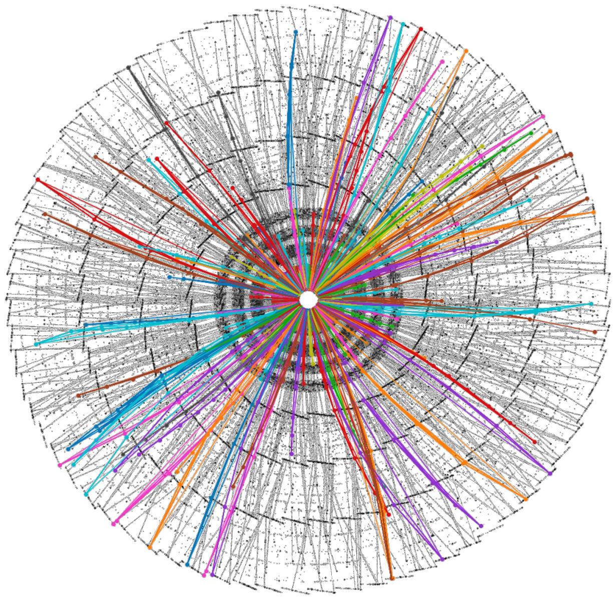
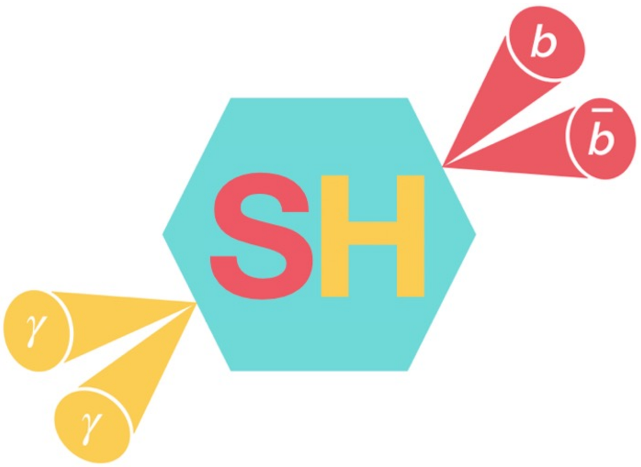
Charged particle trajectography using a novel geometric deep learning algorithm
&
Search for additional Higgs-boson-like particles in ATLAS Run 2 data.

Charline Rougier



Michel Daydé	*President of the jury*
Vladimir Gligorov	*Rapporteur*
Frédéric Machefert	*Rapporteur*
Amber Boehnlein	*Member of the jury*
Marumi Kado	*Member of the jury*
Jan Stark	*Supervisor*



19th September, MRV, Amphithéâtre 2

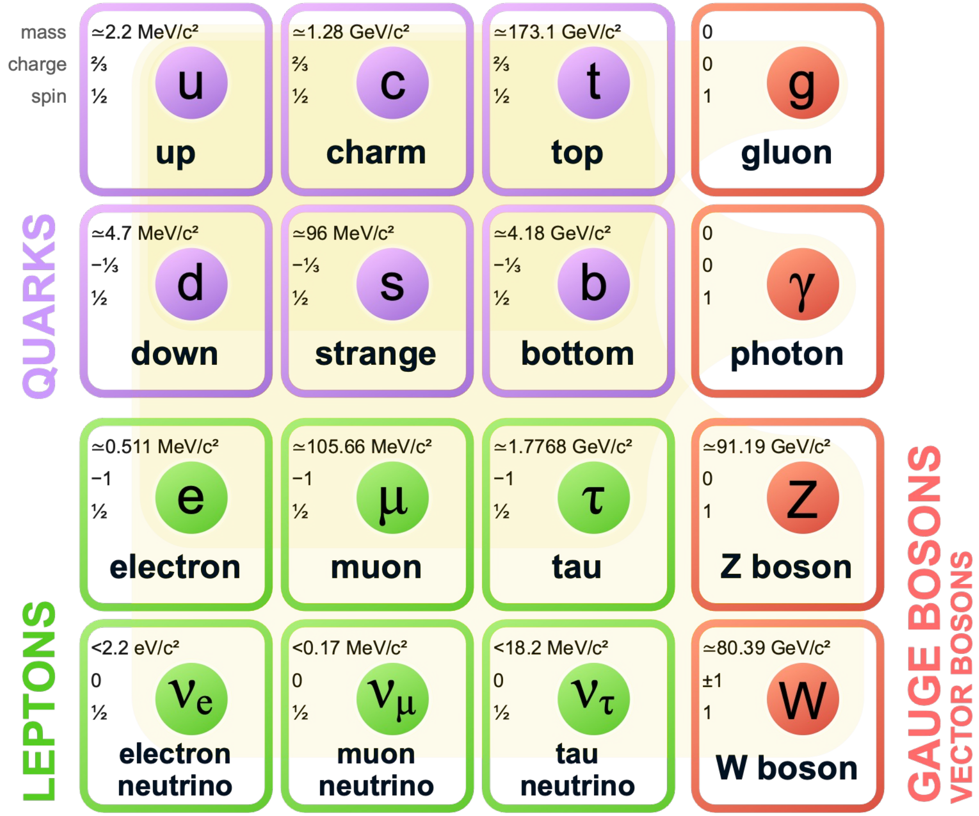
The Standard Model

The **S**tandard **M**odel (SM) is a gauge theory describing the electromagnetic, weak and strong interactions.

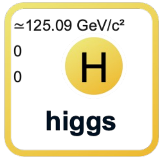
The matter is described by fermions:

- quarks,
- leptons.

Interactions are mediated by bosons.



$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi + \mathrm{h.c} + \bar{\psi}_i y_{ij}\phi\psi_j + h.c + \underbrace{|D^{\mu}\phi|^2 + V(\phi)}_{\text{Scalar sector}}$$



Discovered in 2012, jointly by ATLAS and CMS.

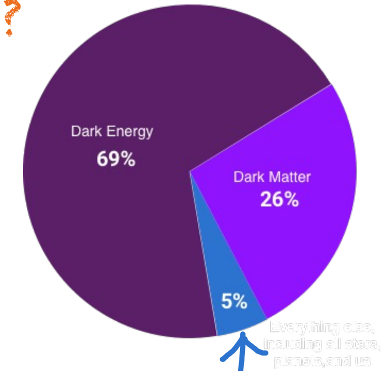
Beyond the Standard Model

We know the SM is not the end of the story. Many questions remain without clear answers:

Matter-antimatter asymmetry ?

Why is there so much more matter than antimatter in the universe ?

What is dark energy ?



What is dark matter ?

Included in the SM.

How to include gravity ?

Does the graviton exist ?

Hierarchy problem ?

Why the mass of the Higgs boson is much lighter than the Plank mass (10^{19} GeV) ?

Number of free parameters quite large

19 free parameters in the SM

Extension of the scalar sector

I am an experimentalist !

If there is an extended scalar sector, we can observe its effects:

- Modifications of the 125 GeV Higgs boson properties

  → Require precise measurements of the H properties.

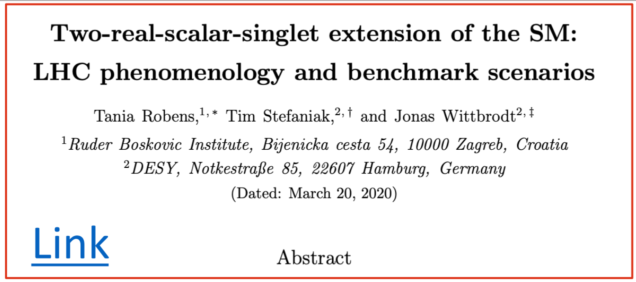
- Discovery of BSM decays of the 125 GeV Higgs boson

  → Searches are ongoing

- Direct discovery of new scalar particles

  → That's what I am looking for !

Additional scalars



Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios

Tania Robens,[1,*] Tim Stefaniak,[2,†] and Jonas Wittbrodt[2,‡]
[1]*Ruder Boskovic Institute, Bijenicka cesta 54, 10000 Zagreb, Croatia*
[2]*DESY, Notkestraße 85, 22607 Hamburg, Germany*
(Dated: March 20, 2020)

Link

Abstract

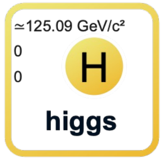
Many others BSM predict additional bosons

The most simple extension of the scalar sector introduces two new real singlets.

→ Two new additional scalar particles h_1 and h_2.

What can we infer about their expected properties ?

➢ Their properties highly depend on the specifics of the beyond the SM theory.



$$\left.\begin{array}{l} m_{\mathrm{H}} > m_{h_2} > m_{h_1} \\ m_{h_2} > m_{\mathrm{H}} > m_{h_1} \\ m_{h_2} > m_{h_1} > m_{\mathrm{H}} \end{array}\right\} ?$$

Additional scalars

Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios

Tania Robens,[1,*] Tim Stefaniak,[2,†] and Jonas Wittbrodt[2,‡]
[1]*Ruder Boskovic Institute, Bijenicka cesta 54, 10000 Zagreb, Croatia*
[2]*DESY, Notkestraße 85, 22607 Hamburg, Germany*
(Dated: March 20, 2020)

Link

Abstract

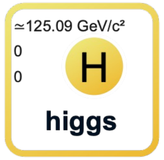
The most simple extension of the scalar sector introduces two new real singlets.

→ Two new additional scalar particles X and S.

Many others BSM predict additional bosons

What can we infer about their expected properties ?

➢ Their properties highly depend on the specifics of the beyond the SM theory.



$$\left.\begin{array}{l} m_H > m_{h_2} > m_{h_1} \\ m_{h_2} > m_H > m_{h_1} \\ m_{h_2} > m_{h_1} > m_H \end{array}\right\} ?$$

$H \to XX$

$H \to SS$

$h_2 \to XX$

$H \to SX$

$X \to HH$

$X \to SH$

$S \to HH$

$X \to SS$

non-exhaustive examples

Additional scalars

The most simple extension of the scalar sector introduces two new real singlets.

Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios

Tania Robens,[1,*] Tim Stefaniak,[2,†] and Jonas Wittbrodt[2,‡]

[1]*Ruder Boskovic Institute, Bijenicka cesta 54, 10000 Zagreb, Croatia*
[2]*DESY, Notkestraße 85, 22607 Hamburg, Germany*
(Dated: March 20, 2020)

Link

Abstract

Many others BSM predict additional bosons

→ Two new additional scalar particles X and S.

What can we infer about their expected properties ?

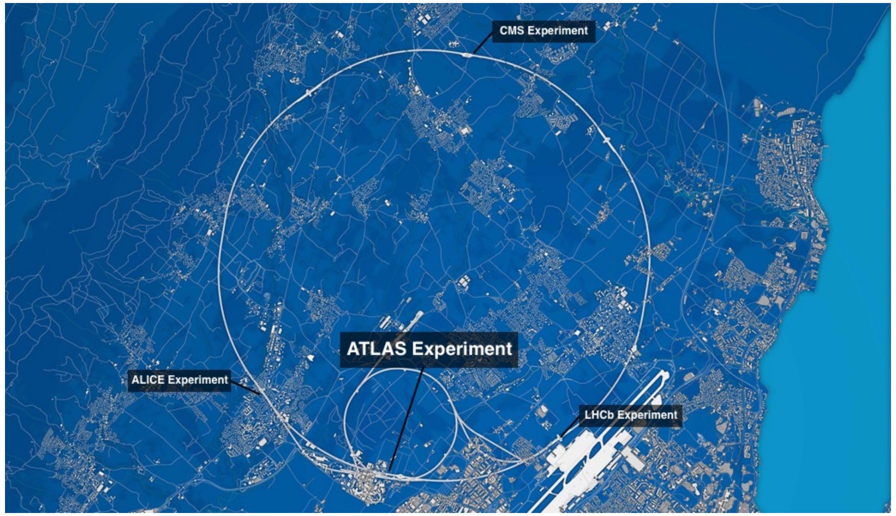
- Their properties highly depend on the specifics of the beyond the SM theory.

