A Machine Learning Framework to improve the search for New Physics in the ATLAS experiment

Alejandro Pérez García

Collaborators: José Salt, Guillem Arbona, Santiago González de la Hoz, Josu Cantero and Miguel Vilaplana

20/11/2025

CPAN Ingenio 2010

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

CSIC Consejo Superior de Investigaciones Científicas

Vniversitat de València

IFIC Institut de Física Corpuscular

Excelencia Severo Ochoa

- Introduction
  - Focus and motivation
- Key elements
  - ATLAS Computing facilities
  - Generated Data
  - ML methods
- Itineraries
  - Classification
  - Regression
  - ARTEMISA implementation
- Conclusion

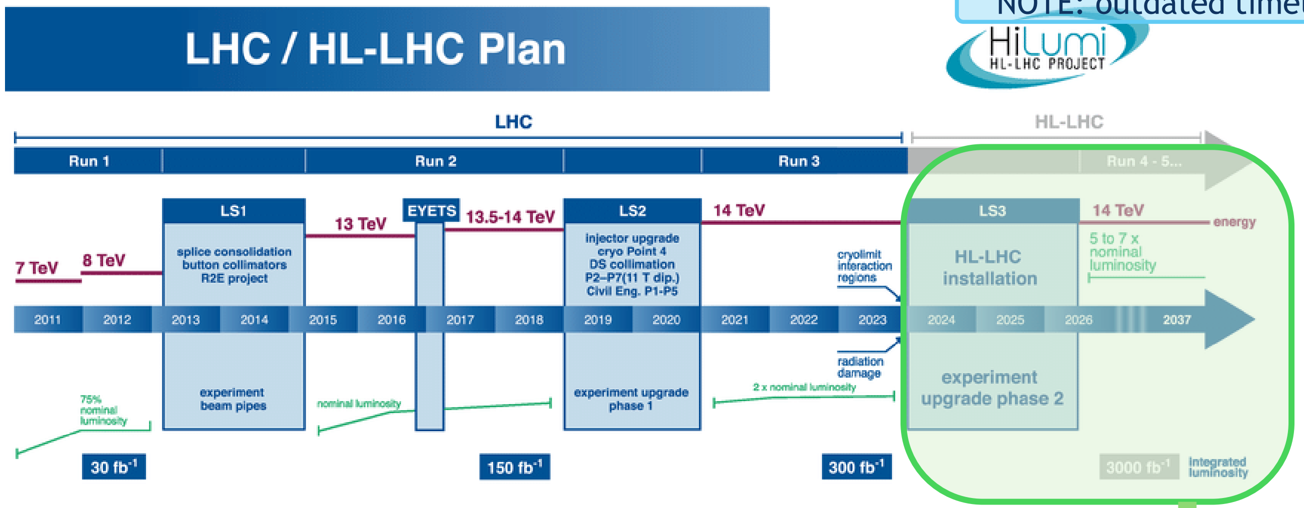
INTRODUCTION

Focus and motivation

INSTITUT DE FÍSICA CORPUSCULAR

Final Run of LHC – software and hardware

NOTE: outdated timeline

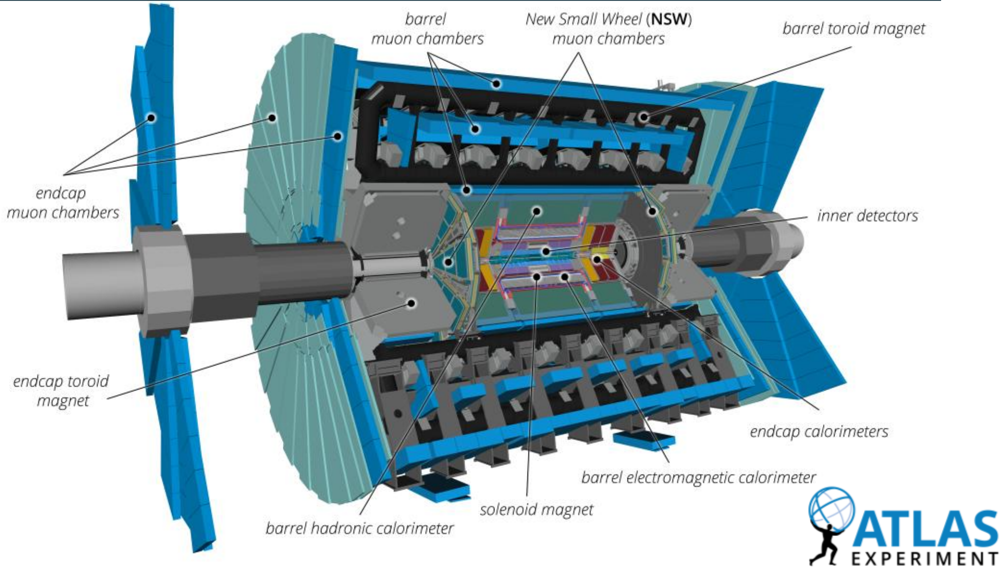


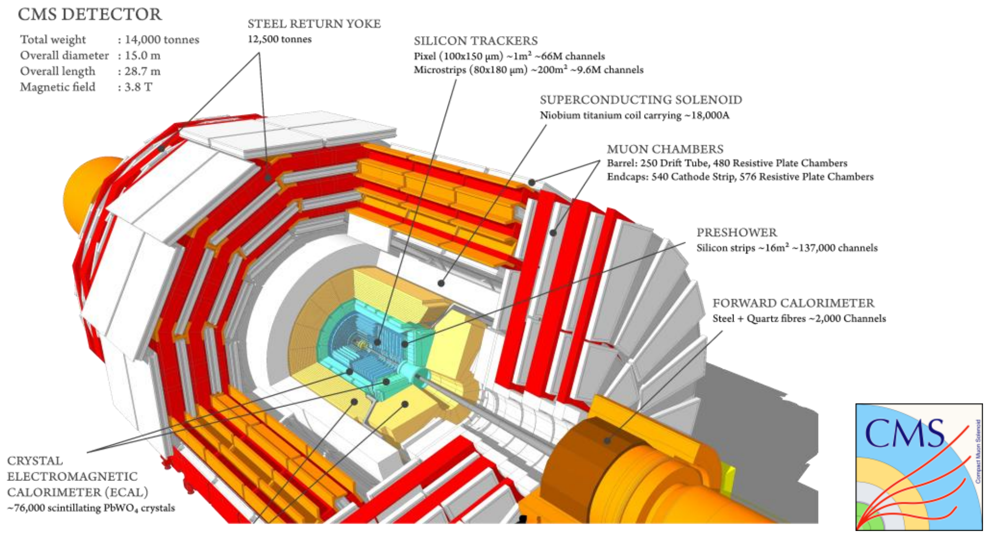
Necessary training in ML/DL for future PhDs

- Analysis will benefit from AI methods
- Broad offer of libraries, ML/DL

TensorFlow scikit learn PyTorch

ATLAS & CMS - general purpose detectors





Precedents, focus and motivation

INSTITUT DE FÍSICA CORPUSCULAR

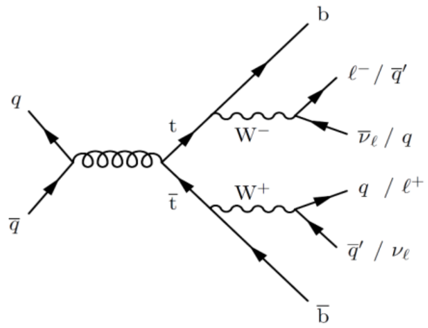
Trajectory of numerous TFGs and TFMs

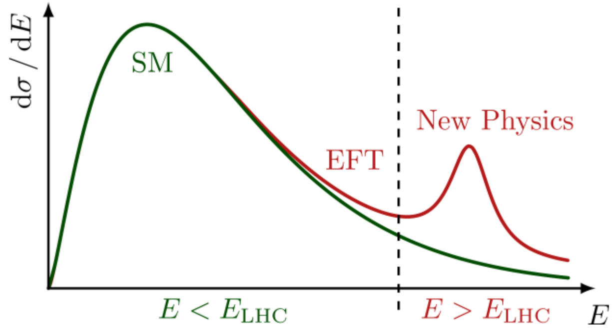
Deep Learning to improve Experimental Sensitivity and Generative Models for Monte Carlo simulations for searching for New Physics in LHC experiments

José Salt[1,*], *Raúl* Balanzá[2], *Azael* Garcia[4], *Jon Ander* Gomez[2], *Santiago* González de la Hoz[1], *Julio* Lozano[3], *Roberto* Ruiz de Austri[1], and *Miguel* Villaplana[1]

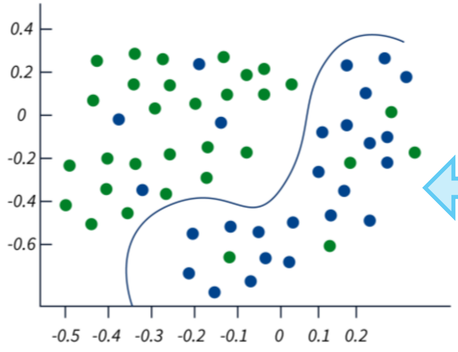
[1]Instituto de Física Corpuscular (IFIC) - Universitat de València (UV) and CSIC, Valencia, Spain
[2]Universitat Politecnica de Valencia (UPV), Valencia, Spain
[3]Universidad de Alcalá de Henares (UAH), Spain
[4]Universitat de València (UV), Spain

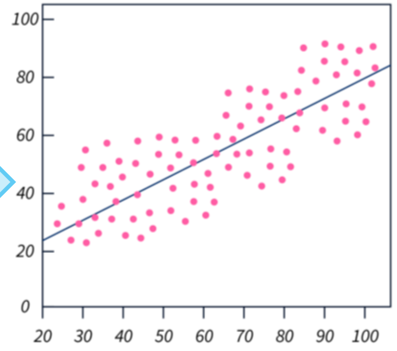
$t\bar{t}$ events → BSM probe





ML itineraries





Artemisa
ARTificial Environment for ML and Innovation in Scientific Advanced Computing

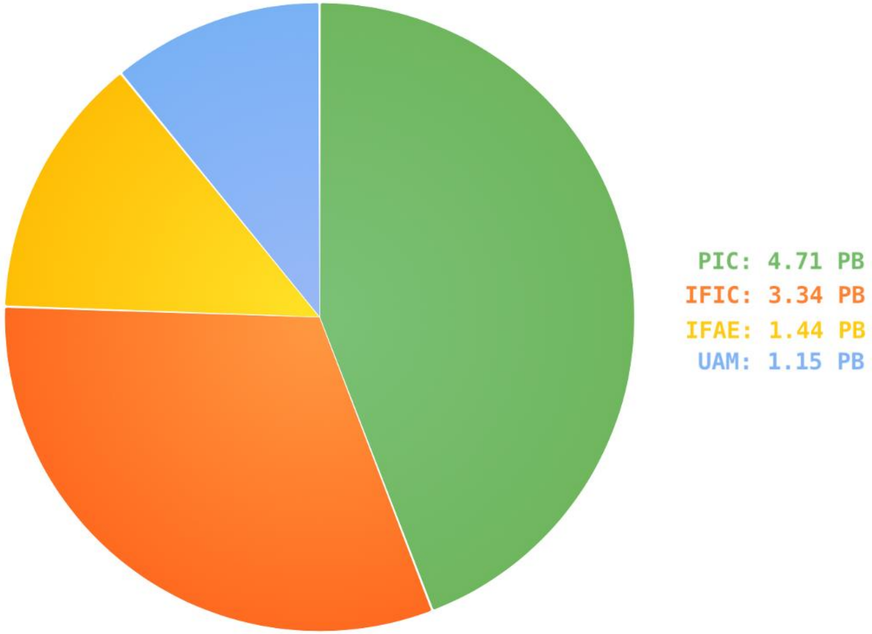
KEY ELEMENTS

CERN Analysis facilities, simulation, ML methods

ATLAS Computing Facilities in Spain

Spanish ATLAS Tier-1 & Tier-2s

Collaboration and coordination

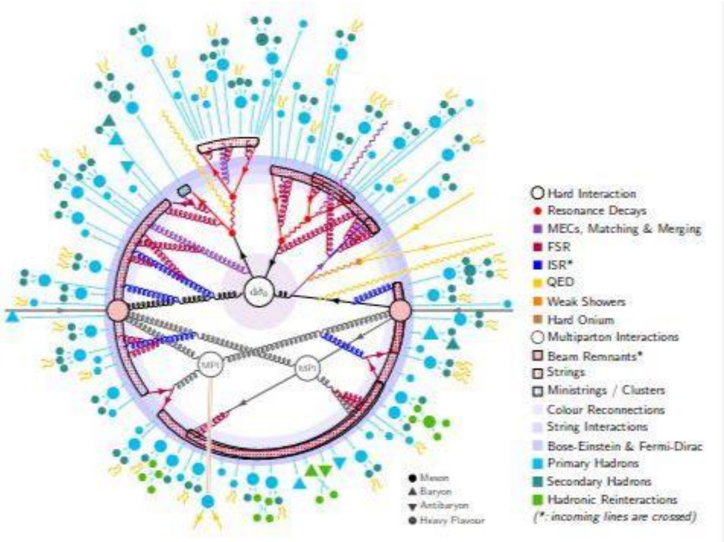


Disk usage 2024

MC Simulated Data

Simulation software

Pythia8



GEANT4 — A SIMULATION TOOLKIT

CPU demanding

Simulation parameters

Setting name	Setting description	CMS default	ATLAS default	Common Proposal
POWHEG				
qmass	top-quark mass [GeV]	172.5	172.5	172.5
twidth	top-quark width [GeV]	1.31	1.32	1.315
hdamp	first emission damping parameter [GeV]	237.8775	258.75	250
wmass	$W^{\pm}$ mass [GeV]	80.4	80.3999	80.4
wwidth	$W^{\pm}$ width [GeV]	2.141	2.085	2.11
bmass	b-quark mass [GeV]	4.8	4.95	4.875
PYTHIA 8				
	PYTHIA 8 version	v240	v230	v240 (CMS) v244 (ATLAS)
	Tune	CP5	A14	Monash
PDF:pSet	LHAPDF6 parton densities to be used for proton beams	NNPDF31_nnlo _as_0118	NNPDF23_lo _as_0130_qed	NNPDF23_lo _as_0130_qed
TimeShower:alphaSvalue	Value of α_s at Z mass scale for Final State Radiation	0.118	0.127	0.1365
SpaceShower:alphaSvalue	Value of α_s at Z mass scale for Initial State Radiation	0.118	0.127	0.1365
MPI:alphaSvalue	Value of α_s at Z mass scale for Multi-Parton Interaction	0.118	0.126	0.130
MPI:pT0ref	Reference p_T scale for regularizing soft QCD emissions	1.41	2.09	2.28
ColourReconnection:range	Parameter controlling colour reconnection probability	5.176	1.71	1.80

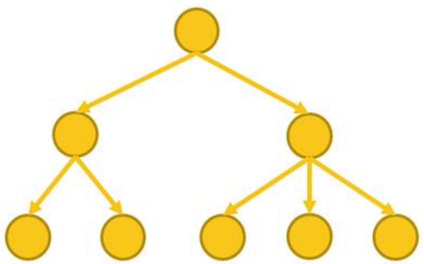
Negro, Giulia. (2022). Top Quark Modelling and Tuning in ATLAS and CMS. 10.48550/arXiv.2201.03517.

