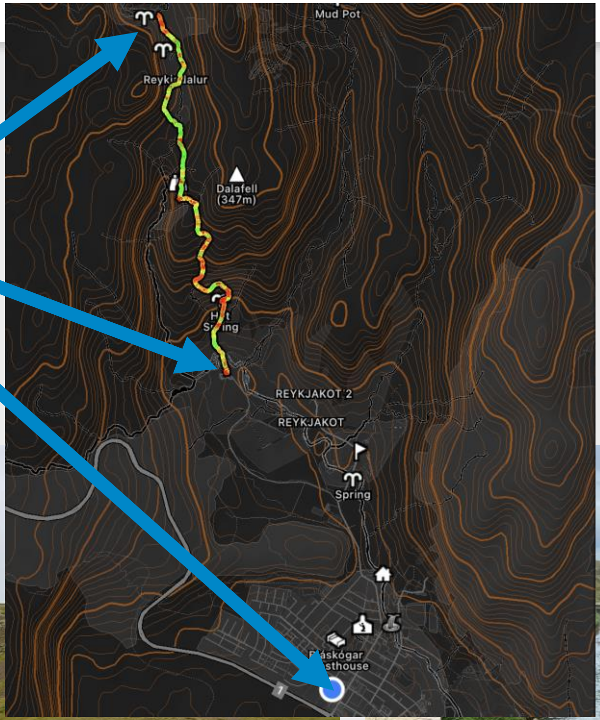
Hot spring hike

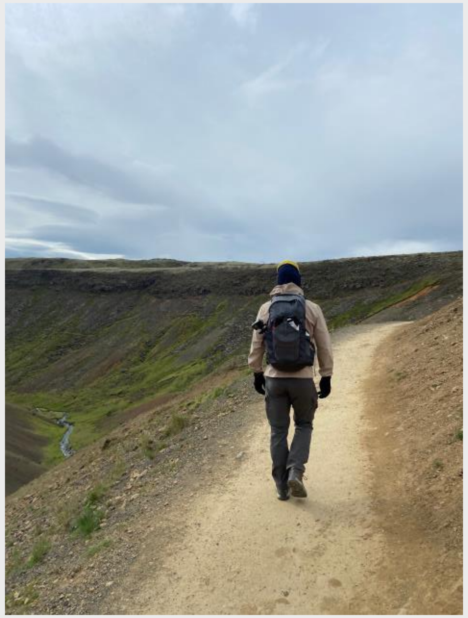
- https://www.alltrails.com/trail/iceland/southern/reykjadalur-hot-spring-thermal-river
- A car would be nice to get to the start of the hike
- Hike took ~2h15m for me, including a ~20 minutes bath in the river and stopping to take photos on the way