$X \rightarrow SH$

$m_X > m_H + m_S$

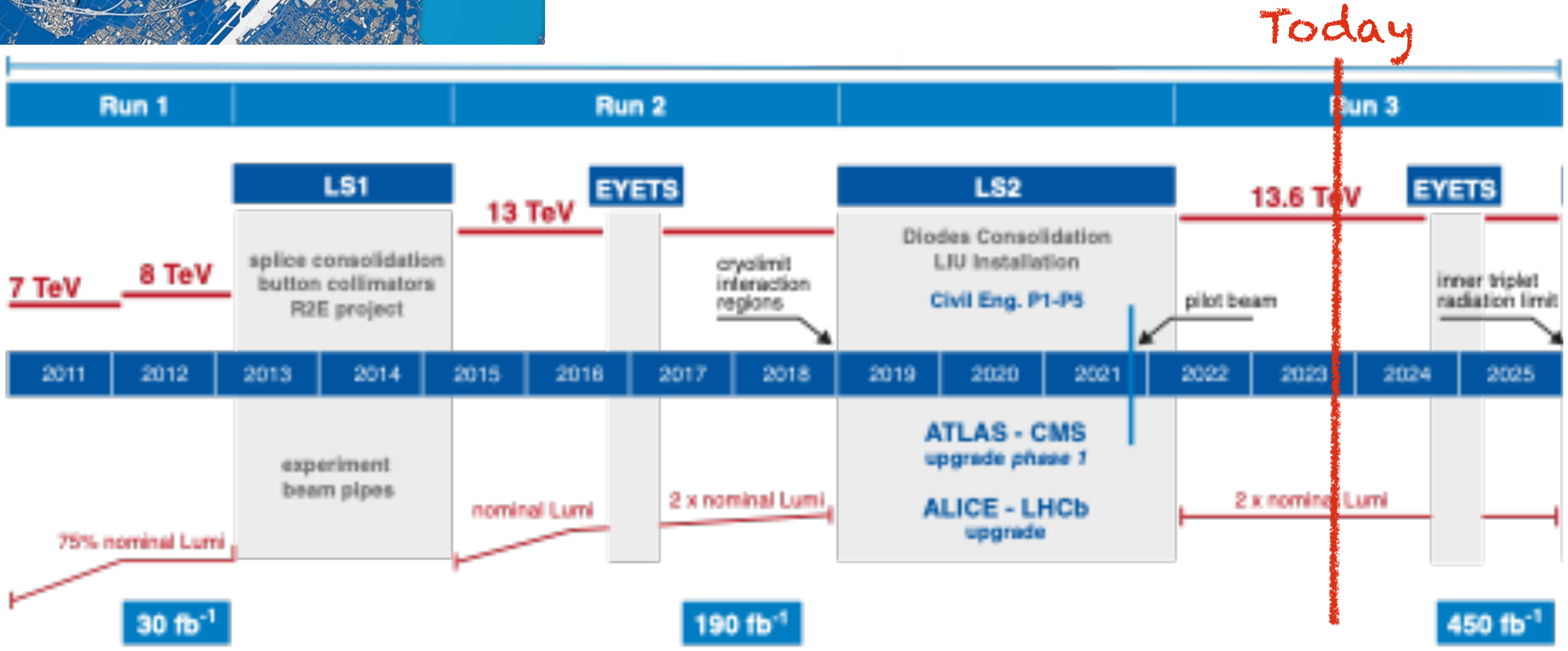
Let's hunt these new additional scalars

The Large Hadron Collider



Currently, the best tool to test the scalar sector is the **L**arge **H**adron **C**ollider (LHC):

- Designed to produce p-p collisions at $\sqrt{s} = 14$ TeV
- Also produce heavy ions collisions



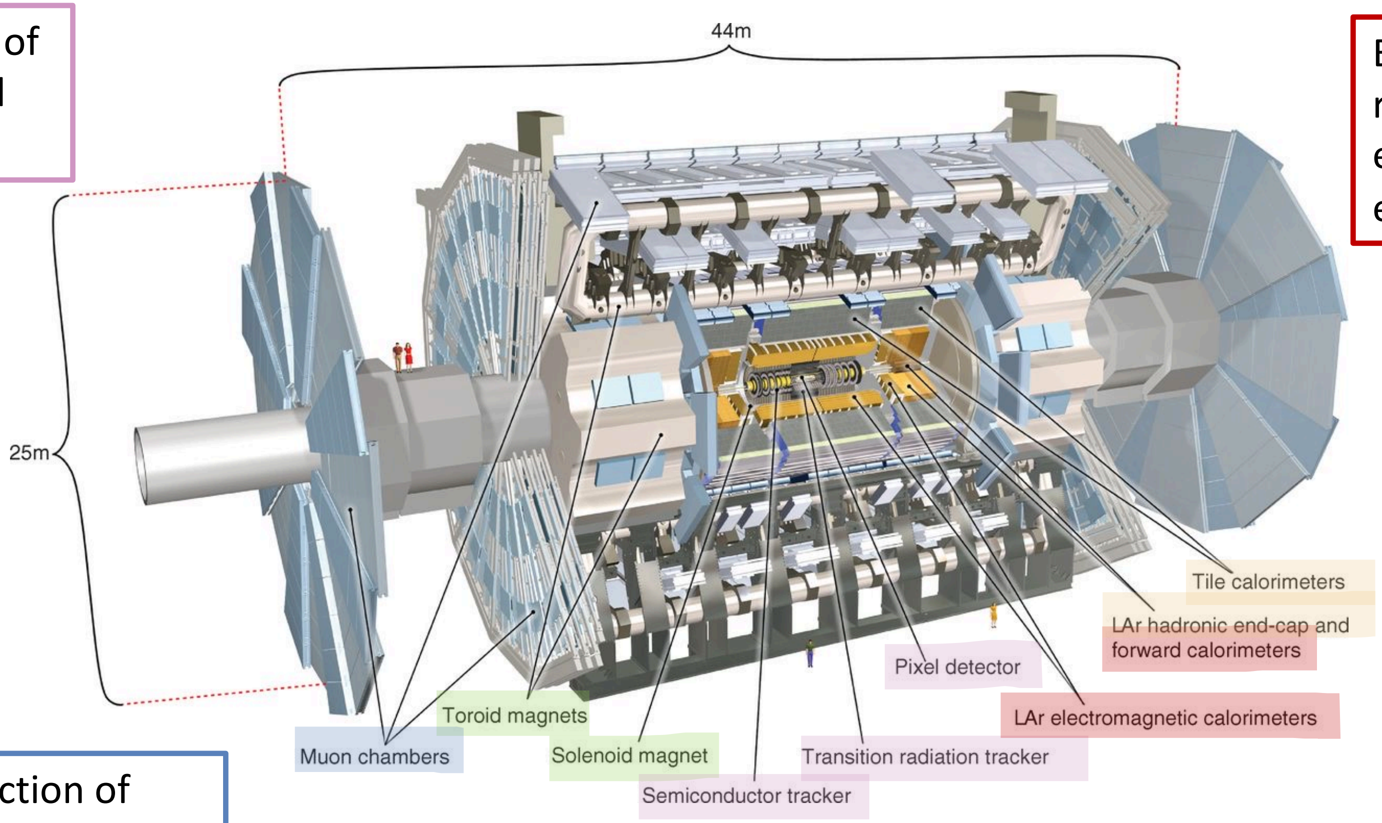
The ATLAS detector

The ATLAS (**A T**oroidal **L**HC **A**pparatu**S**) is a 4π detector with a onion-*layers* cylindrical shape.

Tracking system: reconstruction of charged particle trajectories and vertices.

Magnet system: bends the trajectories of charged particles for momentum measurement.

Muon spectrometer: reconstruction of muon trajectories.



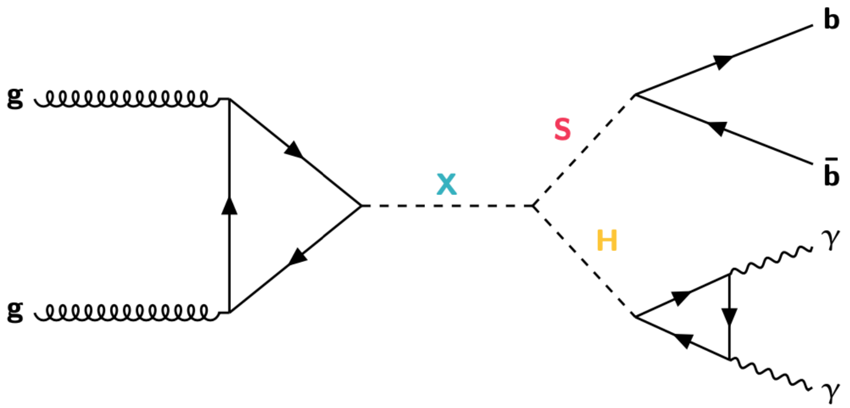
Electromagnetic calorimeter: reconstruction (direction and energy) of photons and electrons.

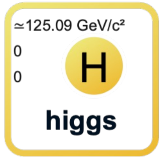
Hadronic calorimeter: reconstruction (direction and energy) of hadrons.

Search for $X \to S(b\bar{b})H(\gamma\gamma)$

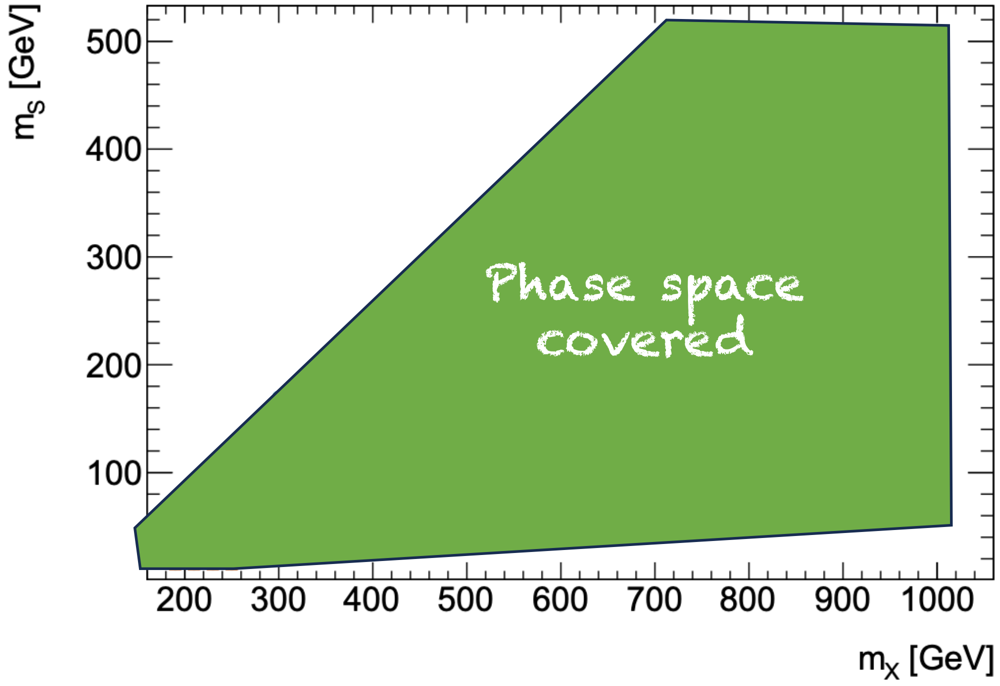
$$X \to SH \to \text{SM, with } m_X > m_S + m_H$$

The search is done using the ATLAS Run 2 dataset, corresponding to an integrated luminosity of $140\ \text{fb}^{-1}$ recorded at $\sqrt{s} = 13$ TeV.





with
- $m_S \in [15, 500]$ GeV
- $m_X \in [170, 1000]$ GeV



CMS already publicly released results on this analysis: no signal is seen.