Experiment-theory loop update

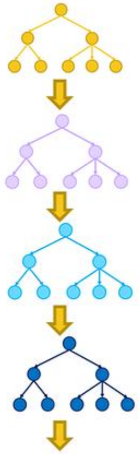
DTs, BDTs, NNs – Deep Sets

DT based methods

Single Decision Tree **DT**

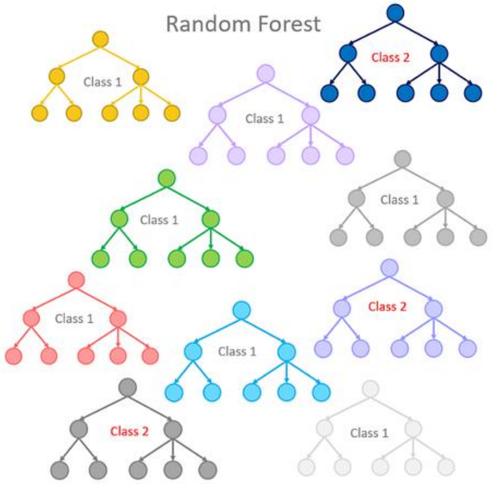


Gradient Boosted Trees **BDT**



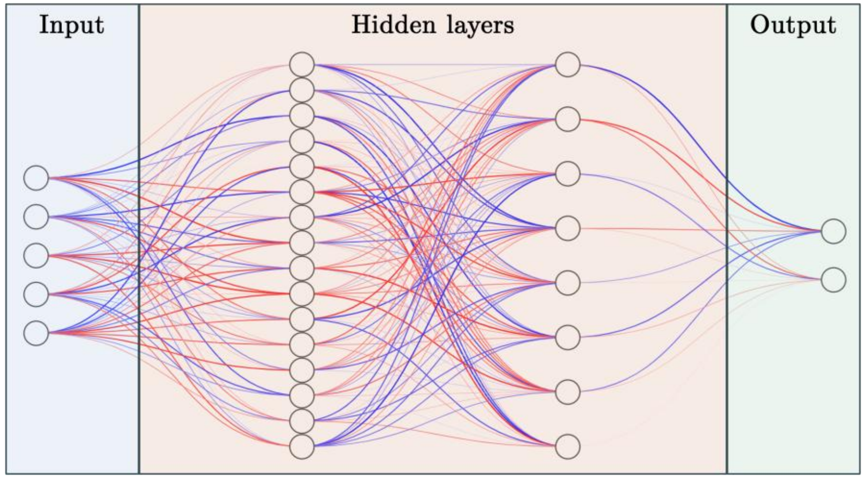
RF

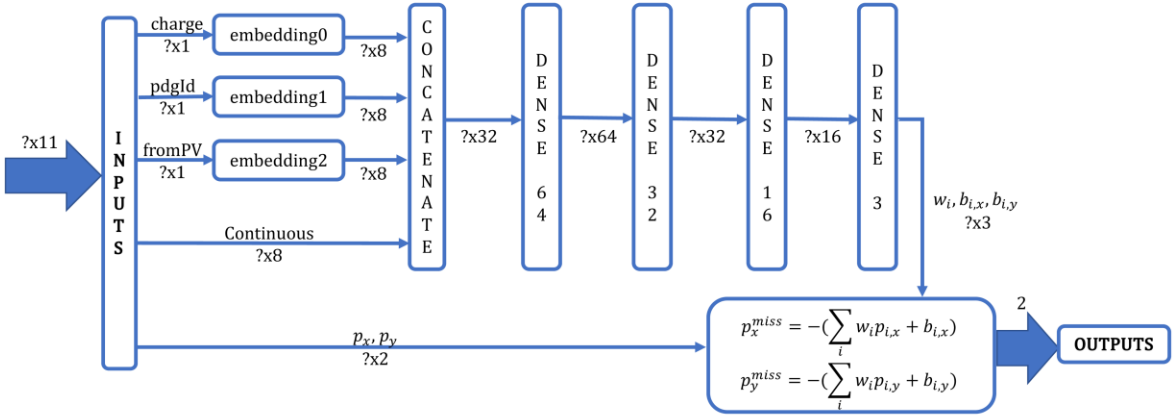
Random Forest



scikit learn

Neural Networks





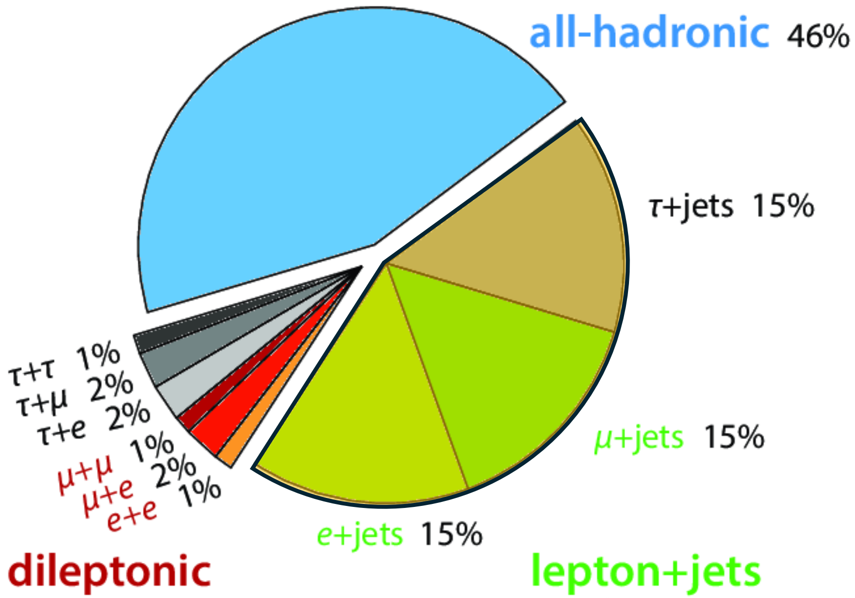
TensorFlow

ITINERARIES

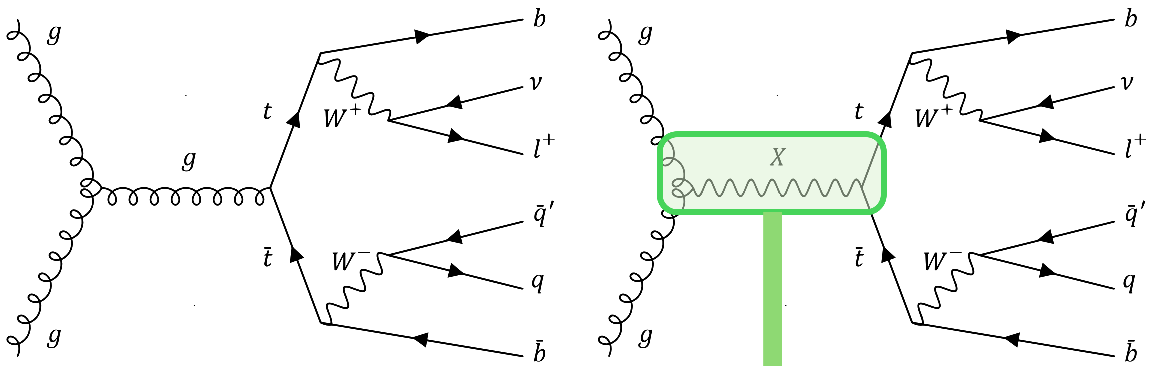
Classification and Regression

Use Case 1: SM/BSM semileptonic $t\bar{t}$ classification

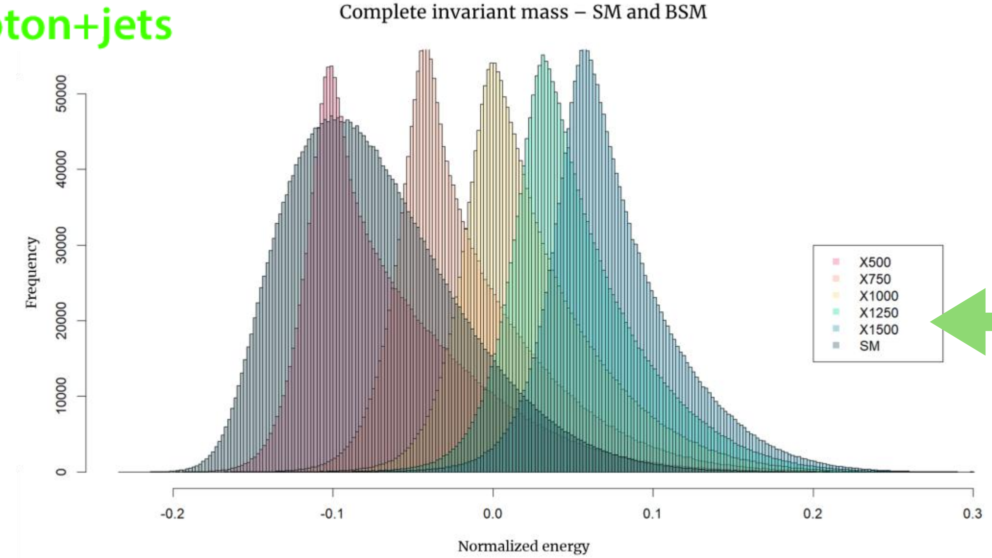
Semileptonic $t\bar{t}$



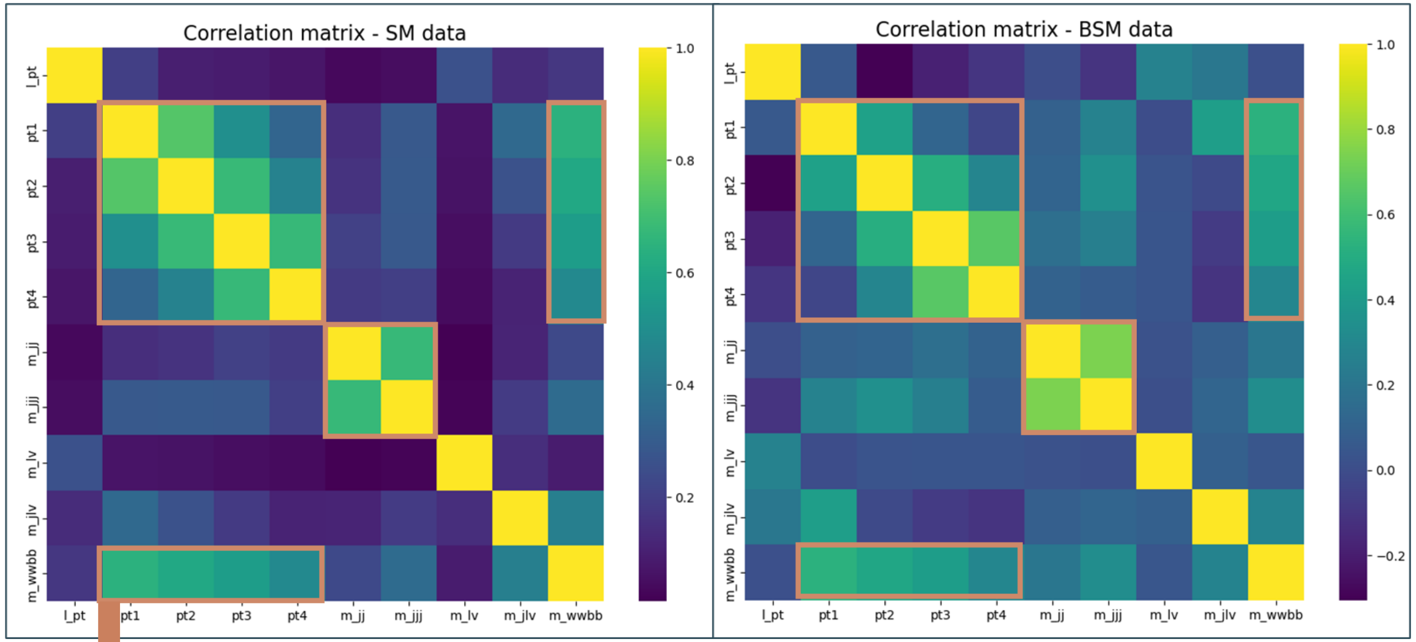
SM (gluon) vs BSM (Z')