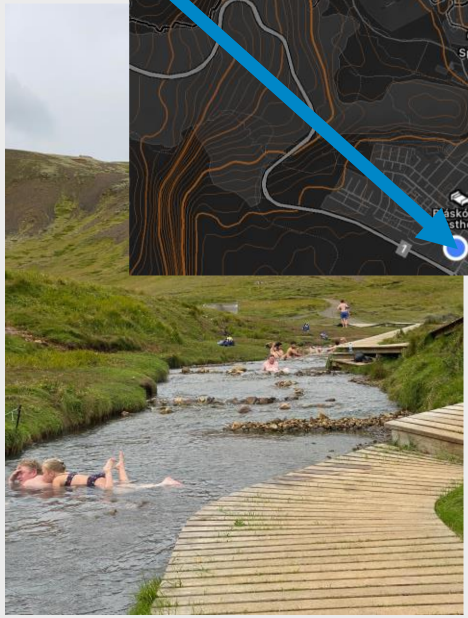
Hot spring

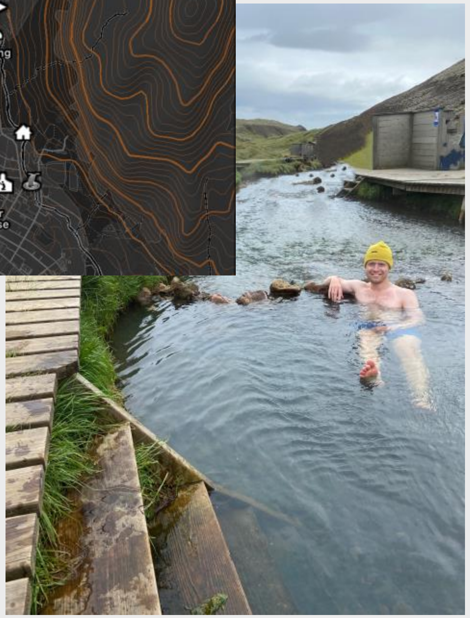
Start of hike

Hotel Örk









drive. enable. innovate.

RAISE
Center of Excellence

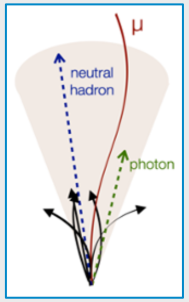
The CoE RAISE project has received funding from
the European Union’s Horizon 2020 –
Research and Innovation Framework Programme
H2020-INFRAEDI-2019-1 under grant agreement no. 951733

Follow us:

Backup Slides

Deep Learning for particle flow reconstruction

Physics simulation

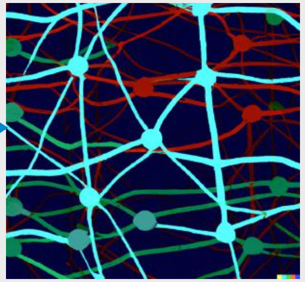


Data selection

Dataset creation

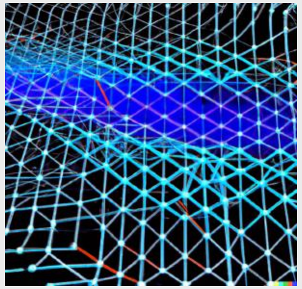
Data pre-processing

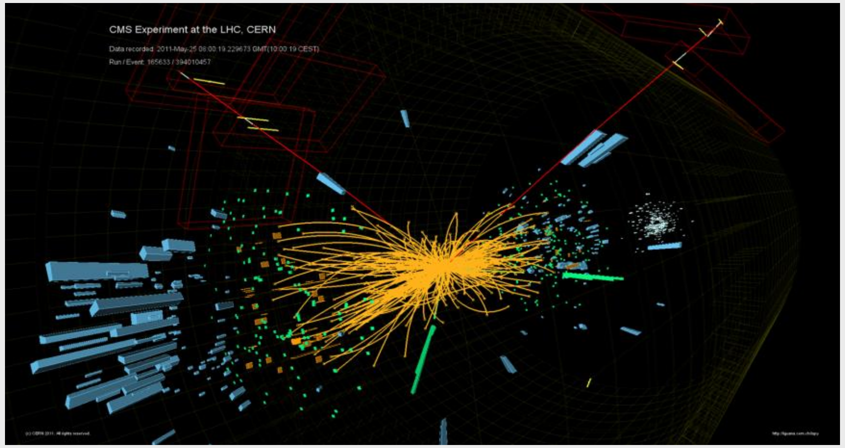
AI training



Model export

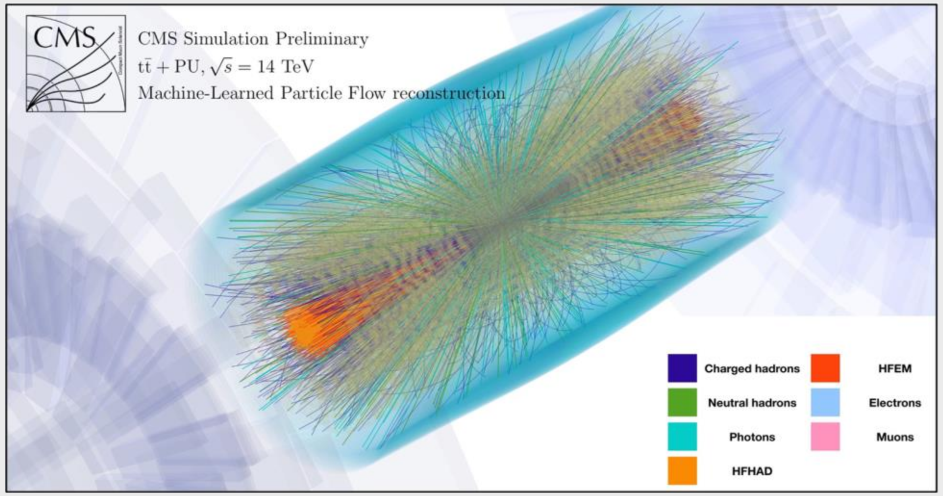
Trained model





CMS Collision event

Event reconstruction



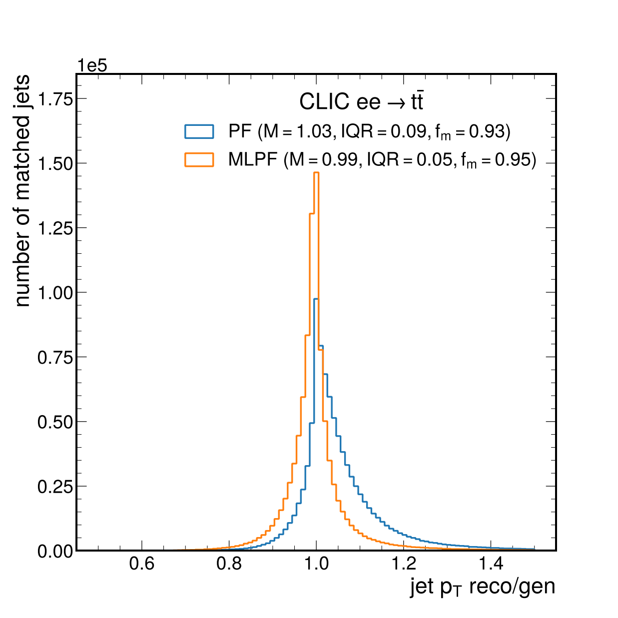
MLPF event reconstruction [1]

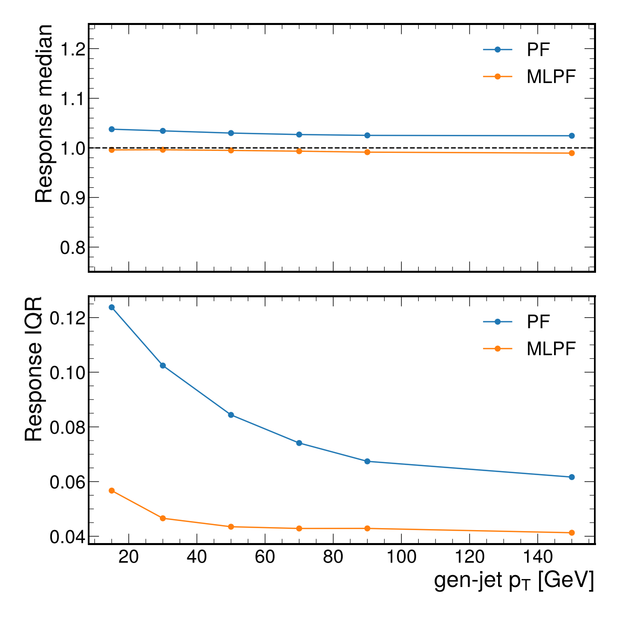
[1] Pata, J., Duarte, J., Mokhtar, F., Wulff, E., Yoo, J., Vlimant, J.-R., Pierini, M., Girone, M. (2022). *Machine Learning for Particle Flow Reconstruction at CMS*. Retrieved from http://arxiv.org/abs/2203.00330