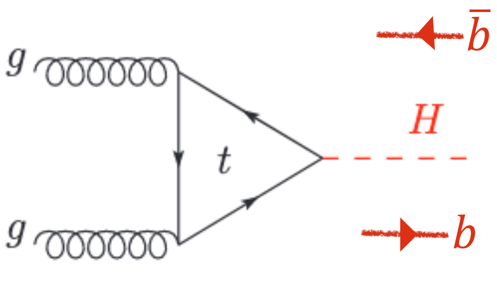
Search for X → $S(b\bar{b})H(\gamma\gamma)$: backgrounds

Final state is $\bar{b}b\gamma\gamma$.

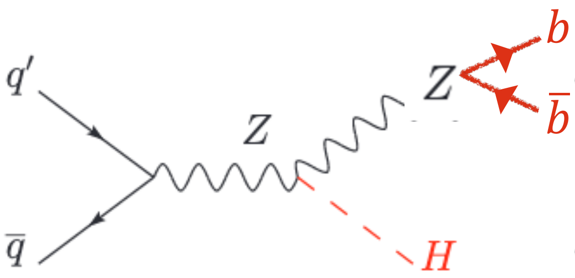
Other SM processes have the same final state: single Higgs production, di-Higgs production and continuum background.

Resonant backgrounds

ggH

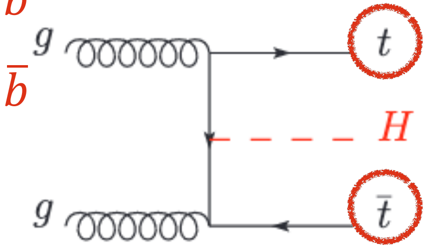


ZH



ttH

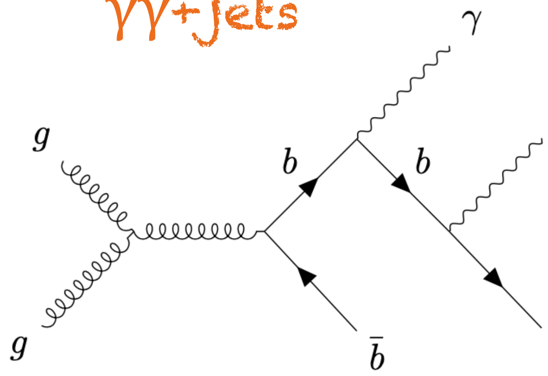
misidentified as b



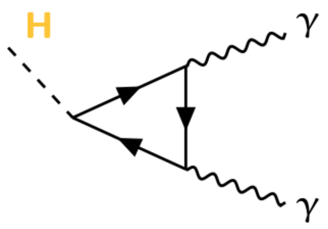
misidentified as b

Continuum background

γγ+jets



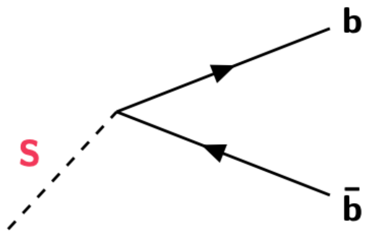
Selection of the signal



Di-photon

We select 2 photon candidates compatible with an Higgs decay:

- ✓ Apply identification criteria
- ✓ Reject photons from the decay $\pi^0 \rightarrow \gamma\gamma$
- ✓ Kinematic requirement to optimize the Higgs resonance peak: $\frac{p_T}{m_{\gamma\gamma}} > 0.35\ (0.25)$
- ✓ $m_{\gamma\gamma} \in [105{,}160]$ GeV

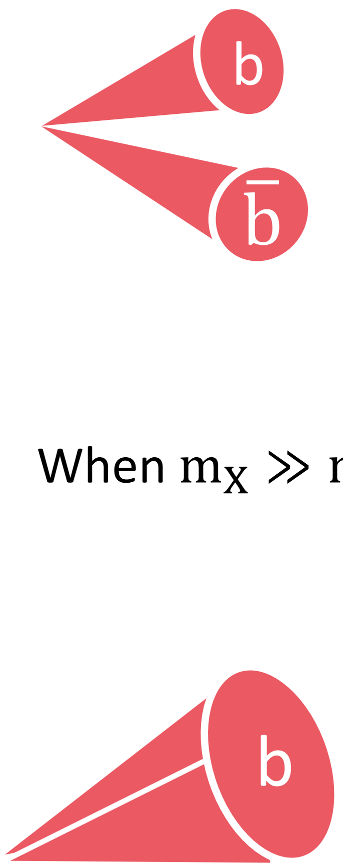


***b*-jet**

We reconstruct b-jet candidates:

- ✓ Reconstruct the stream of particles from the jet using the anti-k_t algorithm with R=0.4
- ✓ p_T > 25 GeV
- ✓ $N_{central\ jets} \in [2{,}5]$
- ✓ *b*-tagging: identify the flavor of the quarks

Selection of the signal: categories



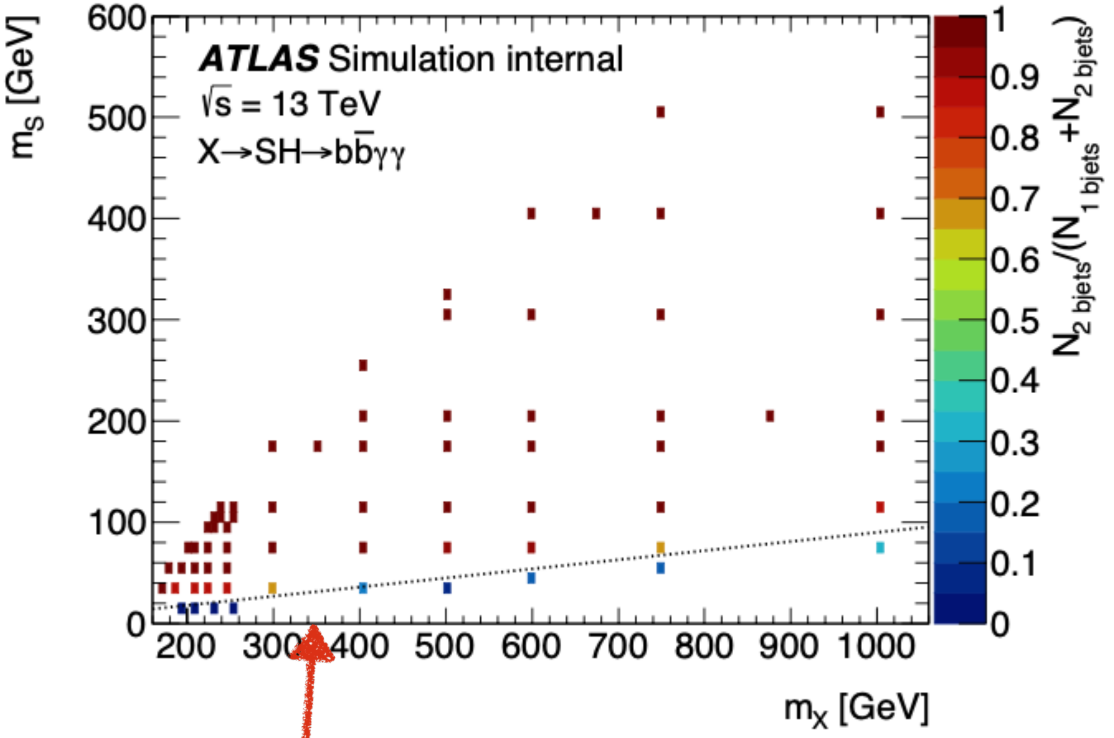
2 b-tagged category

Require **2** b-tagged @77% WP

When $m_X \gg m_S + m_H$:

1 b-tagged category

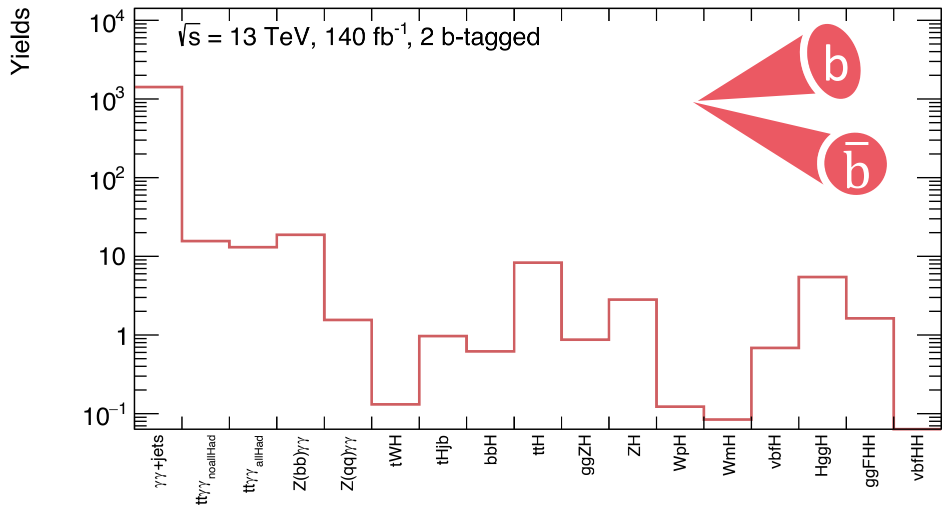
Required **1** b-tagged @77% WP



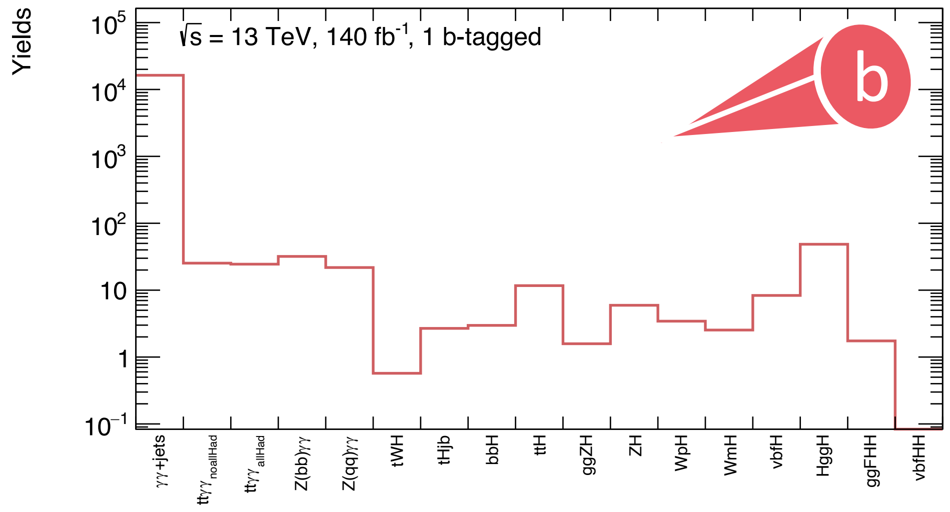
Limit between the two categories empirically found to be $\frac{m_S}{m_X} = 9\%$.