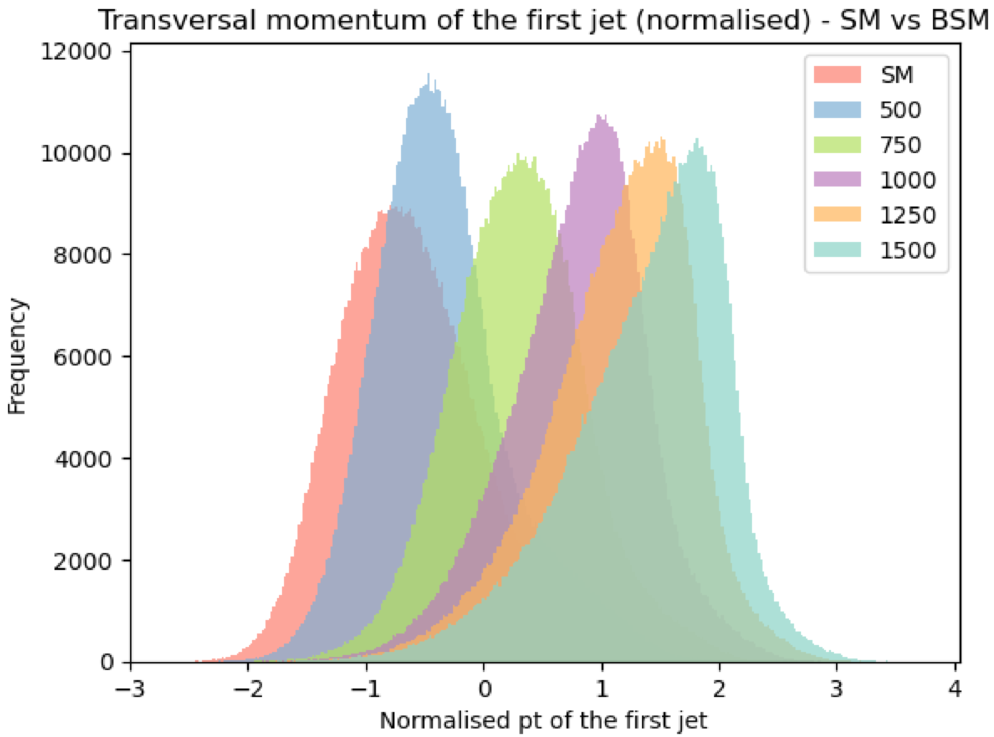


Different $M_{Z'}$



Use Case 1: correlation matrices and first plots



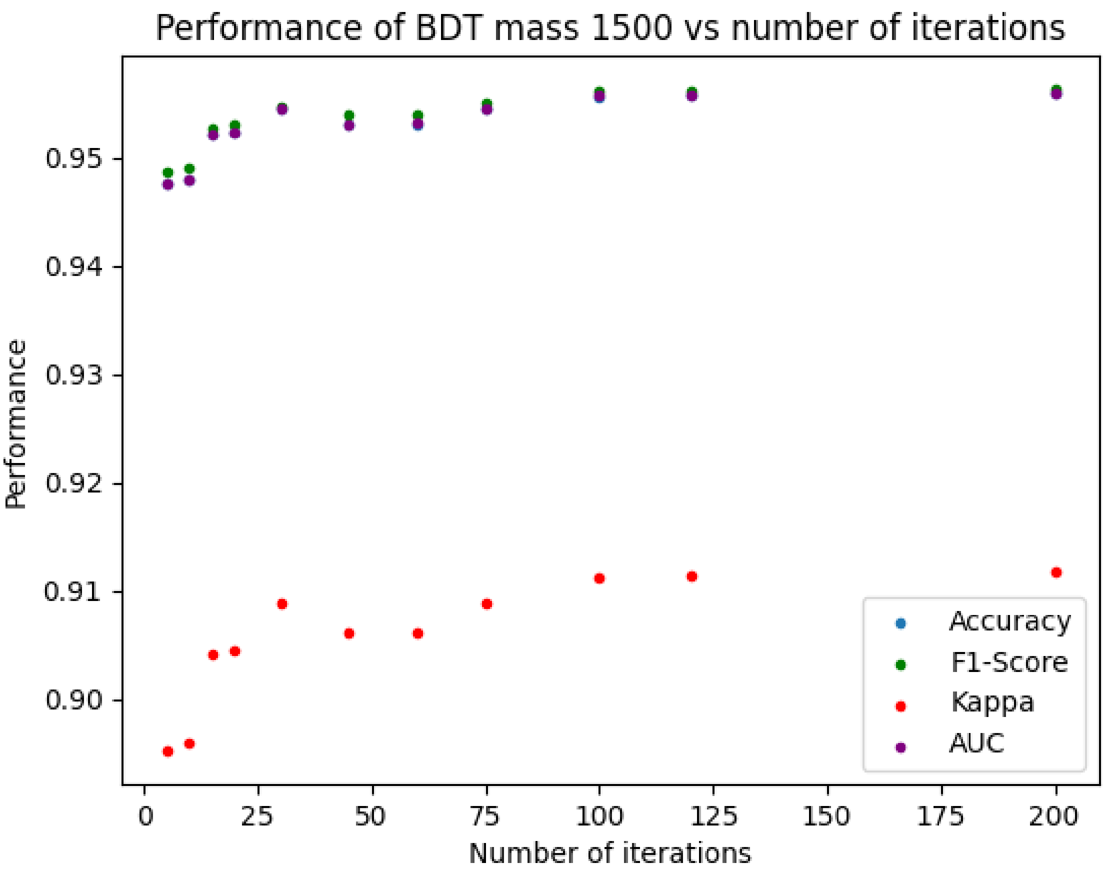


SM Gluon vs $t\bar{t}$ resonance ($M_{Z'}$ = 1000 GeV)

$t\bar{t}$ resonance ($M_{Z'} \in [500, 1500]$ GeV)

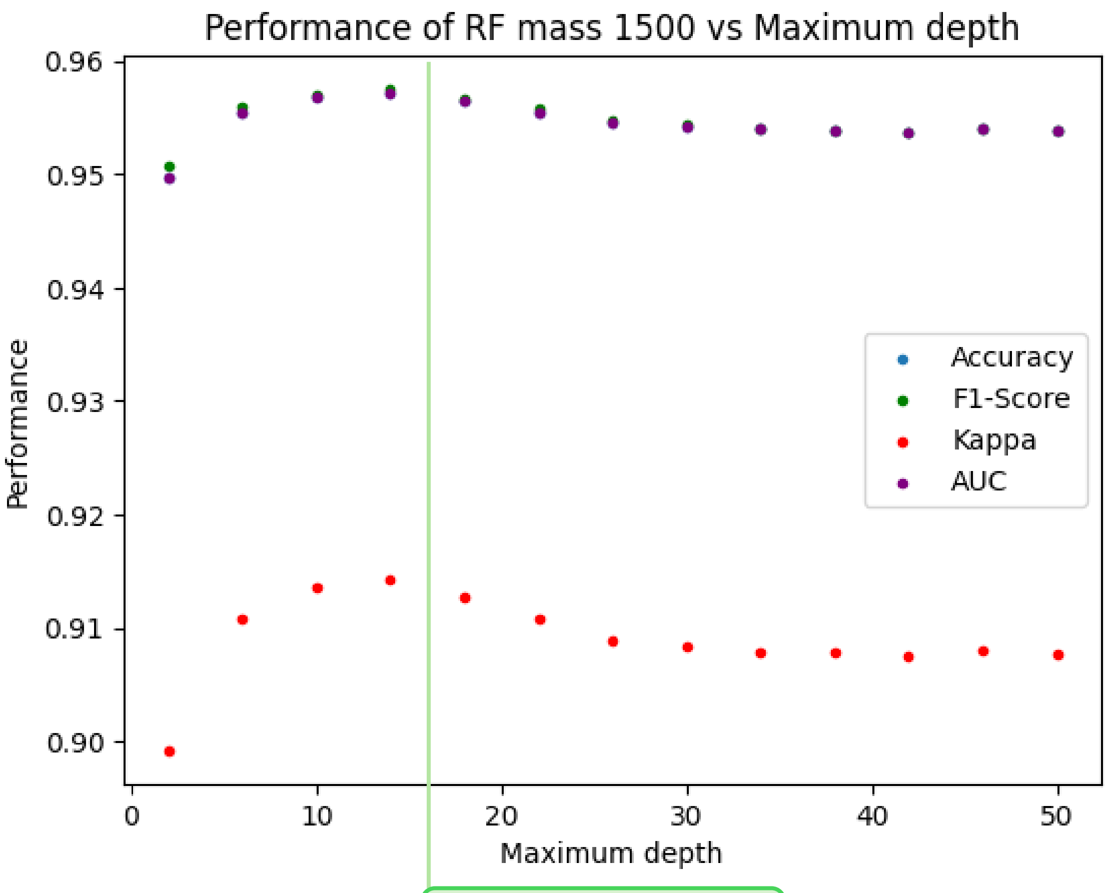
Key tendencies

Number of BDT iterations



Assymptotic improvement

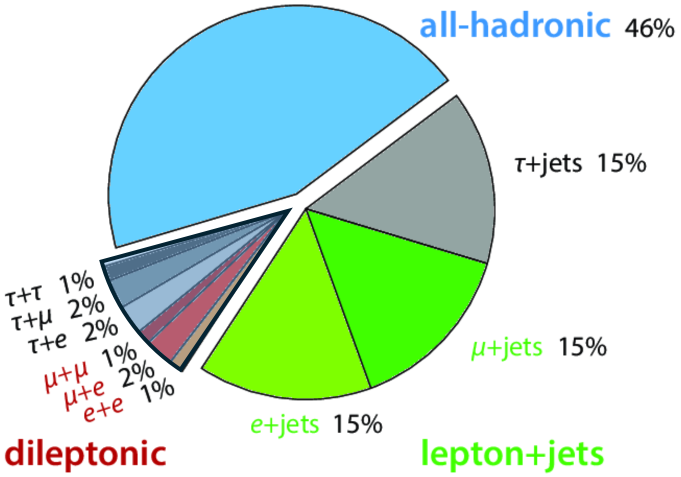
Max depth of RF

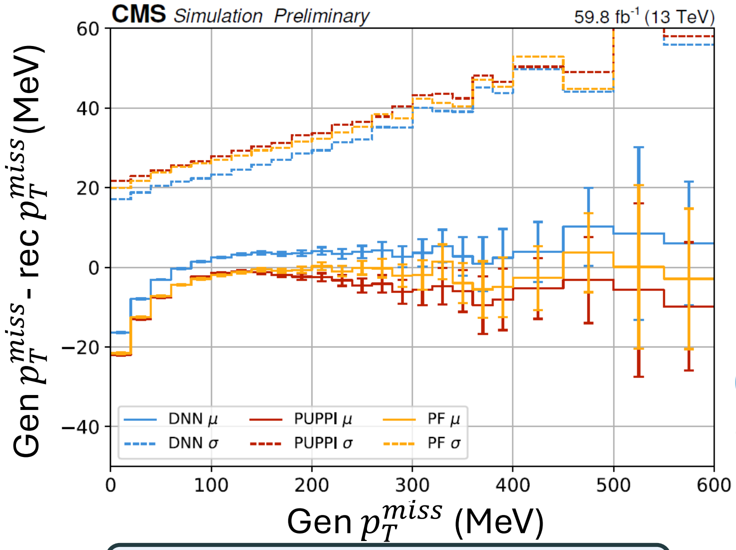


Overfitting

Use Case 2: Regression

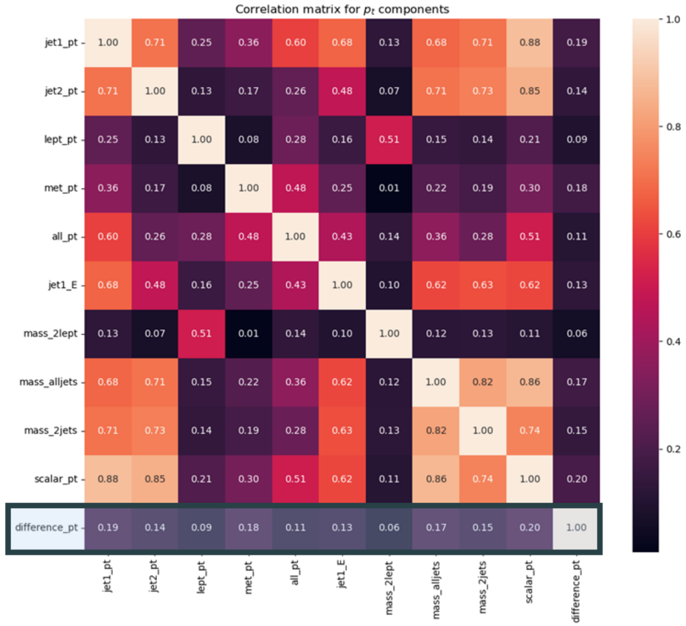
Dileptonic ttbar regression





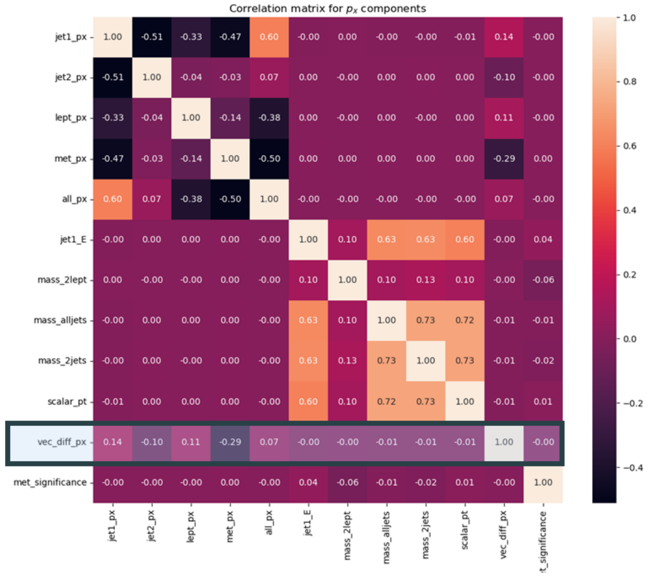
CMS precedent

Correlation matrices



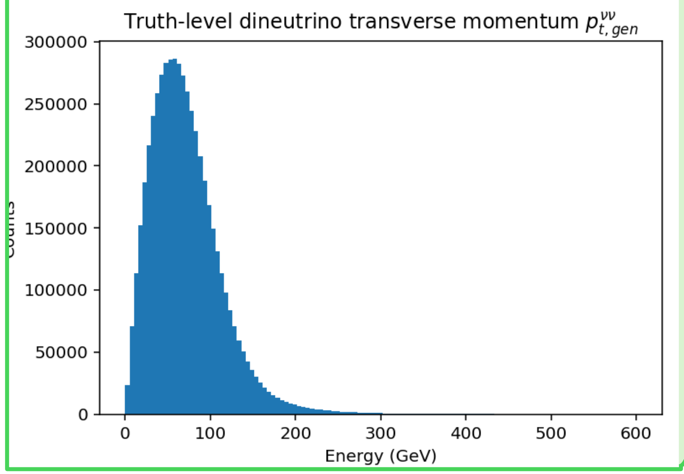
Very low correlation

Best if px/py coordinates



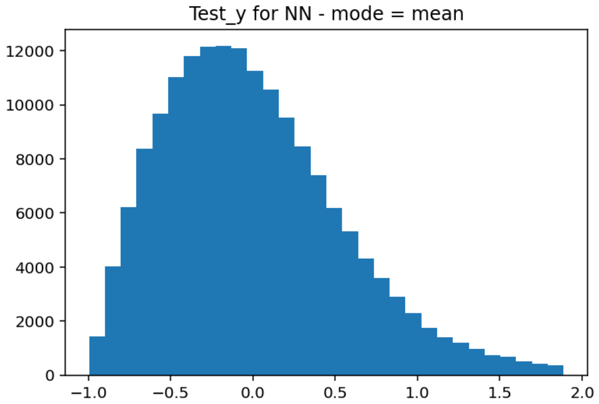
Use Case 2: architecture exploration

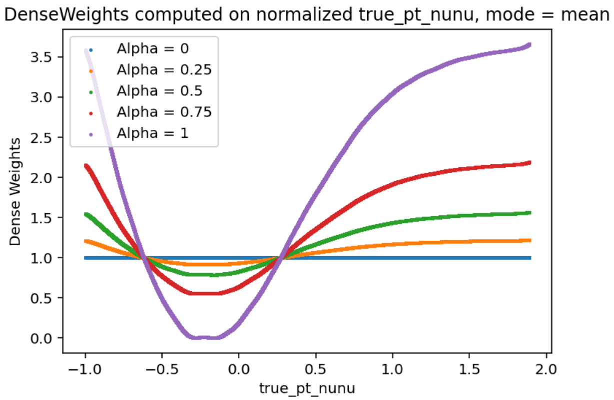
Reweighting procedure



Imbalanced dataset

DenseWeights



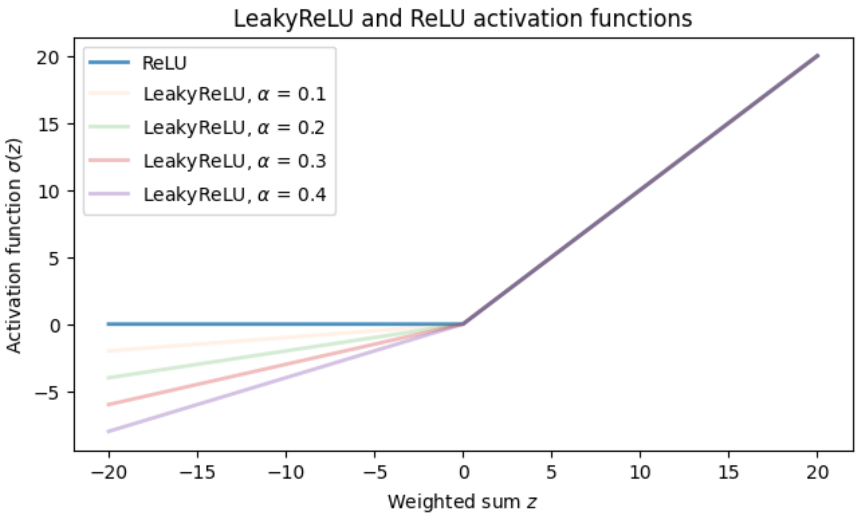


Hyperparameter testing

Hyperparameter	Options
Loss function	[logcosh, huber, mse, mae, custom]
Activation function	[relu, leaky relu, sigmoid]
First hidden layer neurons	[16, 32, 64, 128, 256, 512]
Hidden layers neurons ratio	[1/8, 1/4, 1/2, 1, 2]
Epochs	[50, 100, 500]
Dropout	[0.1, 0.2, 0.3, 0.4]
Alpha (DenseWeights)	[0.25, 0.5, 0.75, 1]
Patience (Early Stopping)	[5, 10, 20]