Jet and MET in ttbar test data

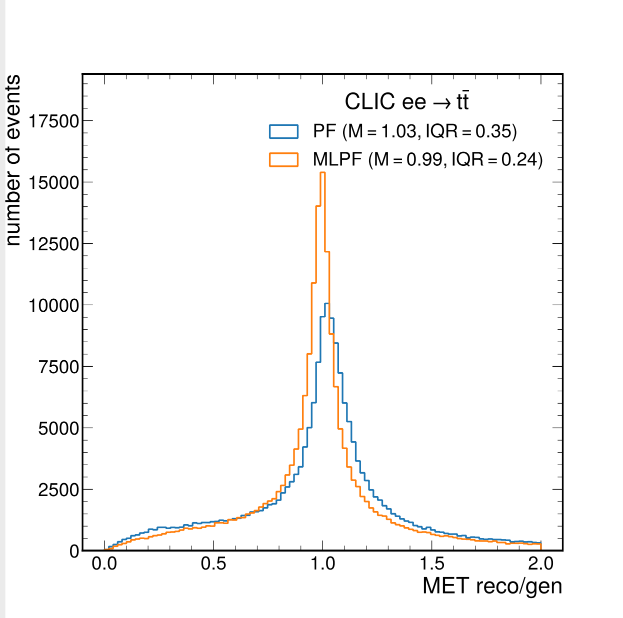
- For all test samples MLPF outperforms PF in Jet and MET reconstruction in terms of response width (quantified by median and interquartile range IQR)
- MLPF also outperforms PF in terms of fraction of reconstructed jets $\left({}^{n_{reco\ jets}}/_{n_{ground-truth\ jets}}\right)$