Multivariate discriminant

After the preselections:



~ 95% background from $\gamma\gamma + jets$

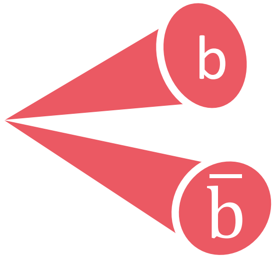


~ 99% background from $\gamma\gamma + jets$

Large phase space targeted by the analysis → design, train and use several Neural Networks ?

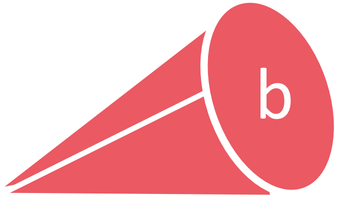
→ **P**arametrized **N**eural **N**etwork

Parametrized Neural Network (PNN)



- Parametrized with m_S and m_X
- Input features → $m_{b\bar{b}}$ and $\tilde{m}_{b\bar{b}\gamma\gamma}$

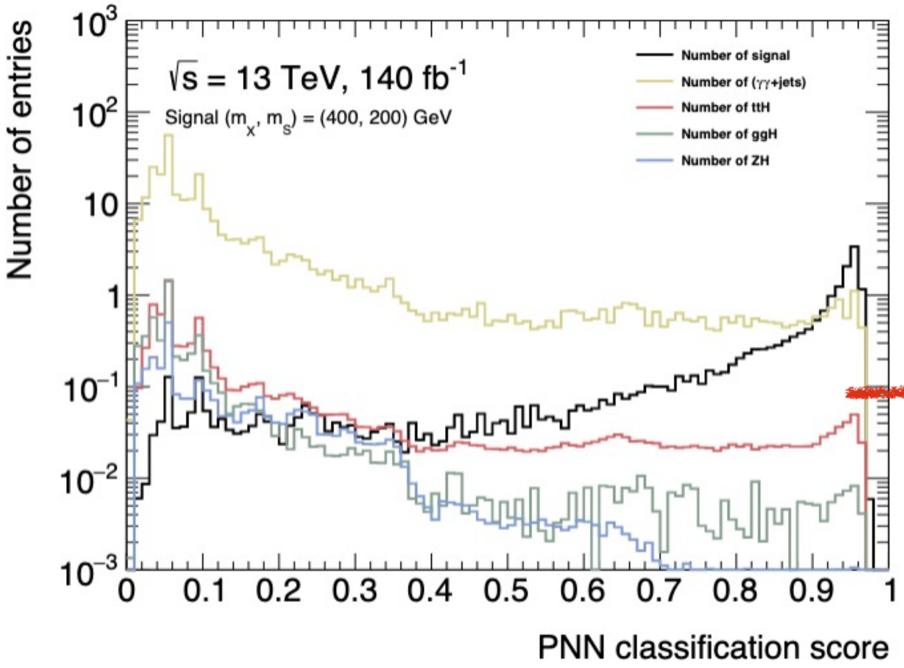
$$\widetilde{m}_{b\bar{b}\gamma\gamma} = m_{b\bar{b}\gamma\gamma} - (m_{\gamma\gamma} - 125\ \mathrm{GeV})$$



- Parametrized with m_X
- Input features → p_T^b and $\tilde{m}_{b\gamma\gamma}$

$$\widetilde{m}_{b\gamma\gamma} = m_{b\gamma\gamma} - (m_{\gamma\gamma} - 125\ \mathrm{GeV})$$

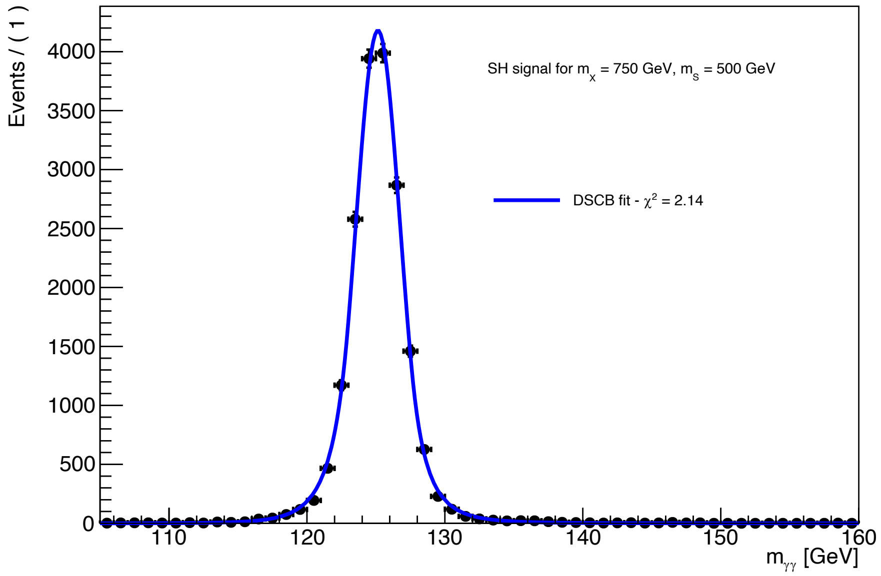
The PNN shape varies depending on the m_S and m_X hypothesis.



Apply for each targeted mass point a dedicated requirement on the PNN classification score.

Signal and background modelling

Signal:



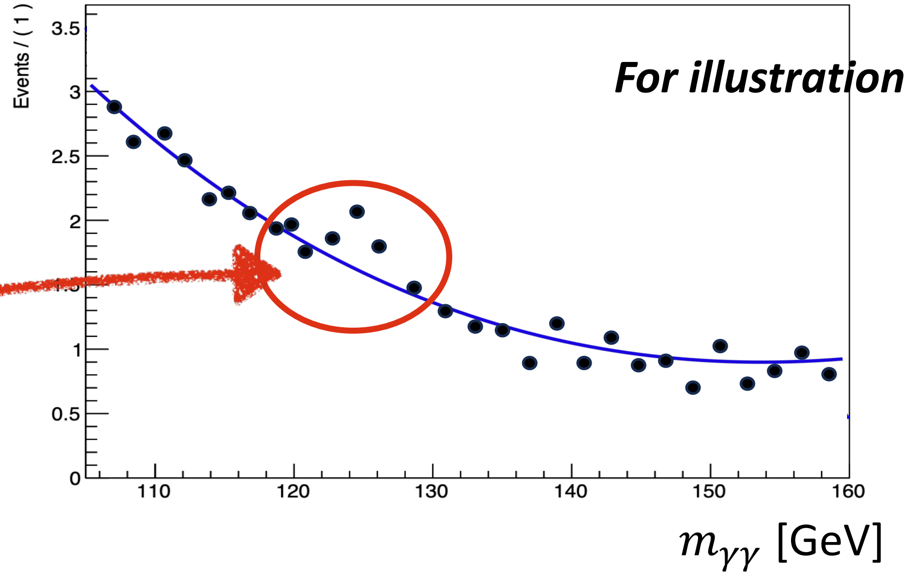
Modeled with a Double-Sided Cristal Ball:

→ Include also all resonant backgrounds.

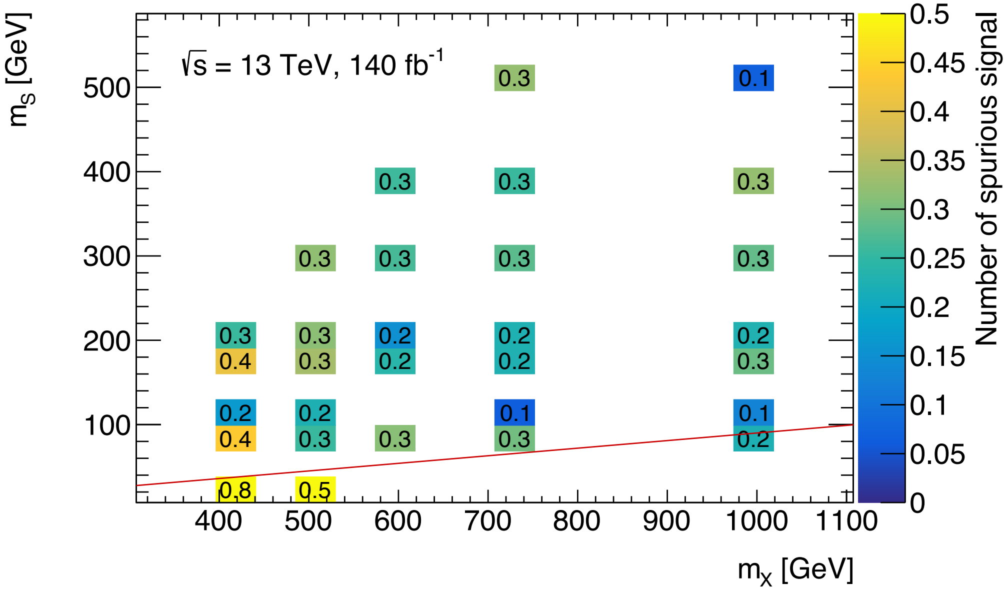
Continuum background:

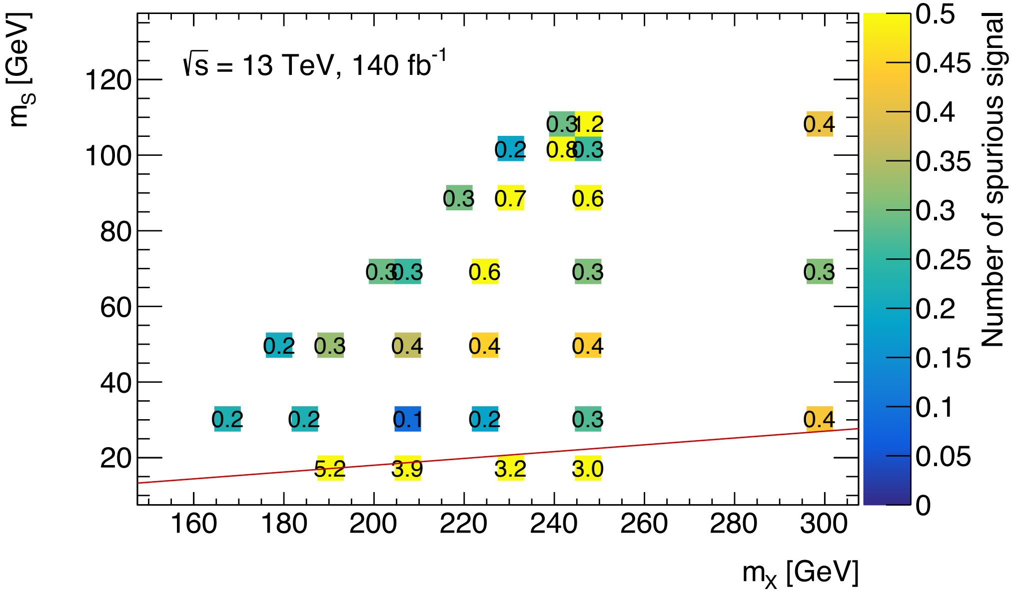
→ Exponential with a normalization taken on the data sideband.