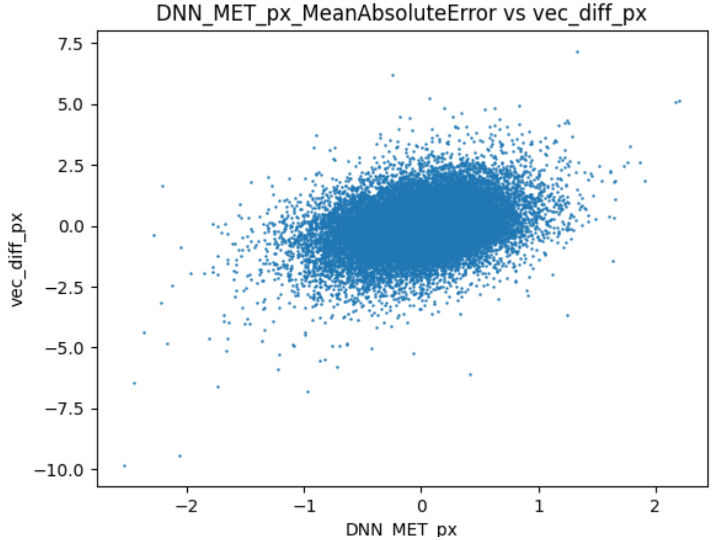
Code	Procedure	Range
`mean`	$x_n = (x - \bar{x})/\bar{x}$	-
`mean_std`	$x_n = (x - \bar{x})/\sigma_x$	-
`min_max`	$x_n = (x - \min(x))/(\min(x) - \max(x))$	$[0, 1]$
`log`	$x_1 = \log(x + \kappa) \rightarrow x_n = (x_1 - \bar{x}_1)/\bar{x}_1$	$[0, +\infty]$

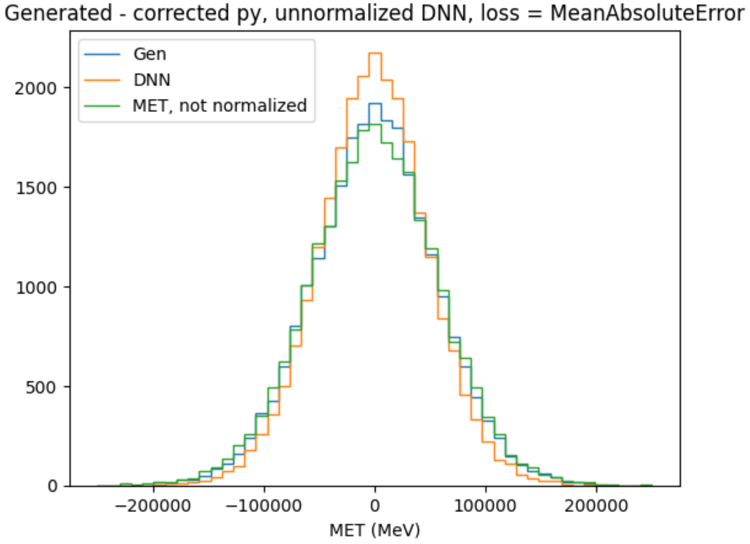
Activation function



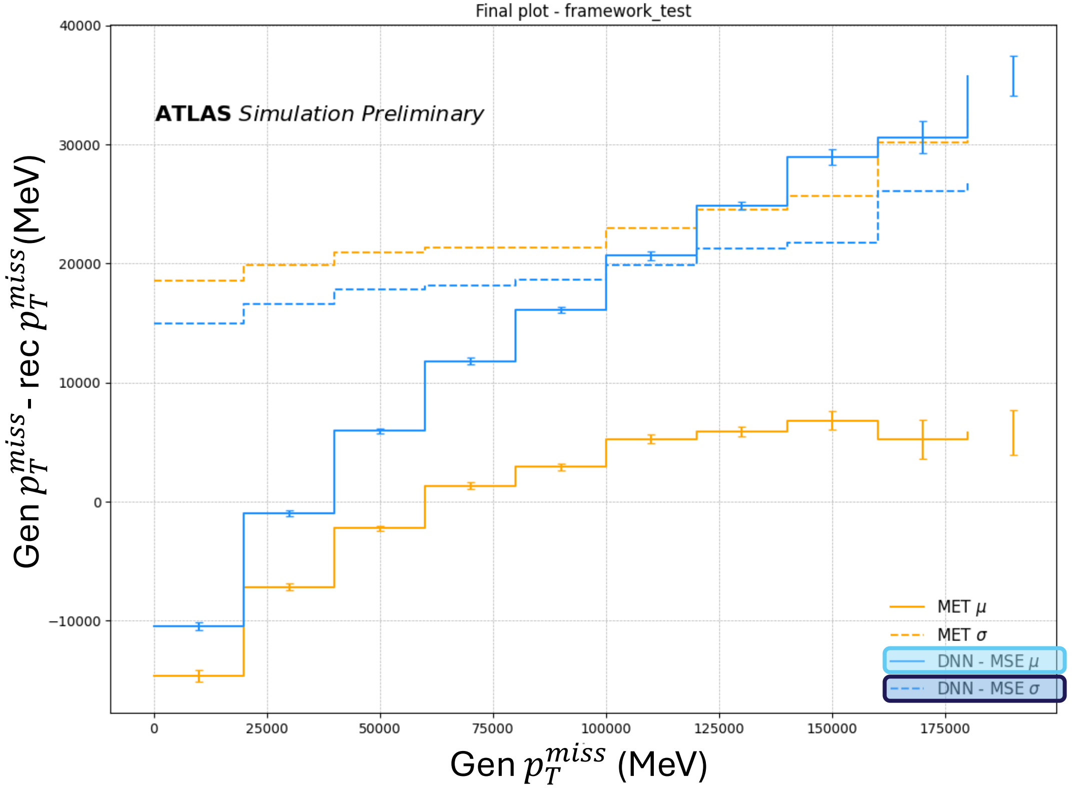
Use Case 2: key results

Strong deviation – ongoing project





Preliminary results

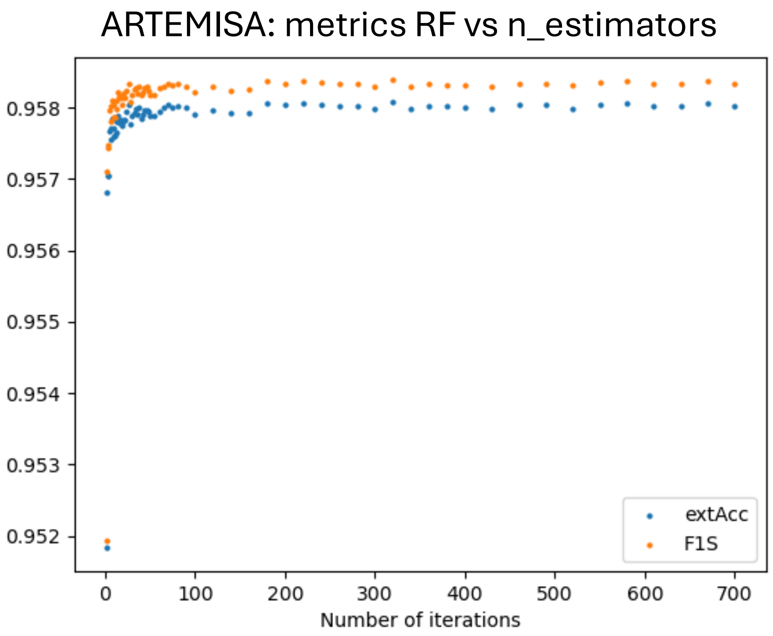


Slight resolution improvement

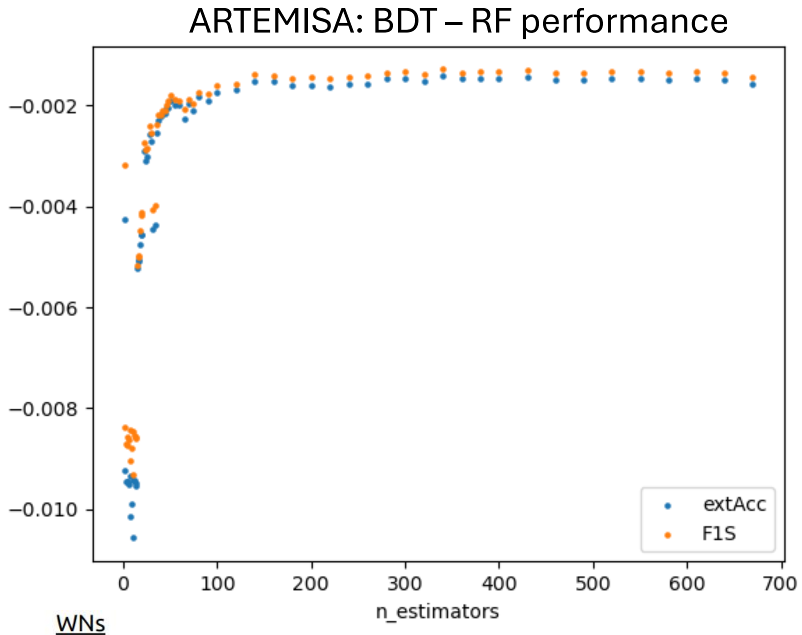
Notably worse reconstruction

Use Case 3: ARTEMISA infrastructure and results

RF number of trees



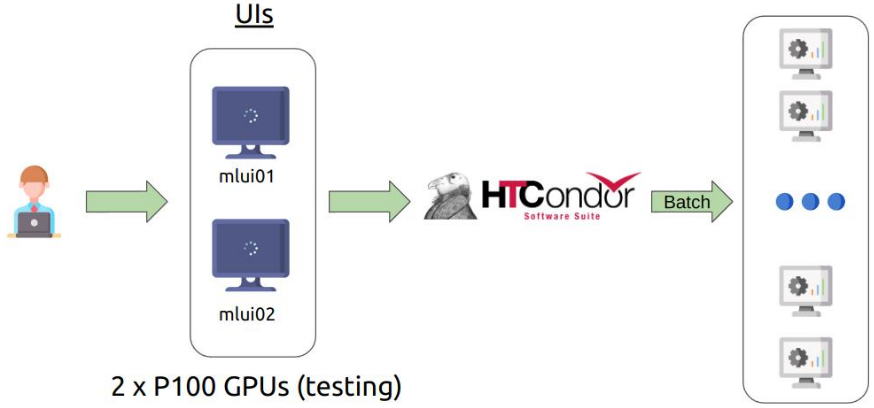
BDT vs RF



Local infrastructure

Artemisa

ARTificial Environment for ML and Innovation in Scientific Advanced Computing



2 x P100 GPUs (testing)

43 x A100/V100 GPUs

CONCLUSION

- Deep initiation into ML/DL applied to Physics
- Useful for Bachelor's / Master's & first steps in HEP analysis
- Classification: ttbar BSM signal
- Regression: MET reconstruction
- ARTEMISA: high-end computing infrastructure

Key references

- “Estudi de l’extracció del senyal ressonant de successos ttbar mitjançant tècniques de Machine Learning en l’experiment ATLAS”, Alejandro Pérez García, Bachelor’s Thesis.

- “Approaching new challenges in the search for New Physics in ATLAS ttbar events”, Alejandro Pérez García, Master’s Thesis.

- “Estimation of Systematic Errors in Deep Learning Methods applied to the extraction of ttbar resonances in ATLAS experiment”, Guillem Arbona Ferrer, Master’s Thesis.

- B. H. Hodkinson, “METNet: A combined missing transverse momentum working point using a neural network with the ATLAS detector” https://cds.cern.ch/record/2777658

- “Measurement of the dineutrino system kinematics in dileptonic top quark pair events in pp collisions at $\sqrt{s} = 13$ TeV”, CMS collaboration, 2024

- D. Meuser, “Measurement of the differential dileptonic $t\bar{t}$ cross section in a BSM phase space with the CMS detector using the full LHC Run 2 data set,” Ph.D. dissertation, Sep. 2023

The authors gratefully acknowledge the computer resources at Artemisa and the technical support provided by the Instituto de Fisica Corpuscular, IFIC(CSIC-UV). Artemisa is co-funded by the European Union through the 2014-2020 ERDF Operative Programme of Comunitat Valenciana, project IDIFEDER/2018/048.
This work was partially supported by MICINN in Spain under grant PID2022-136323OB-C21.