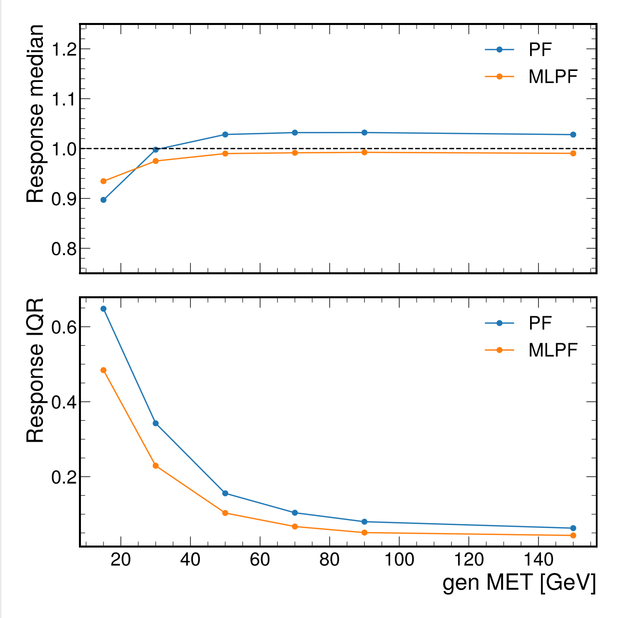
Jets resolution





MET resolution

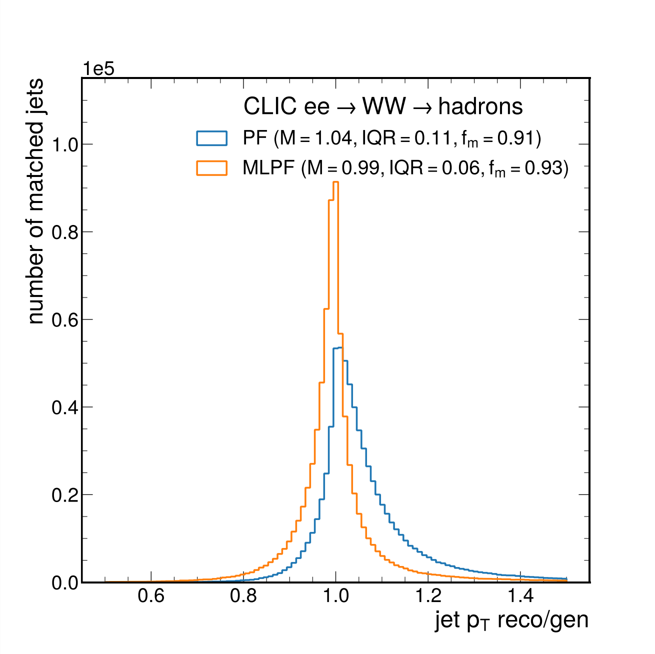




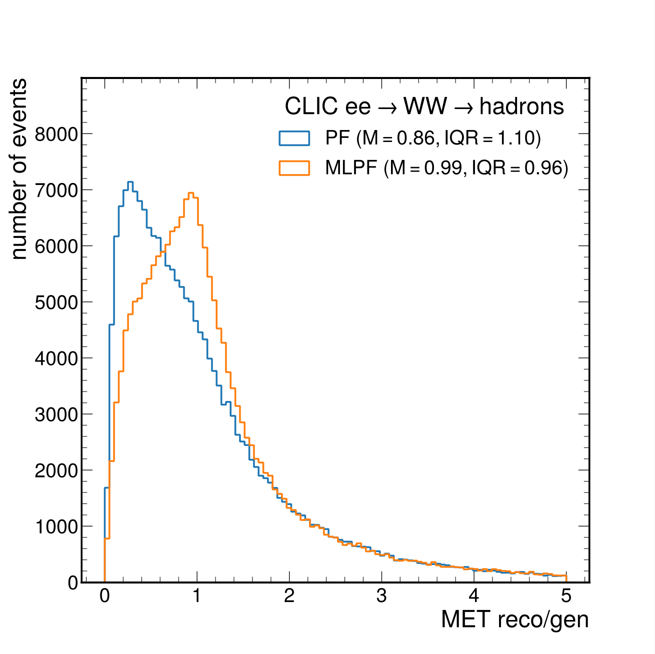
Jet and MET in WW test data

- For all test samples MLPF outperforms PF in Jet and MET reconstruction in terms of response width (quantified by median and interquartile range IQR)
- MLPF also outperforms PF in terms of fraction of reconstructed jets $\left({}^{n_{reco\ jets}}/_{n_{ground-truth\ jets}}\right)$