Spurious signal



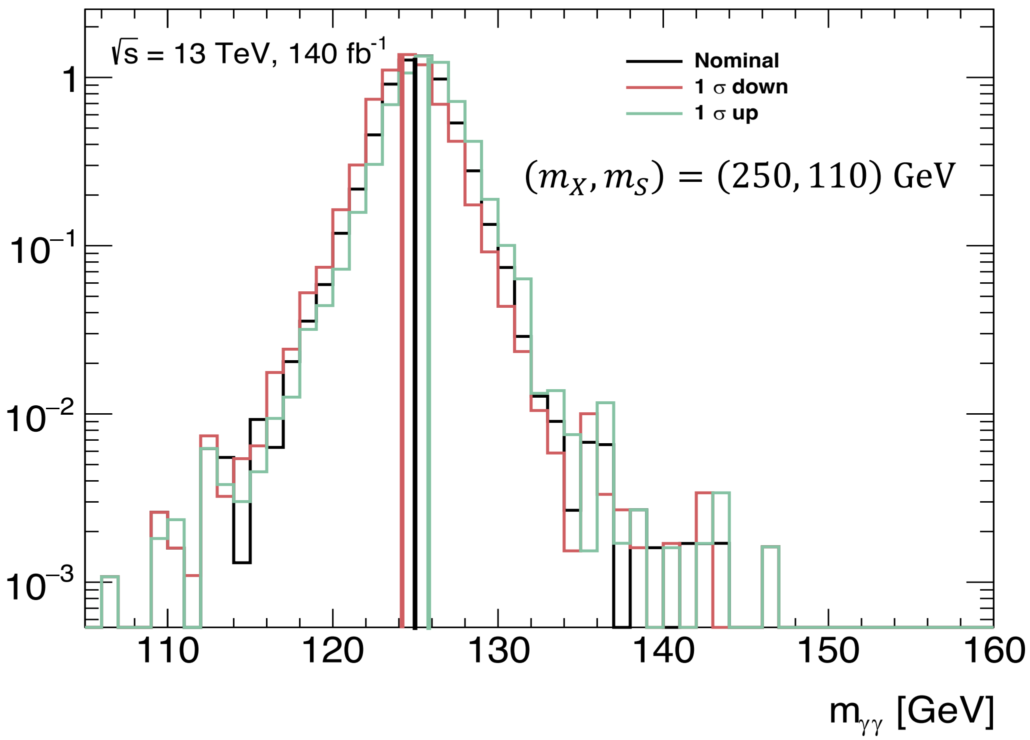
Statistical fluctuation not included in the modeling