THANK YOU FOR YOUR ATTENTION

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

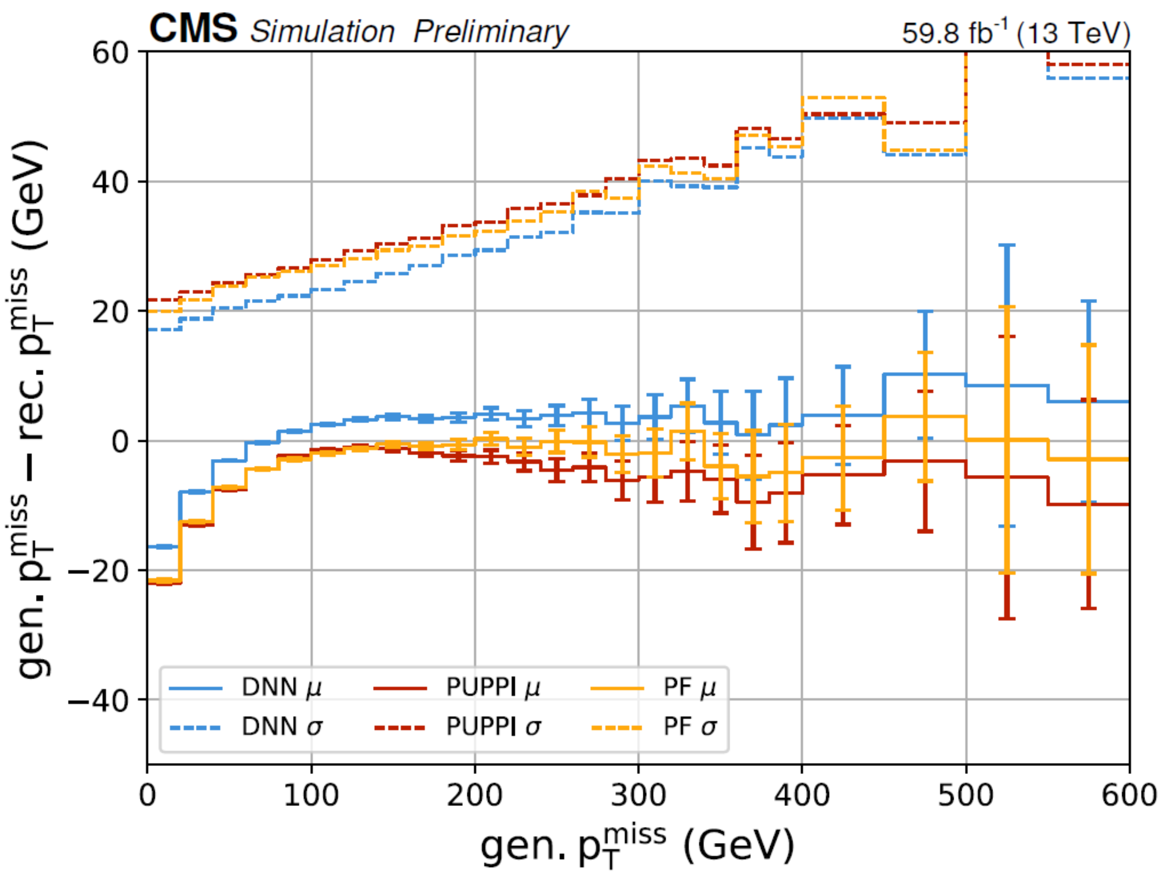
VNIVERSITAT DE VALÈNCIA

IFIC INSTITUT DE FÍSICA CORPUSCULAR

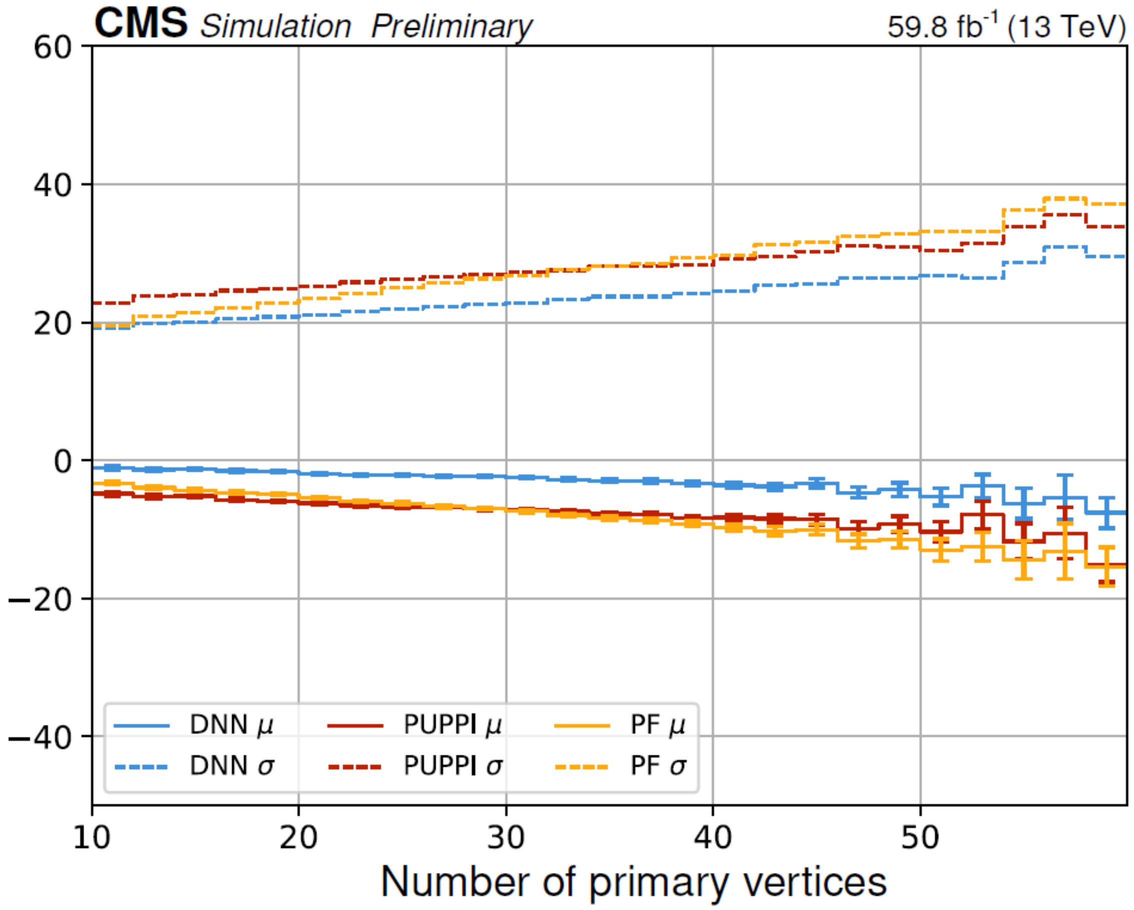
EXCELENCIA SEVERO OCHOA

ADDITIONAL SLIDES

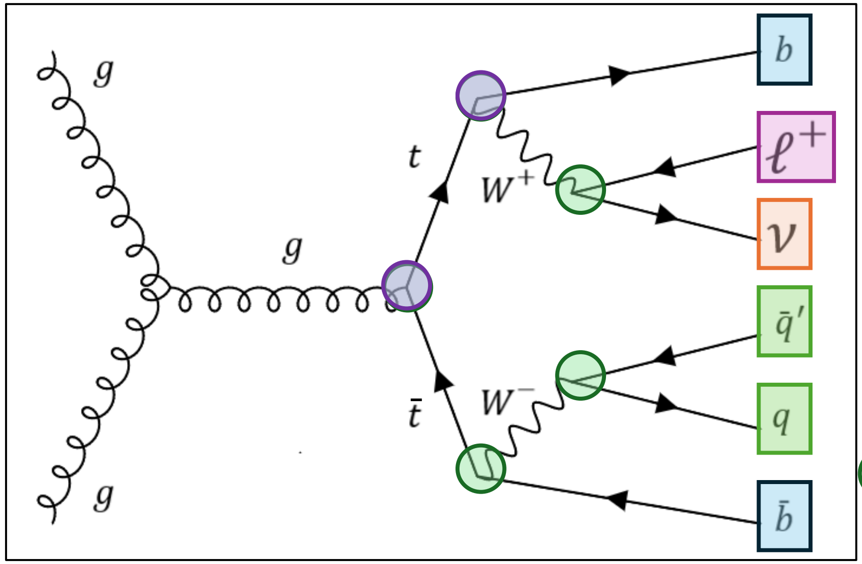
CMS precedent

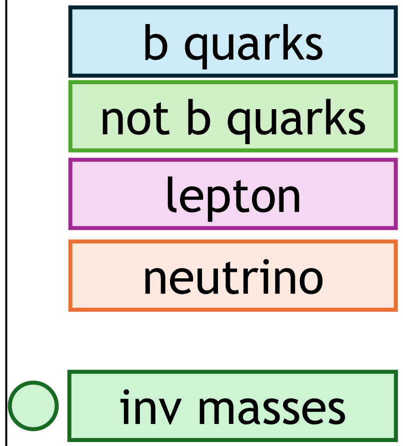


Successful attempt with Run 2 data at CMS



Pileup resilience studies at CMS





$$p_t = \sqrt{p_x^2 + p_y^2}$$
$$\varphi = \arctan(p_y/p_x)$$
$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$$
$$s_{ij}^2 = m_i^2 + m_j^2 + 2(\vec{p_i} \cdot \vec{p_j})$$
$$\approx 2(\vec{p_i} \cdot \vec{p_j})$$

HEPMASS repositori: 7M for training - 3.5M for testing

22 LL variables			5 HL variables
p_t, φ, η, btag 4 jets	p_t, φ, η lepton	p_t, φ, η MET (neutrino)	Invariant masses

Use Case 2: input variables

Pseudorapidity and invariant masses

$$p_t = \sqrt{p_x^2 + p_y^2}$$

$$\varphi = \arctan(p_y/p_x)$$

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$$

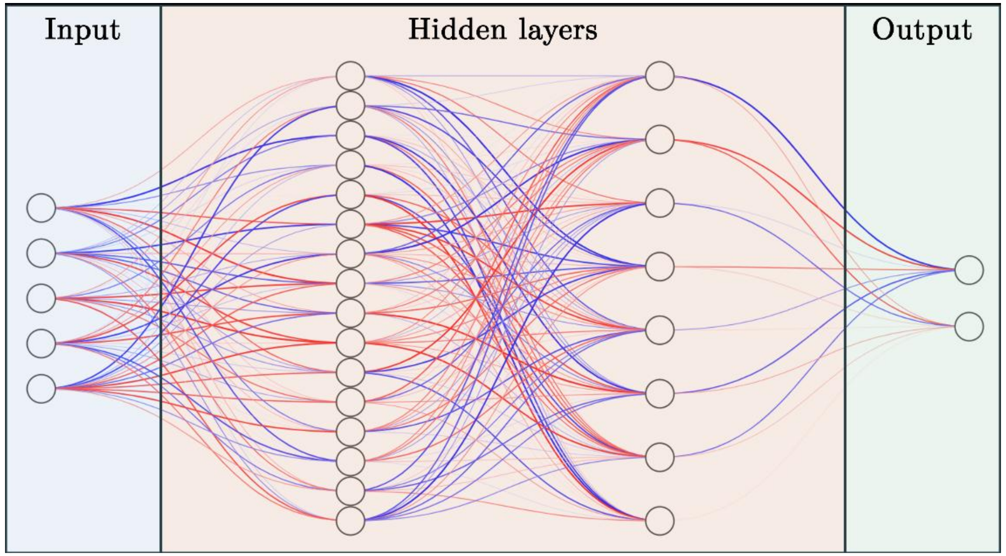
$$s_{ij}^2 = m_i^2 + m_j^2 + 2(\overrightarrow{p_i} \cdot \overrightarrow{p_j}) \approx 2(\overrightarrow{p_i} \cdot \overrightarrow{p_j})$$

New High-Level variables

2D pt	Invariant masses	Others
Leading jets	2 leptons	Leading jet E
Leading leptons	All jets	Scalar sum of jet pt
MET	2 leading jets	
All jets		

Neural Network - MLP

MLP architecture



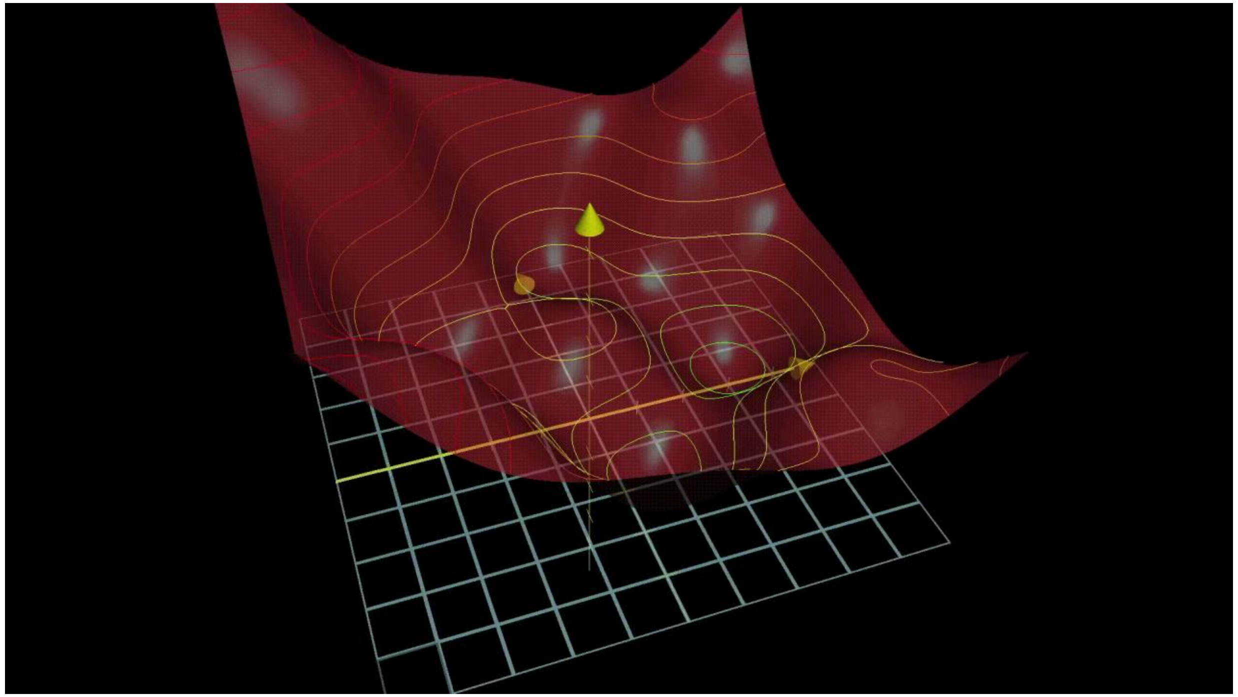
Backpropagation

$$\frac{\partial C_0}{\partial w^{(L)}} = \frac{\partial z^{(L)}}{\partial w^{(L)}} \cdot \frac{\partial a^{(L)}}{\partial z^{(L)}} \cdot \frac{\partial C_0}{\partial a^{(L)}} = a^{(L-1)} \cdot \sigma'(z^{(L)}) \cdot 2(a^{(L)} - y)$$

$$\frac{\partial C_0}{\partial b^{(L)}} = \frac{\partial z^{(L)}}{\partial b^{(L)}} \cdot \frac{\partial a^{(L)}}{\partial z^{(L)}} \cdot \frac{\partial C_0}{\partial a^{(L)}} = 1 \cdot \sigma'(z^{(L)}) \cdot 2(a^{(L)} - y)$$