Jets resolution



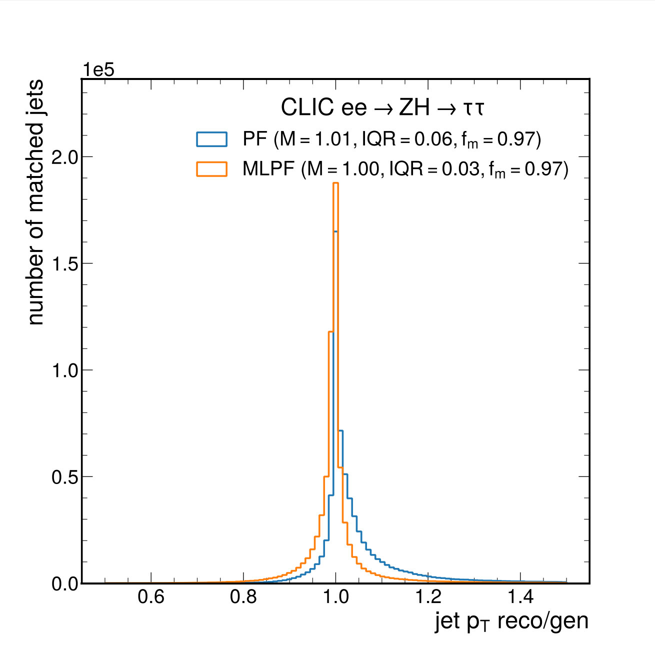
MET resolution



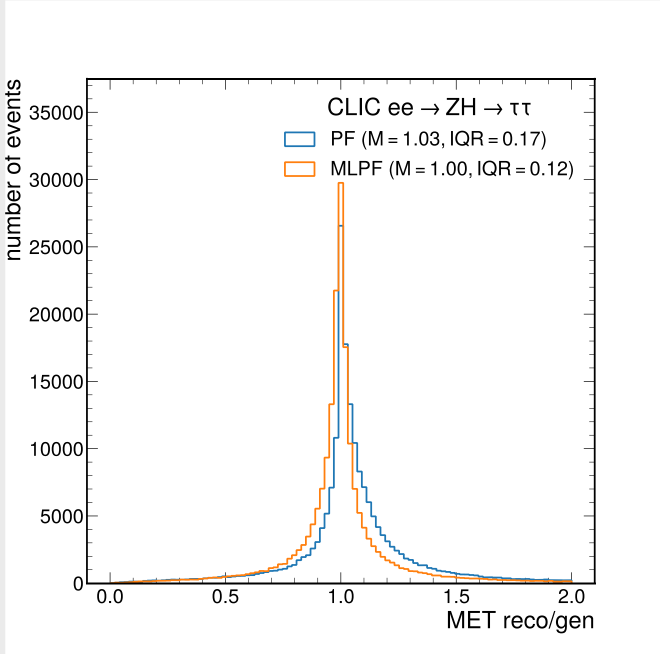
Jet and MET in ZH test data

- For all test samples MLPF outperforms PF in Jet and MET reconstruction in terms of response width (quantified by median and interquartile range IQR)
- MLPF also outperforms PF in terms of fraction of reconstructed jets ($^{n_{reco\ jets}}/_{n_{ground-truth\ jets}}$)