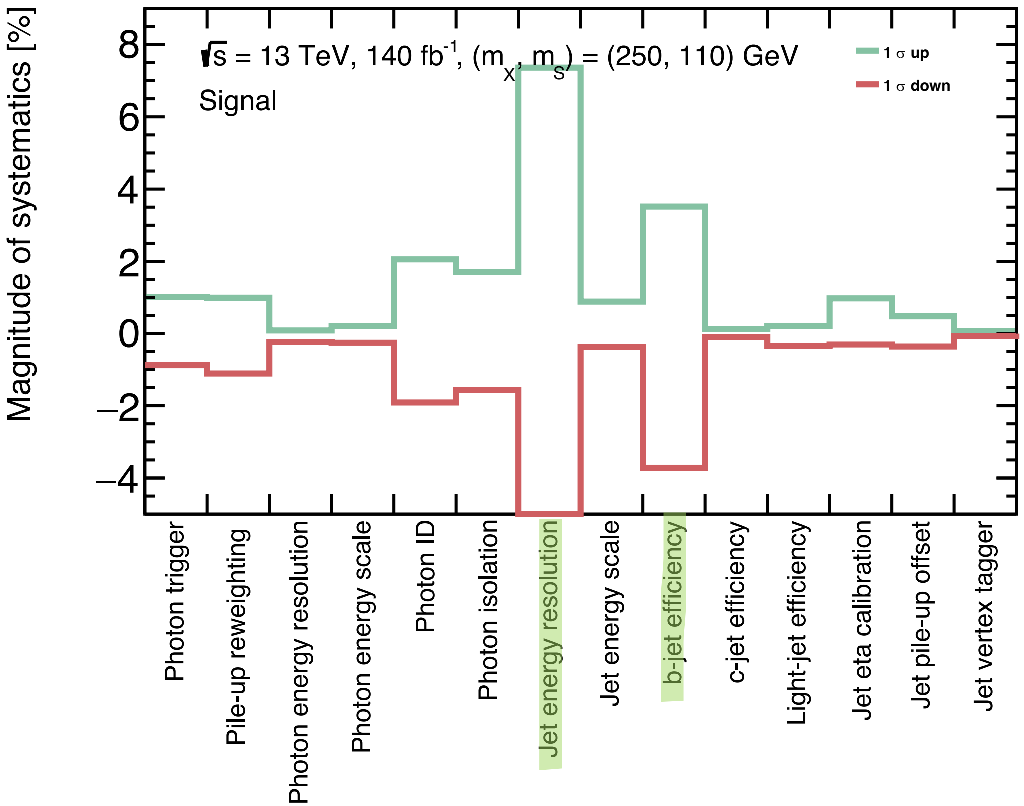




Systematic uncertainties

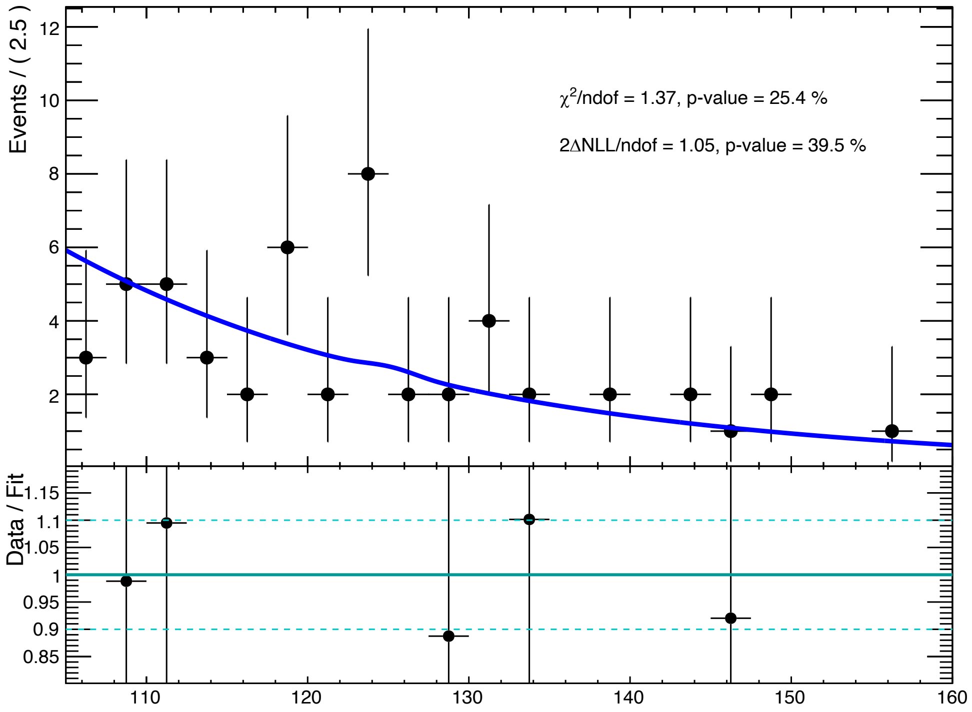
Photon energy scale





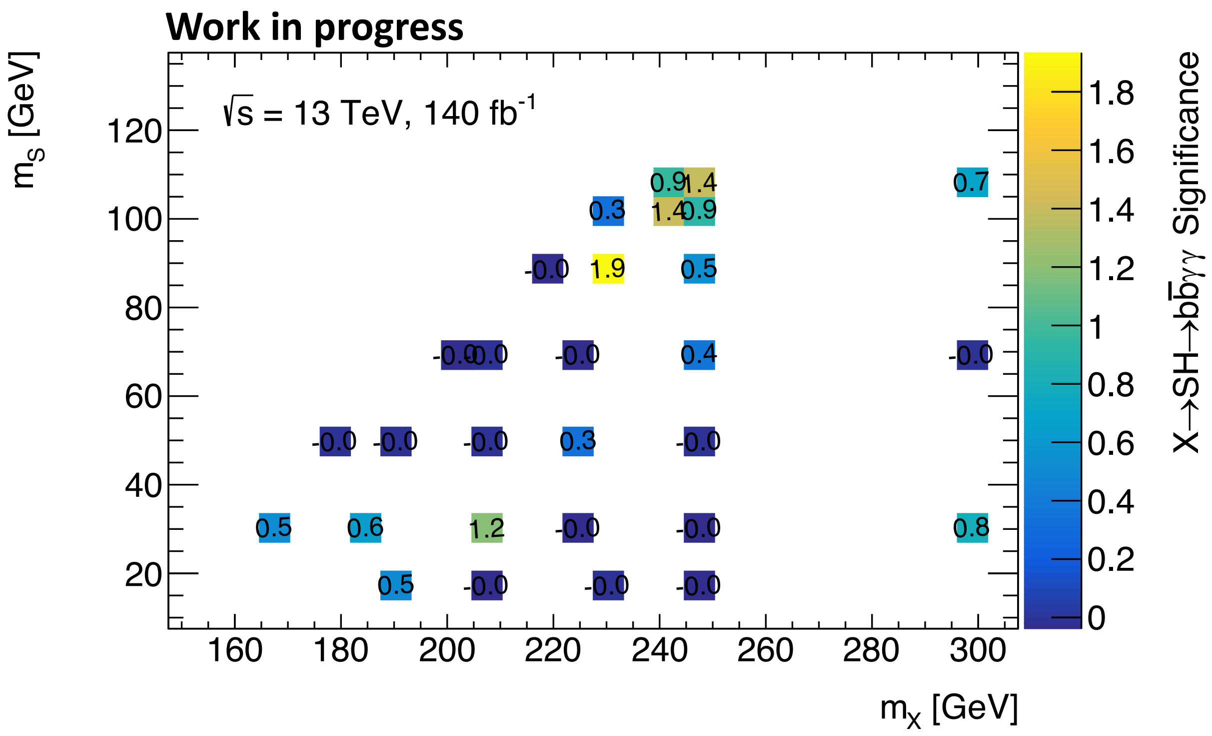
Mass spectrum

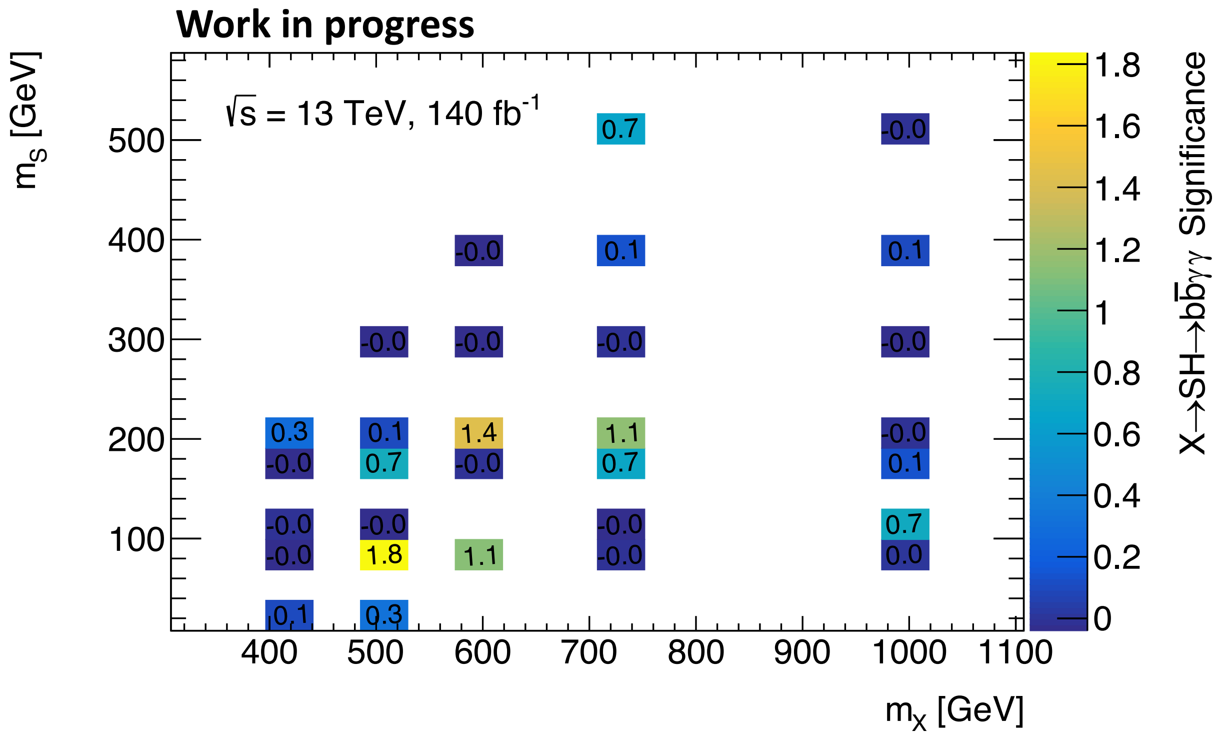
$(m_X, m_S) = (230, 90)$ GeV



$m_{\gamma\gamma}$ $[GeV]$

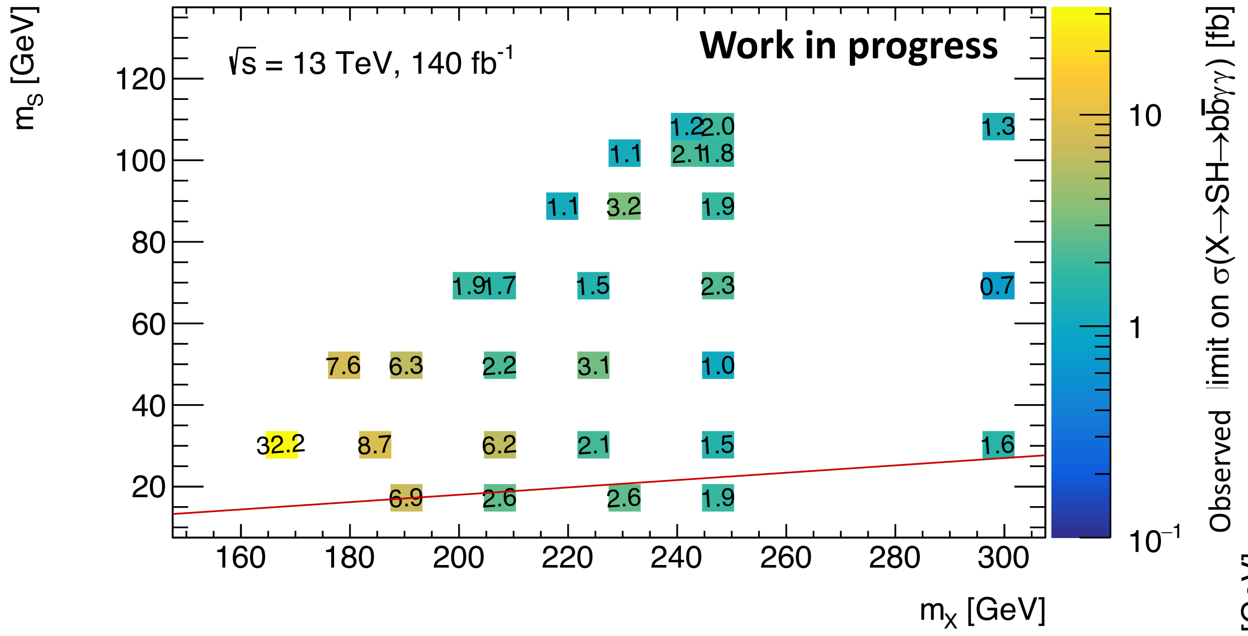
Observed local significance





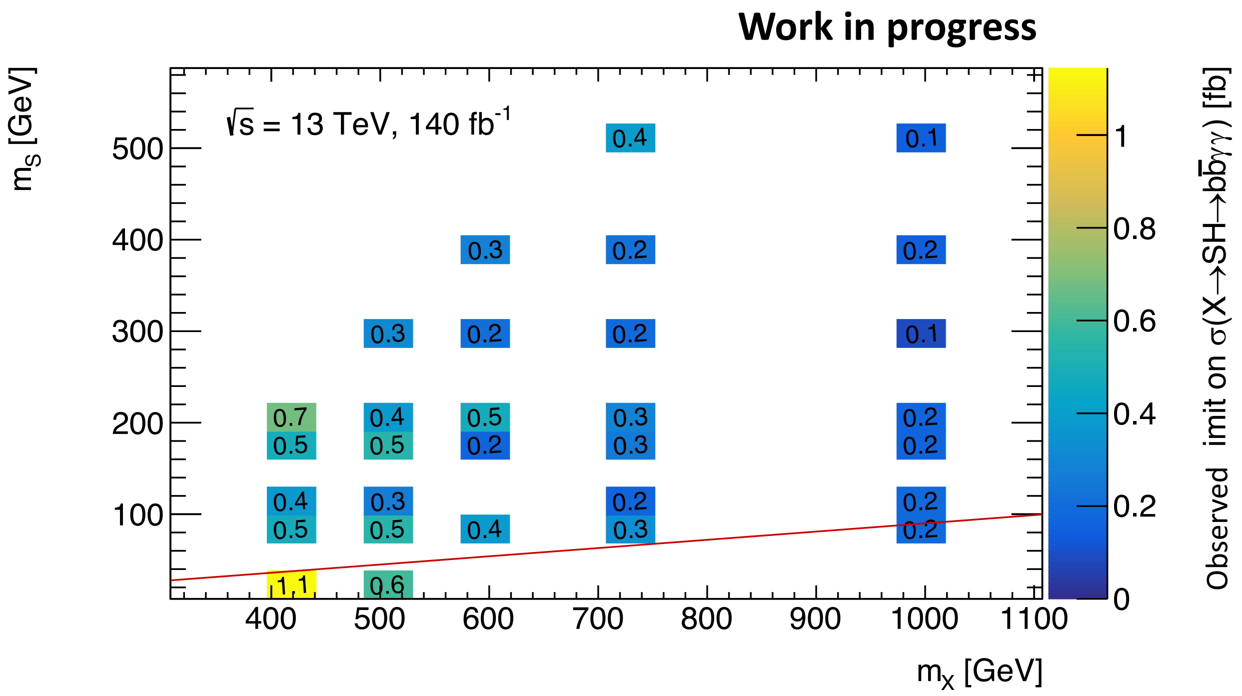
No significant excess is seen in ATLAS Run 2 Data for the $X \rightarrow S(b\bar{b})H(\gamma\gamma)$.

Upper limits on the cross-section



Observed 95% exclusion limits on $\sigma(pp \rightarrow X \rightarrow SH \rightarrow b\bar{b}\gamma\gamma)$.

The red line is the border between the 1 b-tagged and 2 b-tagged category.

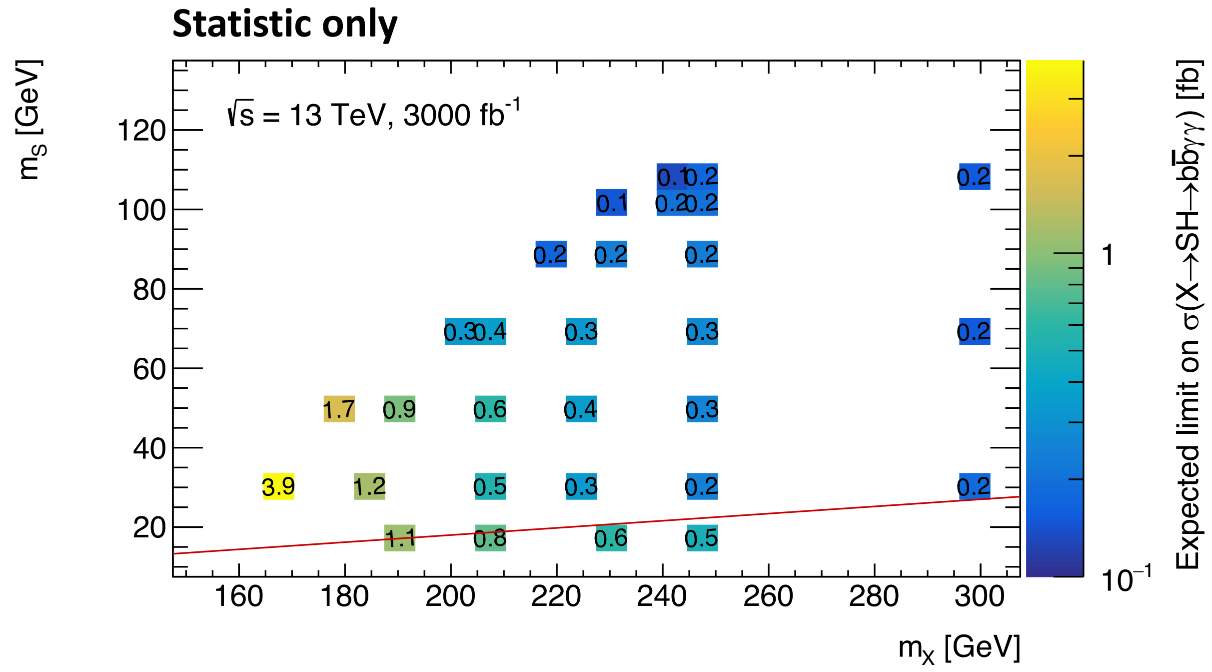


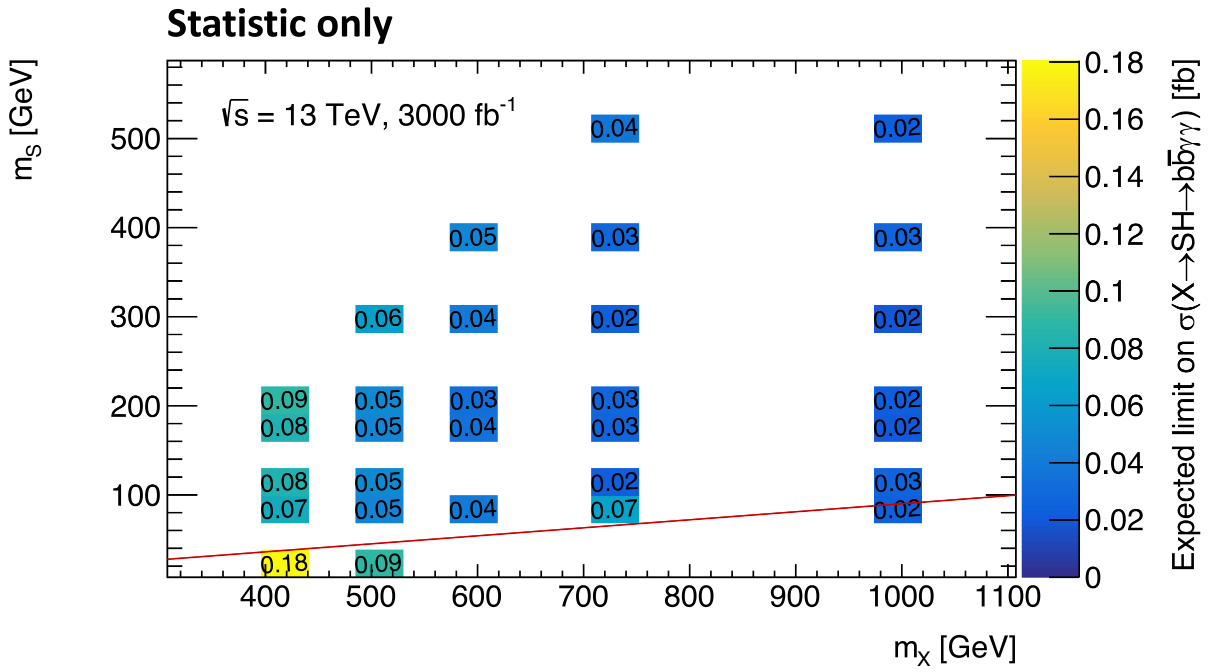
Limits are found to be better at higher masses due to a better signal efficiency and lower background.