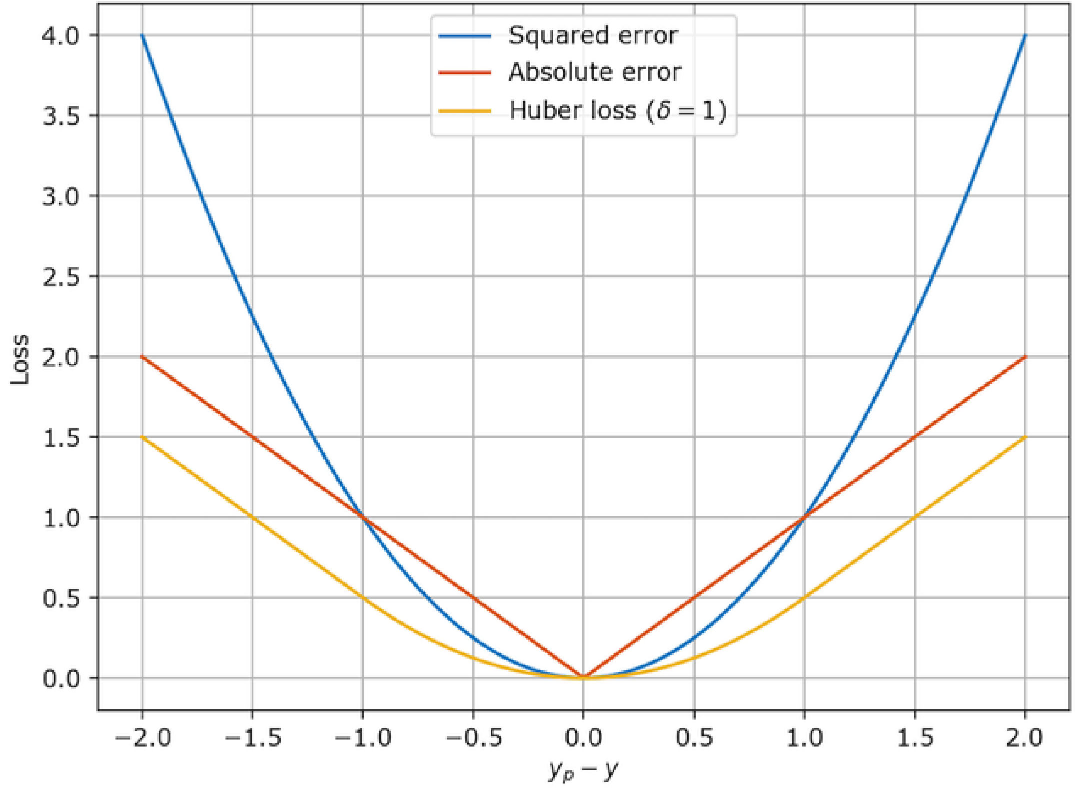
$$\frac{\partial C_0}{\partial a^{(L-1)}} = \frac{\partial z^{(L)}}{\partial a^{(L-1)}} \cdot \frac{\partial a^{(L)}}{\partial z^{(L)}} \cdot \frac{\partial C_0}{\partial a^{(L)}} = w^{(L)} \cdot \sigma'(z^{(L)}) \cdot 2(a^{(L)} - y)$$

Stochastic Gradient Descent

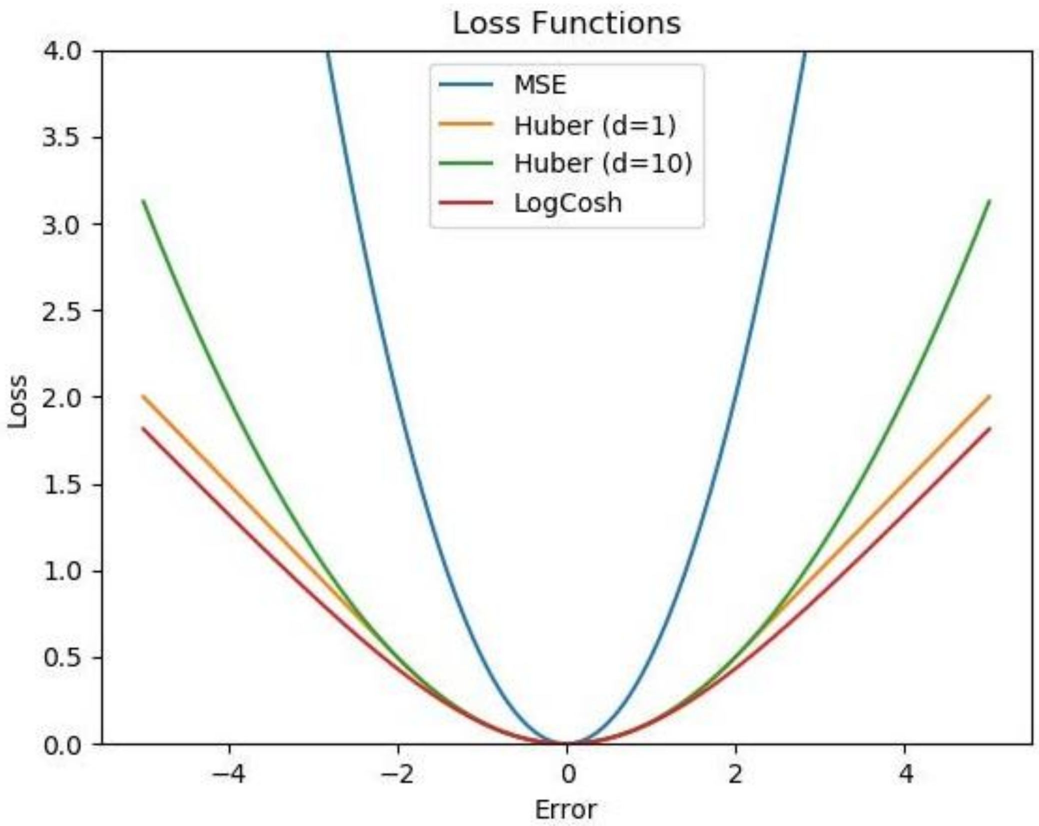


https://www.3blue1brown.com/lessons/backpropagation

Loss function plot



MSE - MAE - Huber



MSE - Huber - lgcosh

https://htcondor.readthedocs.io/en/latest/users-manual/submitting-a-job.html

HTCondor is the job manager at ARTEMISA

```
# Comunicació amb HTCondor - el gestor de treballs dins d'Artemisa
condor_submit NNprova.sub # Carrega en la cua de treball el "submit" NNprova
condor_q                  # Mostra els treballs en cola
condor_rm 862149          # Elimina de la cola el treball 6862149
```

.sh file copy

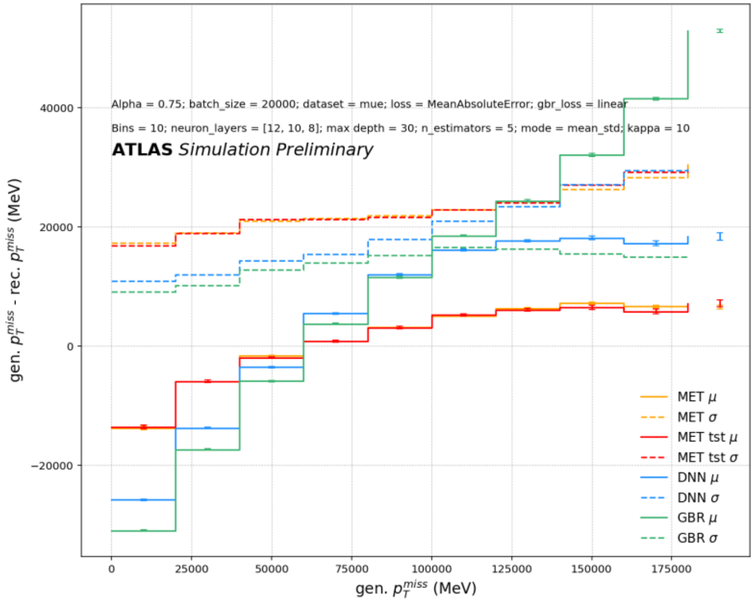
```
#!/bin/bash
python3 BDTprova.py
cp BDTprova.csv /lhome/ific/a/alexpg/
# Per tal de copiar l'output en la meua carpeta
```

,sub file example:

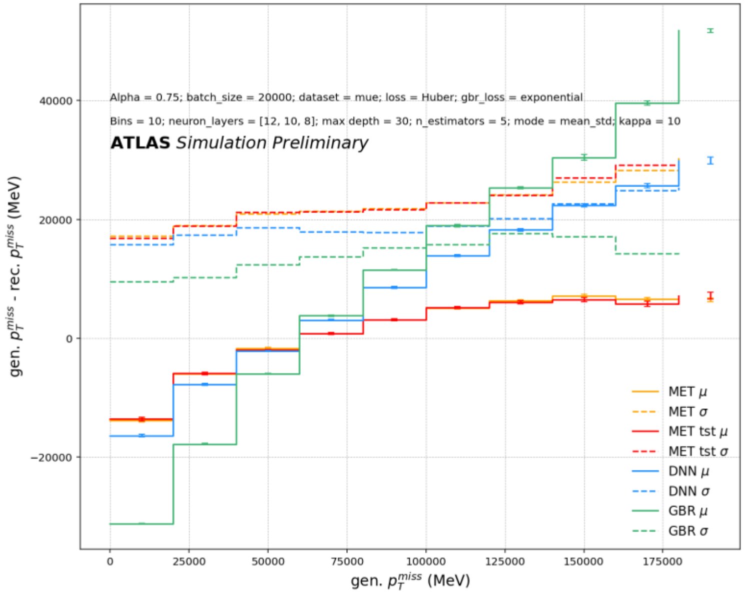
```
universe = vanilla
// No tinc massa clar quines opcions hi ha disponibles, però sembla que en general aquest és correcte.

executable              = BDTprova.sh
// Executarà l'script BDTprova.sh (podem considerar arxius sh bash o directament python)

log                     = condor_logs/BDTprova.log
output                  = condor_logs/BDTprova.$(Cluster).$(Process).out
error                   = condor_logs/BDTprova.$(Cluster).$(Process).err
// Guarda un registre del treball en log, el que eix en terminal en .out i els errors en .err

notify_user = alejandro.perez@ific.uv.es // Es pot demanar que notifique al usuari quan acabe
notification = always

transfer_output_files   = BDTprova.csv   // Podem guardar els files en un arxiu .csv
queue             // Ací podem incloure quantes vegades volem executar el treball (queue 3)
```

INSTITUT DE FÍSICA CORPUSCULAR

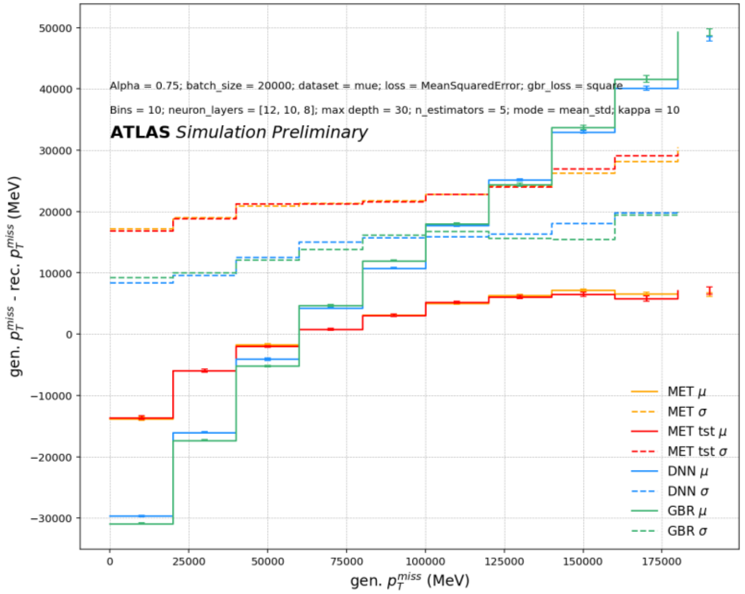
Loss function exploration → Custom loss definition → Custom weights



DNN loss: MeanAbsoluteError
AdaBoost loss: linear



DNN loss: Huber
AdaBoost loss: exponential



DNN loss: MeanSquaredError
AdaBoost loss: square

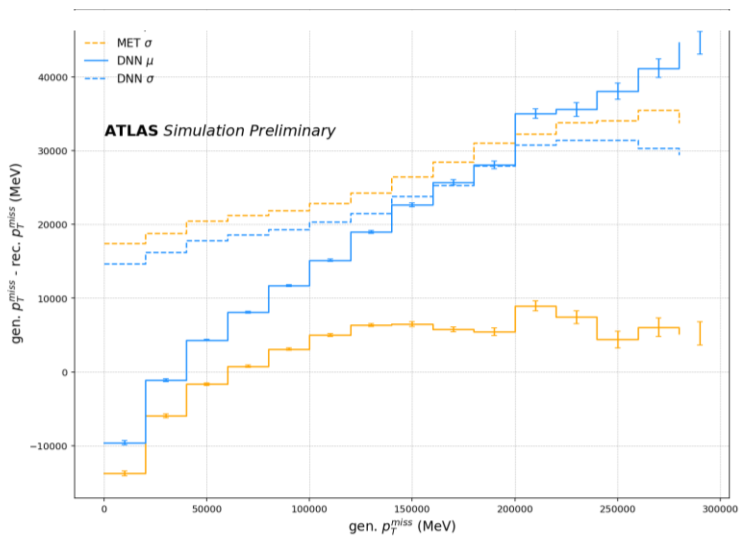
Use Case 2: Architecture exploration II

Loss function exploration

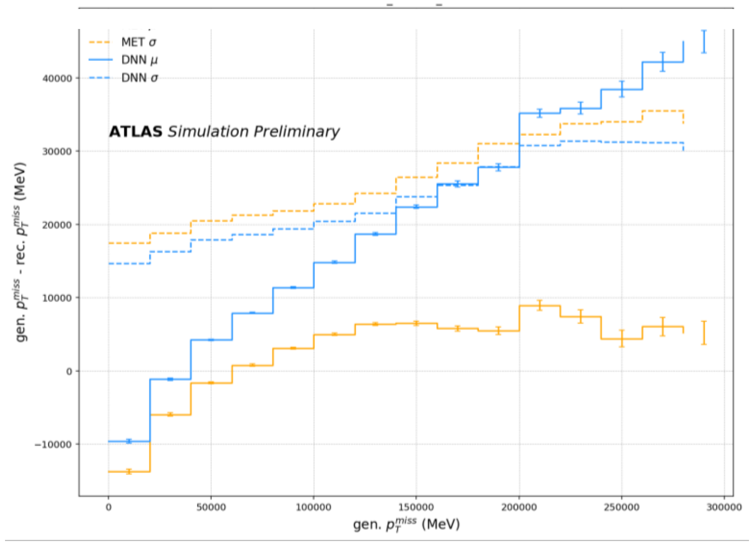
Custom loss definition

Custom weights

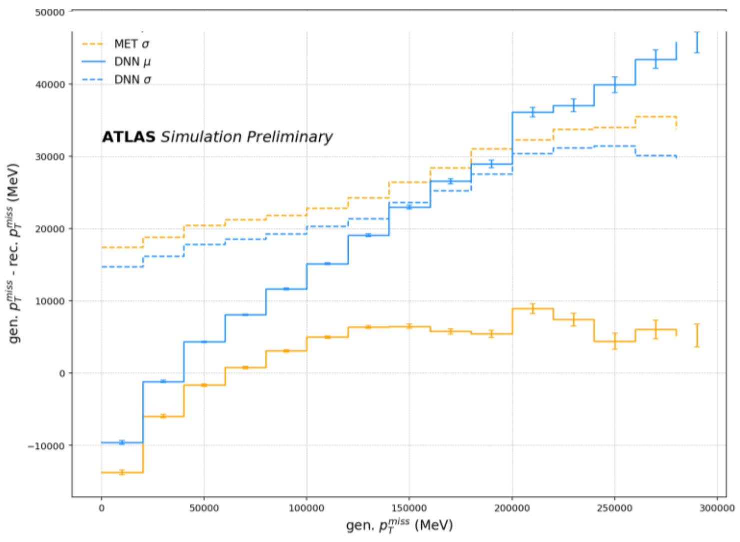
$$C_{custom} = \lambda \cdot \sum_{b} (\text{ mean } [a]_b)^2 + C_{MSE}$$