Jets resolution



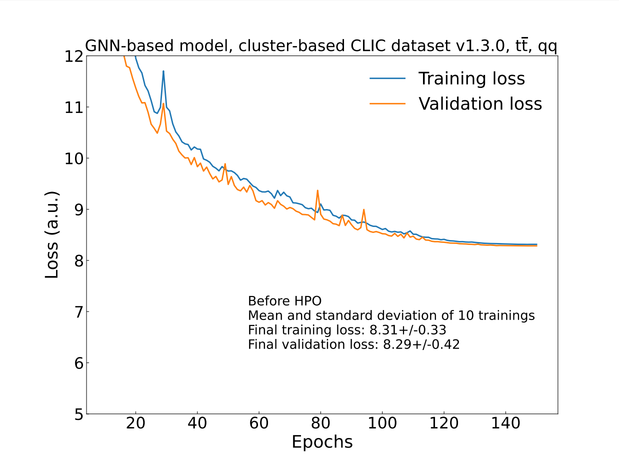
MET resolution

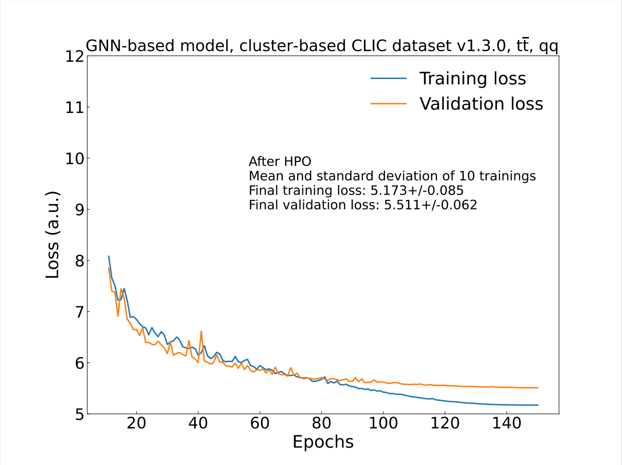


Improvement in training from HPO

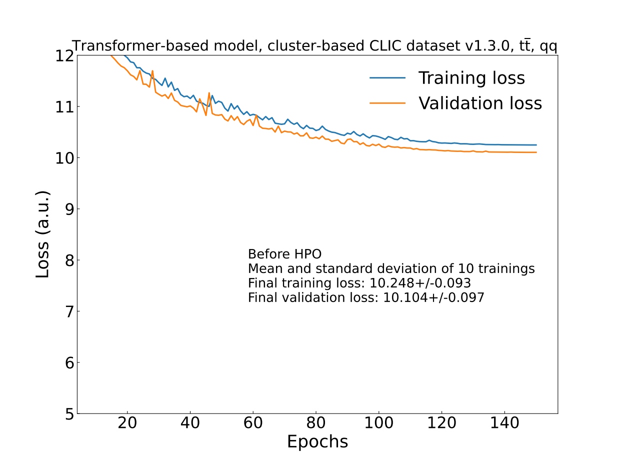
- Same learning curves as in the previous slide but separated into one plot per model and HP-set