Discussion

No $X \rightarrow S(b\bar{b})H(\gamma\gamma)$ signal using ATLAS Run 2 Data corresponding to an integrated luminosity of 140 fb^{-1}.

Does not mean that our X and S do not exist.

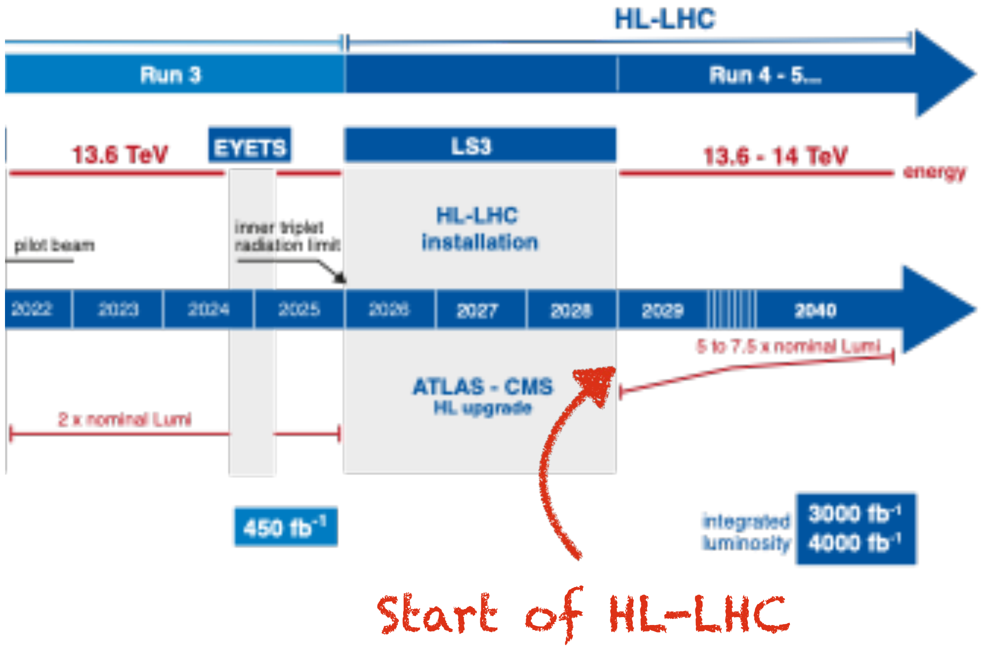




About a factor 10 of improvement in sensibility.

High-luminosity era

→ We need more data !



Upcoming upgrade of the LHC:

- Increase the luminosity of the accelerator by a factor 3
- Increase the data taking rate of the experiment

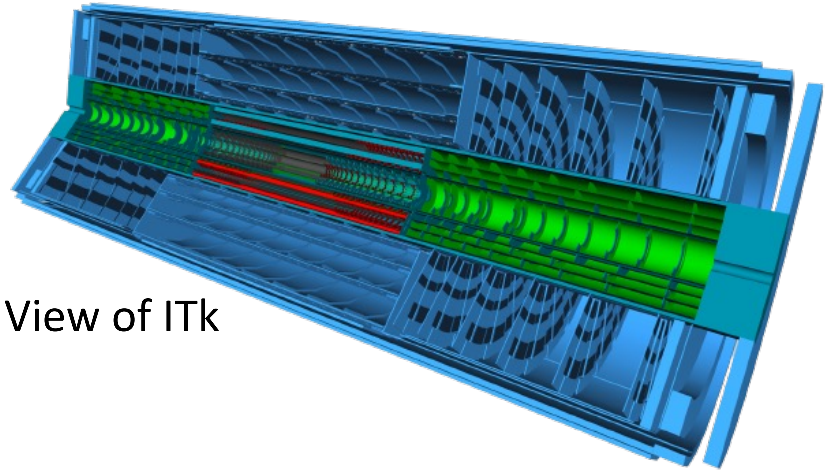
Sharp increase of the event complexity:
Number of simultaneous proton-proton interactions:
$40 \rightarrow 200$

High-luminosity era

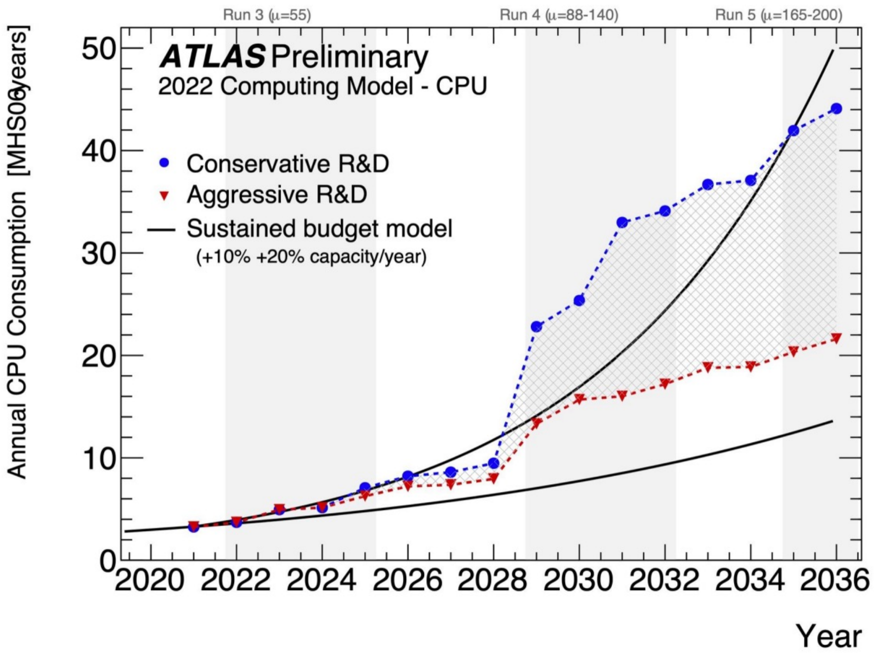
The LHC experiments are upgrading their detector to cope with the harsher environment.

In ATLAS, this includes:

- Installation of a new **I**nner **T**rac**k**er ITk
- Installation of a a High-Granularity Timing Detector
- Upgrade of trigger and data acquisition system
- Replacement of the muon small wheel

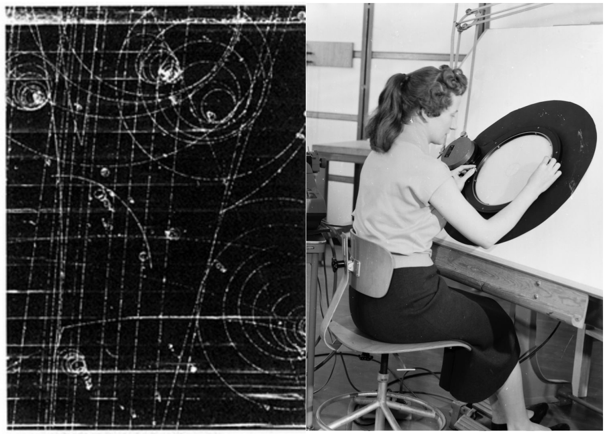


View of ITk



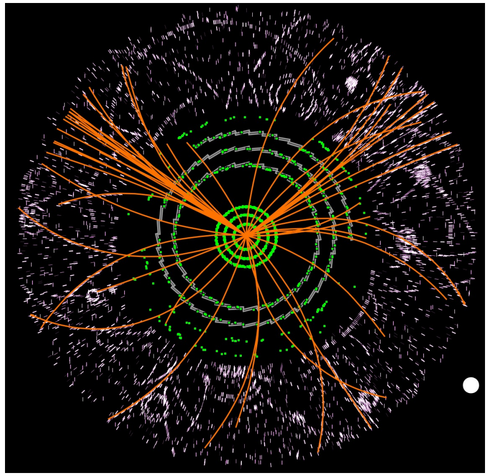
Tracking during HL-LHC: a computing challenge

Track reconstruction is a key step in the event reconstruction, allows to estimate physics parameters of charged particles (p_T, d_0, z_0, φ, …).



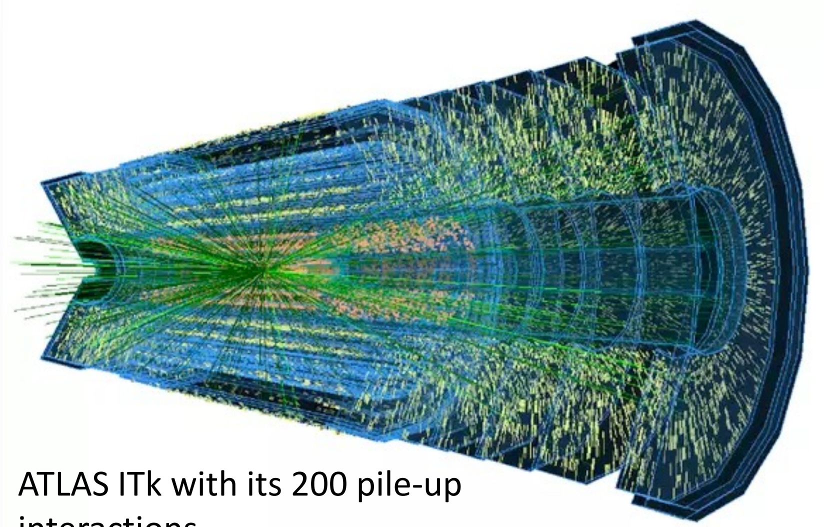
Sixty years before this thesis : track reconstruction with a pen and a ruler.

Picture from: M. Elsing



ATLAS current inner detector
Run 1 ~ 12 pile-up interactions

Tracking done using digital readout and algorithms.



ATLAS ITk with its 200 pile-up interactions.

The algorithms will become too slow to cope with the combinatorics within the CPU budget.

Can we accelerate the algorithm ?

Machine learning applied to tracking

How to deploy a Machine Learning solution for tracking ?

➢ Ask the ML: organize a tracking challenge on the Kaggle platform: TrackML .

The challenge provides an open source tracking dataset, using a detector geometry mimicking the one from the LHC general purpose experiments.

Leaderboard

1		Top Quarks	0.92182	10	2y
2	—	outrunner	0.90302	9	2y
3	—	Sergey Gorbunov	0.89353	6	2y
4	—	demelian	0.87079	35	2y
5	—	Edwin Steiner	0.86395	5	2y
6	—	Komaki	0.83127	22	2y
7	—	Yuval & Trian	0.80414	56	2y
8	—	bestfitting	0.80341	6	2y
9	—	DBSCAN forever	0.80114	23	2y
10	—	Zidmie & KhaVo	0.76320	26	2y

✓ About 650 participants !