$\lambda = 0.1$



$\lambda = 1$

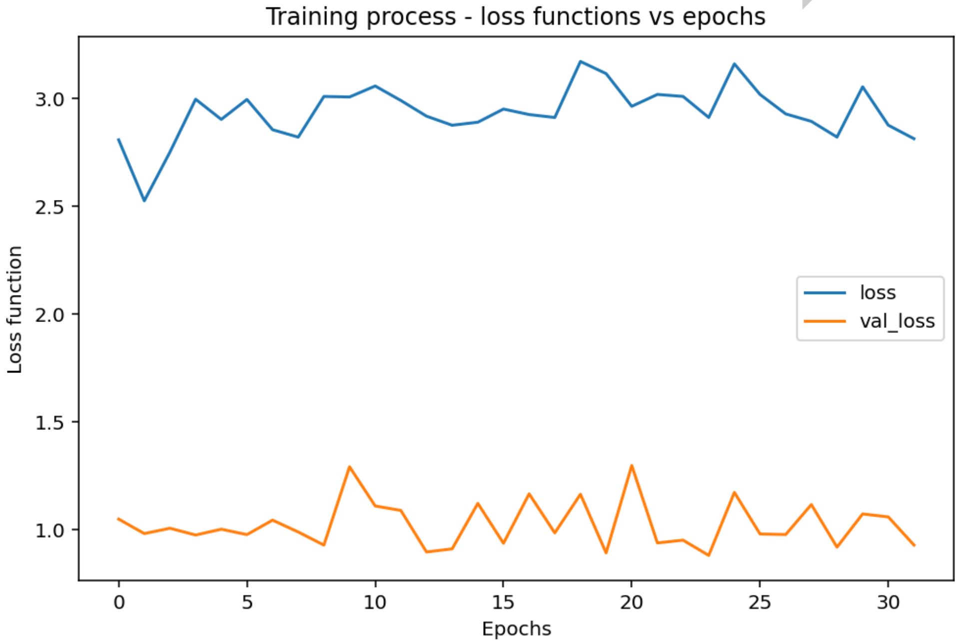


$\lambda = 1000$

Loss function exploration

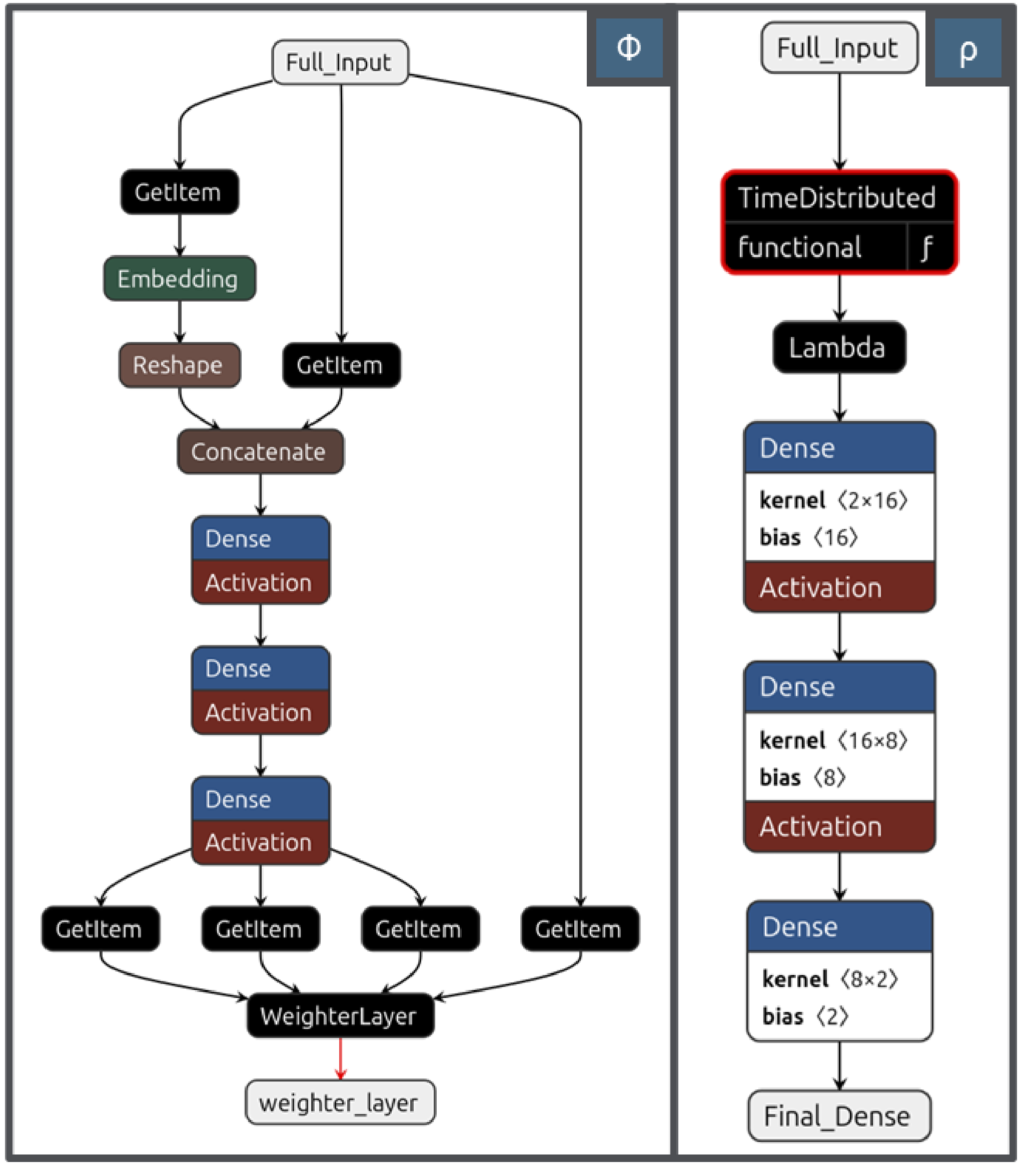
Custom loss definition

Custom weights



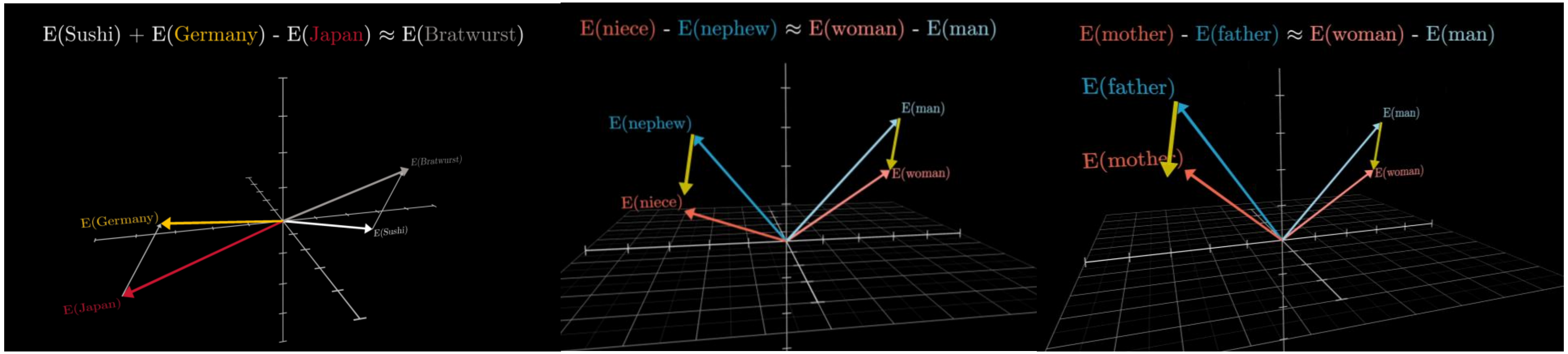
[Determination of missing transverse momentum using multivariate methods for a differential top pair production cross section measurement at CMS]()

Use Case 2: DeepSets exploration – Key findings

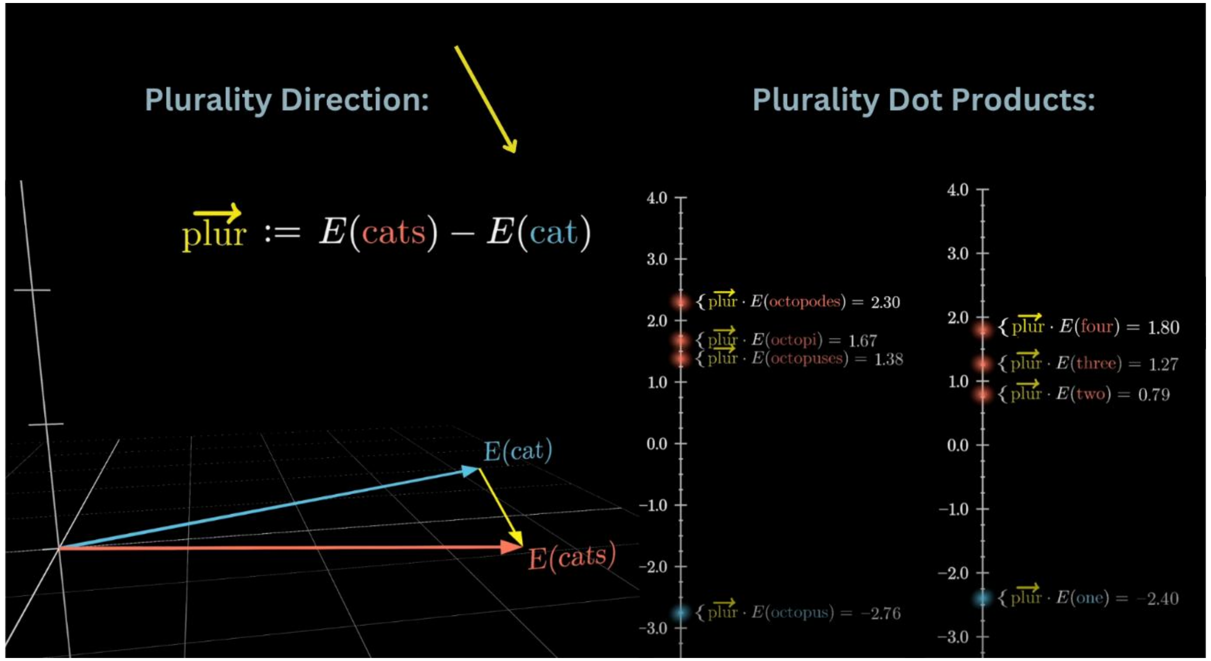


- Inclusion of all simulated neutrinos may be beneficial
- Comparison between ATLAS and CMS might be key
- Lower-level reconstruction may yield better results

Embedding examples

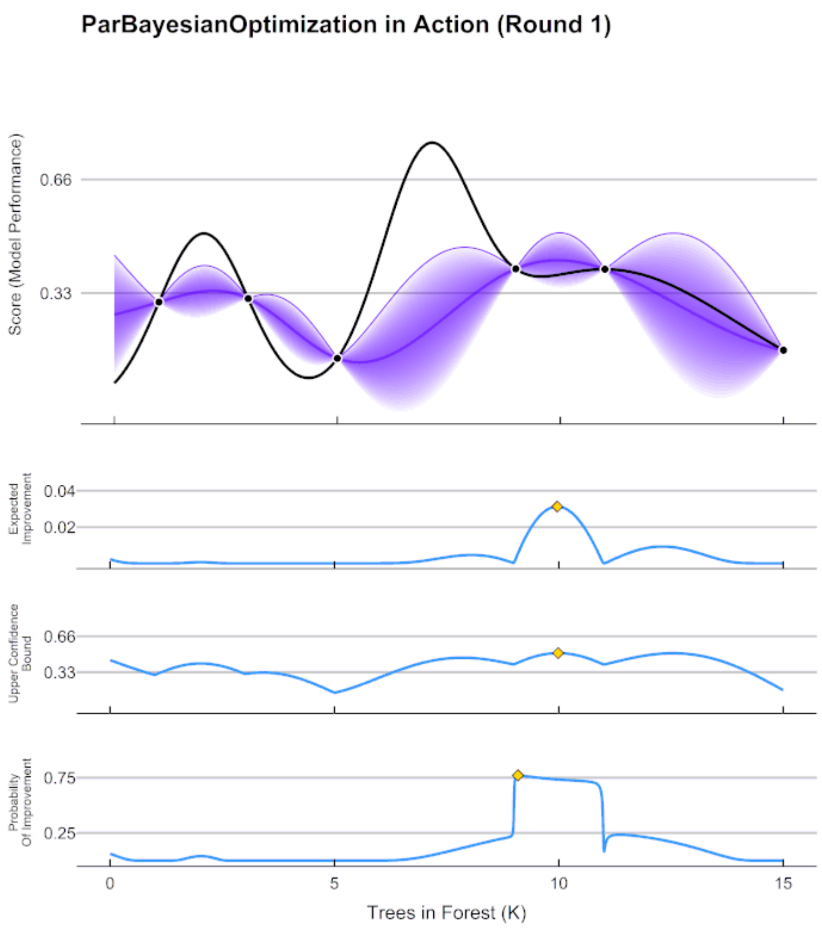


Embedding word differences using Chat-GPT2 - https://www.3blue1brown.com/lessons/gpt

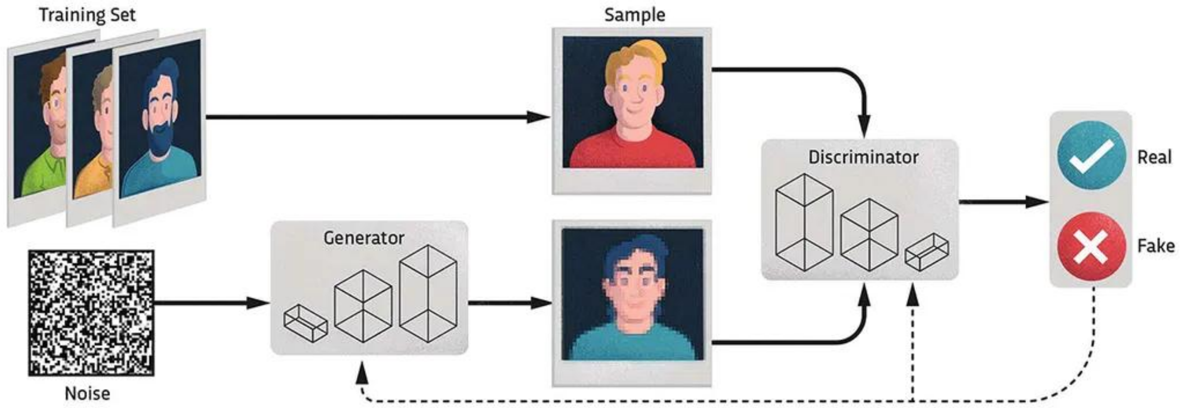


Embedding concept of “plurality” using Chat-GPT2 - https://www.3blue1brown.com/lessons/gpt