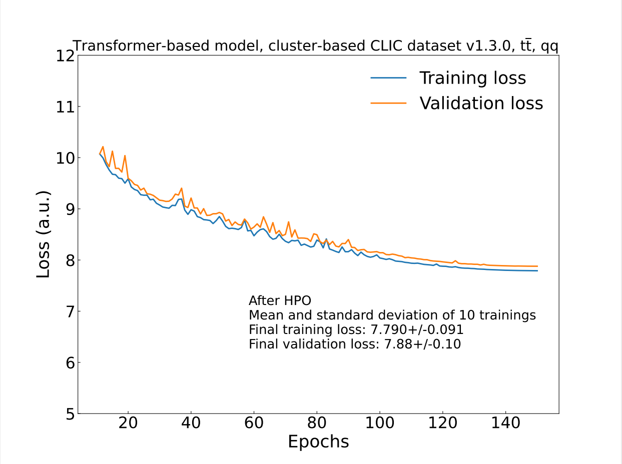
GNN





Transformer



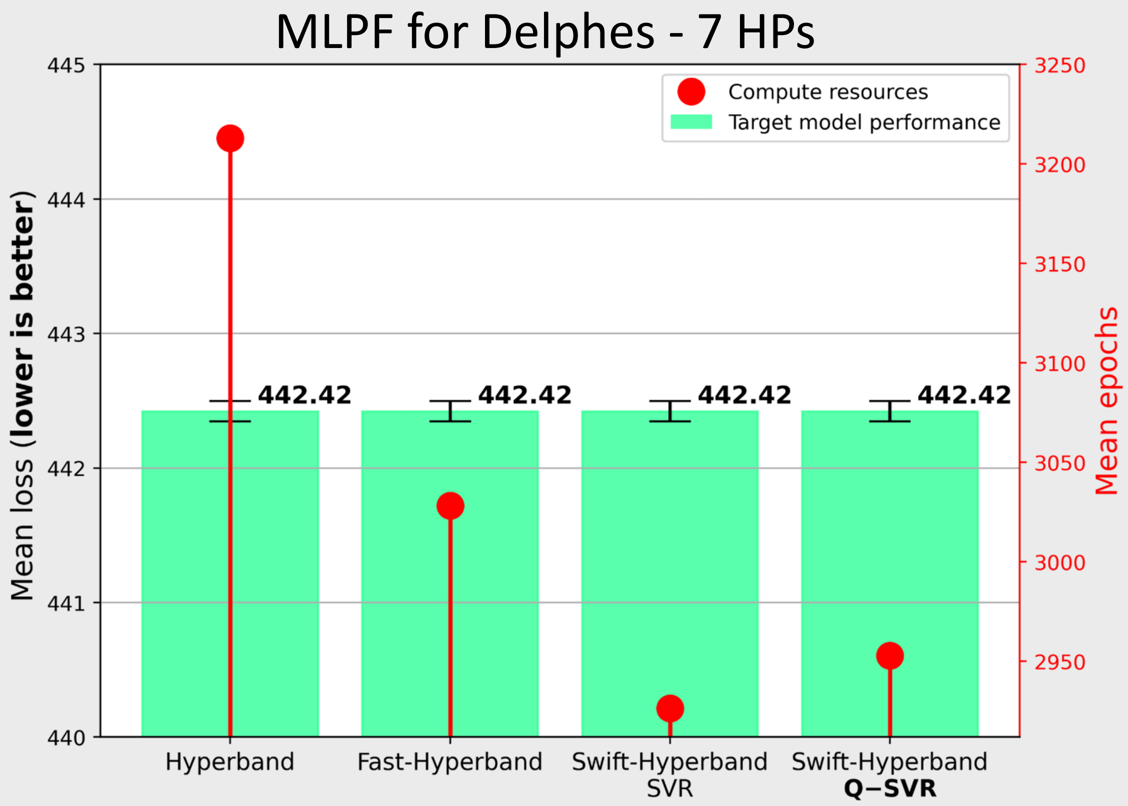


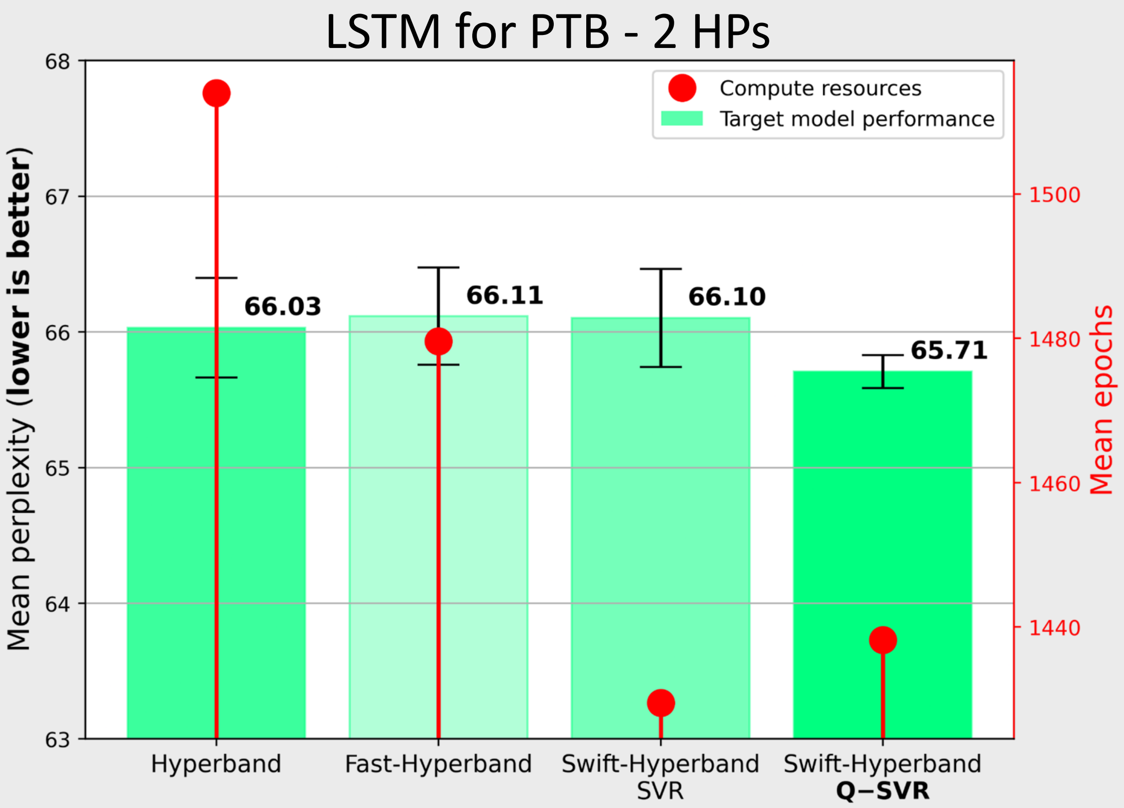
Before HPO

After HPO

Algorithm Comparison

- Simulated results using datasets of existing learning curves





Algorithm Comparison

- Simulated results using datasets of existing learning curves