Exploration alternative - BayesOpt



Fast simulation alternatives – GAN and VAE



GAN - diagram

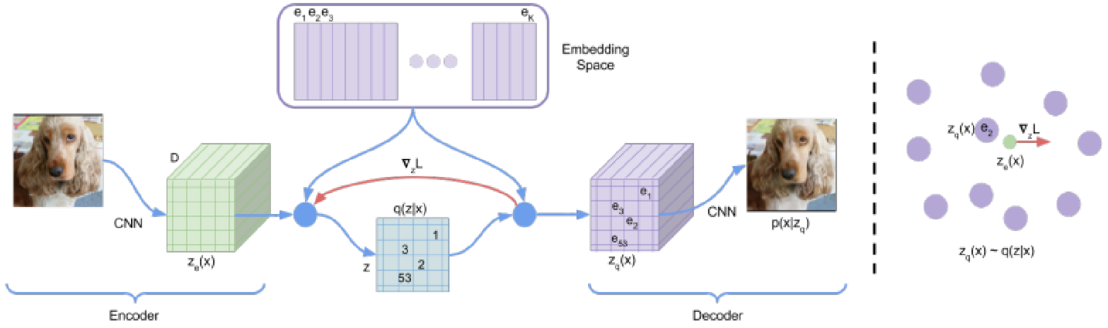


Figure 1: Left: A figure describing the VQ-VAE. Right: Visualisation of the embedding space. The output of the encoder $z(x)$ is mapped to the nearest point e_2. The gradient $\nabla_z L$ (in red) will push the encoder to change its output, which could alter the configuration in the next forward pass.

VAE - diagram

POWHEG-HVQ + PYTHIA8 set up – CMS and ATLAS

Setting name	Setting description	CMS default	ATLAS default	Common Proposal
POWHEG				
qmass	top-quark mass [GeV]	172.5	172.5	172.5
twidth	top-quark width [GeV]	1.31	1.32	1.315
hdamp	first emission damping parameter [GeV]	237.8775	258.75	250
wmass	$W^{\pm}$ mass [GeV]	80.4	80.3999	80.4
wwidth	$W^{\pm}$ width [GeV]	2.141	2.085	2.11
bmass	b-quark mass [GeV]	4.8	4.95	4.875
PYTHIA 8				
	PYTHIA 8 version	v240	v230	v240 (CMS) v244 (ATLAS)
	Tune	CP5	A14	Monash
PDF:pSet	LHAPDF6 parton densities to be used for proton beams	NNPDF31_nnlo _as_0118	NNPDF23_lo _as_0130_qed	NNPDF23_lo _as_0130_qed
TimeShower:alphaSvalue	Value of α_s at Z mass scale for Final State Radiation	0.118	0.127	0.1365
SpaceShower:alphaSvalue	Value of α_s at Z mass scale for Initial State Radiation	0.118	0.127	0.1365
MPI:alphaSvalue	Value of α_s at Z mass scale for Multi-Parton Interaction	0.118	0.126	0.130
MPI:pT0ref	Reference p_T scale for regularizing soft QCD emissions	1.41	2.09	2.28
ColourReconnection:range	Parameter controlling colour reconnection probability	5.176	1.71	1.80

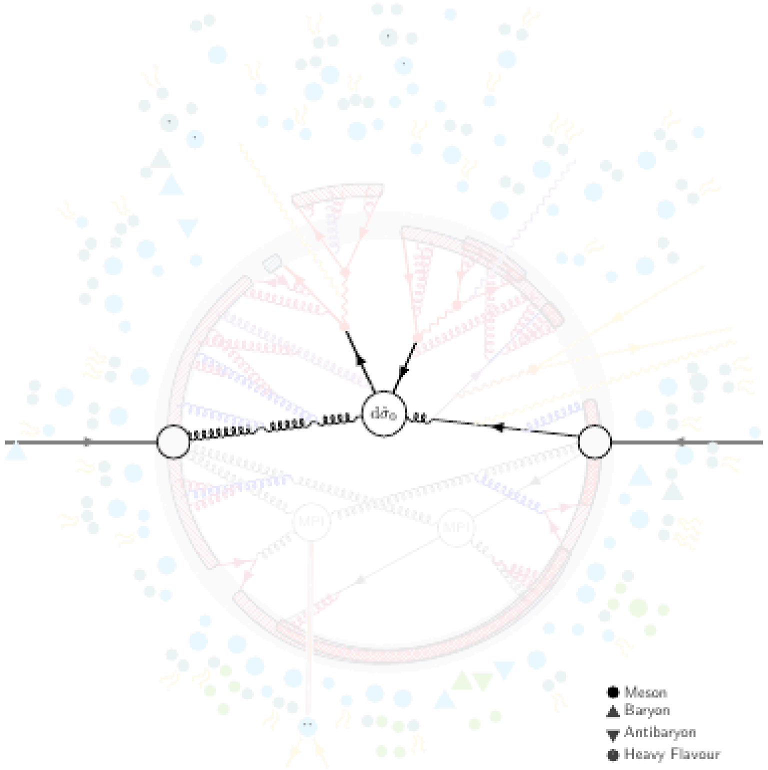
Negro, Giulia. (2022). Top Quark Modelling and Tuning in ATLAS and CMS. 10.48550/arXiv.2201.03517.

Physics modelled within Pythia 8

Classify event generation in terms of "hardness"

1. Hard Process (here $t\bar{t}$)

[figure credit: P. Skands]



2

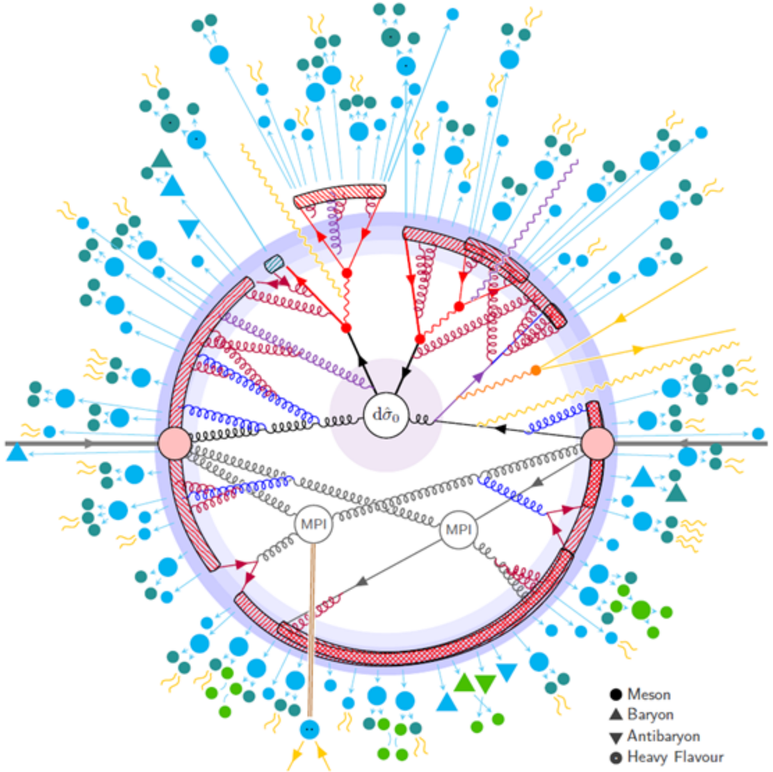
INSTITUT DE FÍSICA CORPUSCULAR

Physics modelled within Pythia 8

Classify event generation in terms of "hardness"

1. Hard Process (here $t\bar{t}$)
2. Resonance decays ($t, Z, \ldots$)
3. Matching, Merging and matrix-element corrections
4. Multiparton interactions
5. Parton showers: ISR, FSR, QED, Weak
6. Hadronization, Beam remnants
7. Decays, Rescattering

[figure credit: P. Skands]

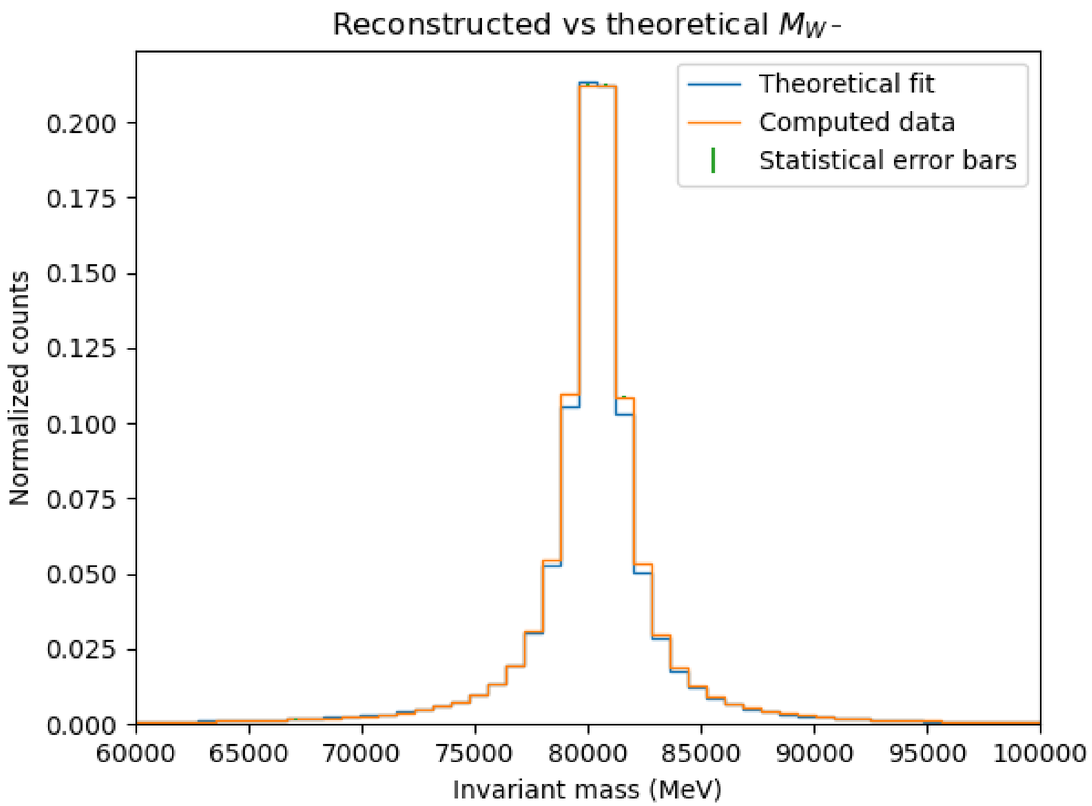


2

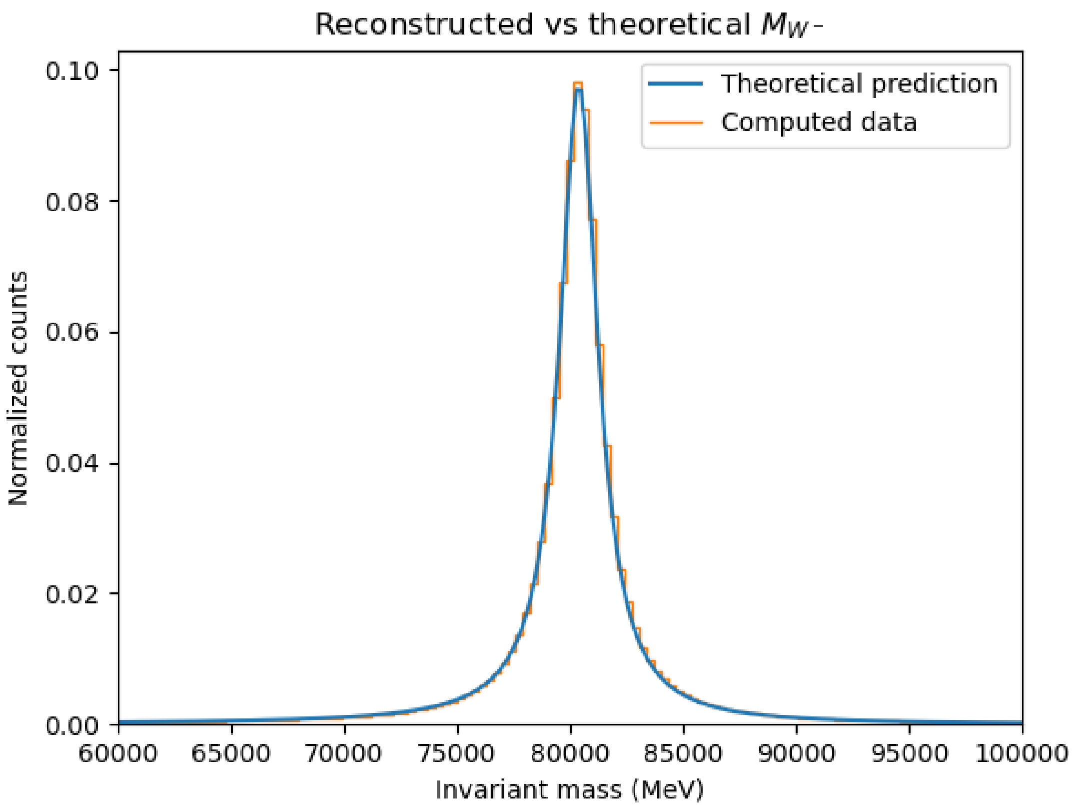
Table 6.2. Cross sections, at next-to-leading-order in QCD, for top-quark production via the strong interaction at the Tevatron and the LHC [14]. Also shown is the percentage of the total cross section from the quark-antiquark-annihilation and gluon-fusion subprocesses.

	σ_{NLO} (pb)	$q\bar{q} \to t\bar{t}$	$gg \to t\bar{t}$
Tevatron ($\sqrt{s} = 1.8$ TeV $p\bar{p}$)	$4.87 \pm 10\%$	90%	10%
Tevatron ($\sqrt{s} = 2.0$ TeV $p\bar{p}$)	$6.70 \pm 10\%$	85%	15%
LHC ($\sqrt{s} = 14$ TeV pp)	$803 \pm 15\%$	10%	90%

M_W statistical error



Successful attempt with Run 2 data at CMS



Pileup resilience studies at CMS

THANK YOU FOR YOUR ATTENTION

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
VNIVERSITAT DE VALÈNCIA
INSTITUT DE FÍSICA CORPUSCULAR
EXCELENCIA SEVERO OCHOA