✖ No groundbreaking solution: ML is not at the core of the best algorithms

Why is it so difficult to use ML for tracking ?

Deep machine learning applied to tracking

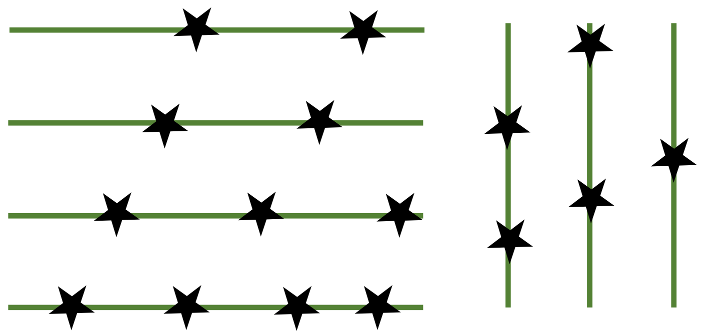
Why is it so difficult to use ML for tracking ?

All the problem lies in the data representation: raw data from tracking detectors are **sparse** data.

Data fed to Neural Networks are frequently represented as an image.

Unsuitable for our tracking detector data.

More natural representation:



Deep machine learning applied to tracking

Why is it so difficult to use ML for tracking ?

All the problem lies in the data representation: raw data from tracking detectors are **sparse** data.

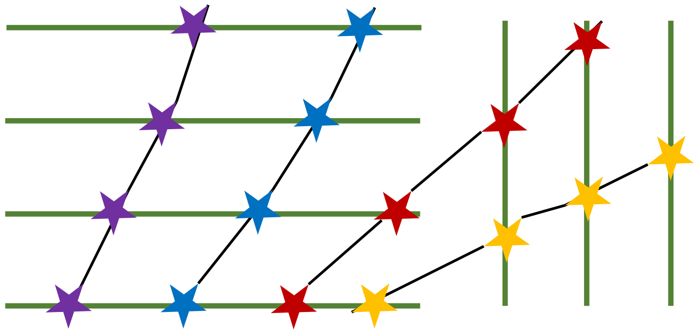
Data fed to Neural Networks are frequently represented as an image.

↓

Unsuitable for our tracking detector data.

More natural representation:

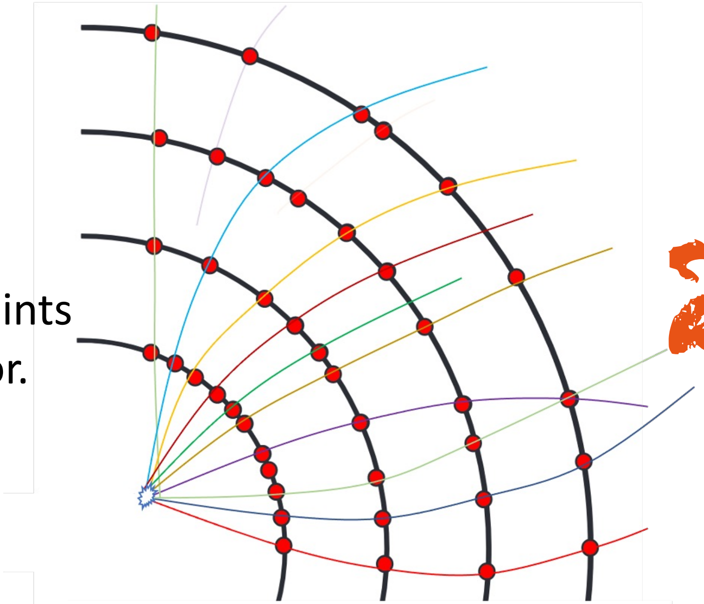
Proof of principle by Exa.TrkX project



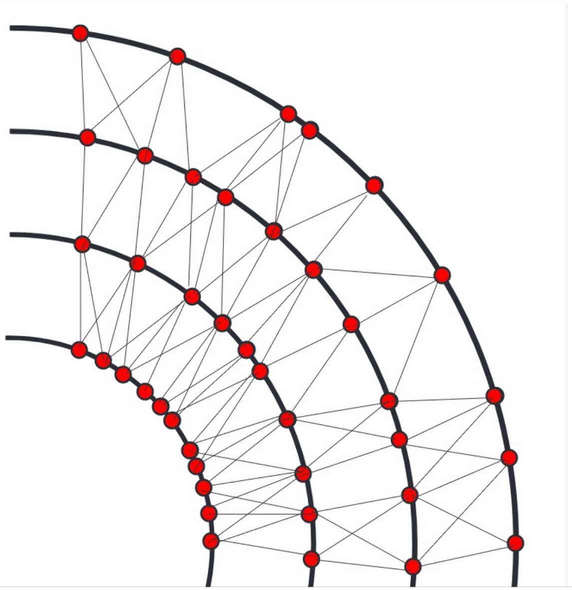
A graph representation !

Tracking with a Graph Neural Network

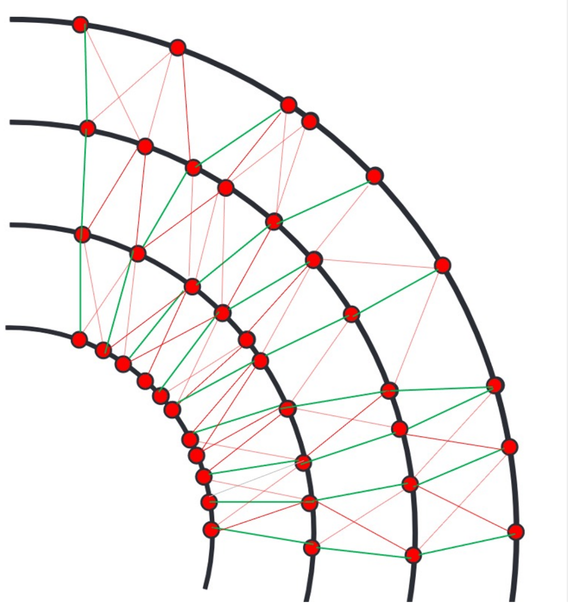
1 Charged particles produce space-points inside the detector.



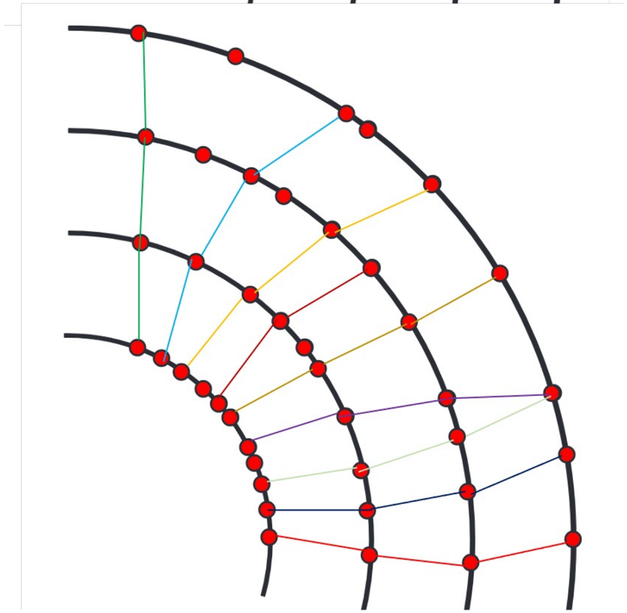
2 Graph creation.



3 Edge classification using a Graph Neural Network.



4 Track building.

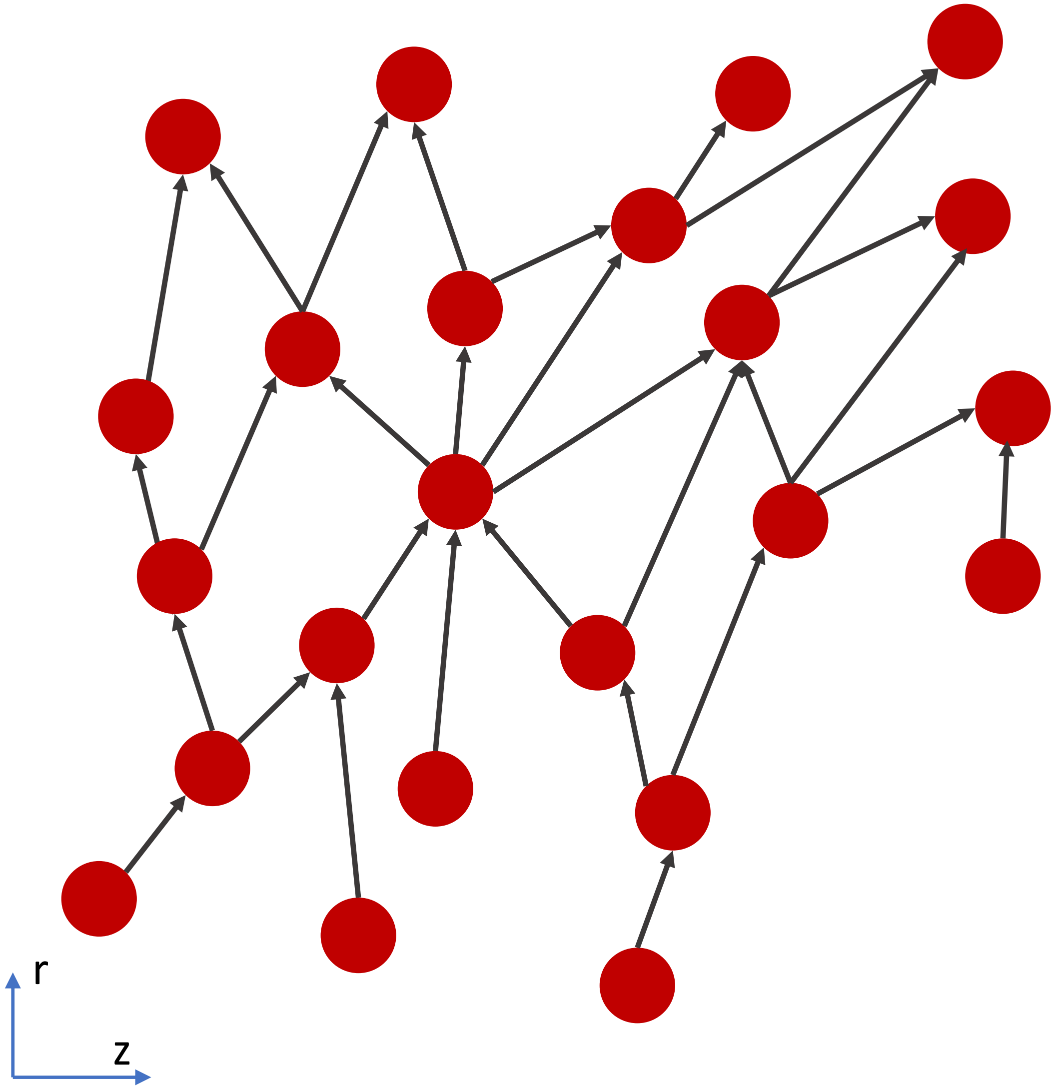


The interaction Network model

Tracking data represented as a graph

$$N_{\text{nodes}}\begin{bmatrix} r & \varphi & z \\ \vdots & \vdots & \vdots \end{bmatrix}$$

$$N_{edges}\begin{bmatrix} \Delta\eta & \Delta\varphi & \Delta r & \Delta z \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$



The interaction Network model

Tracking data represented as a graph

$$N_{nodes}\begin{bmatrix} r & \varphi & z \\ \vdots & \vdots & \vdots \end{bmatrix}$$

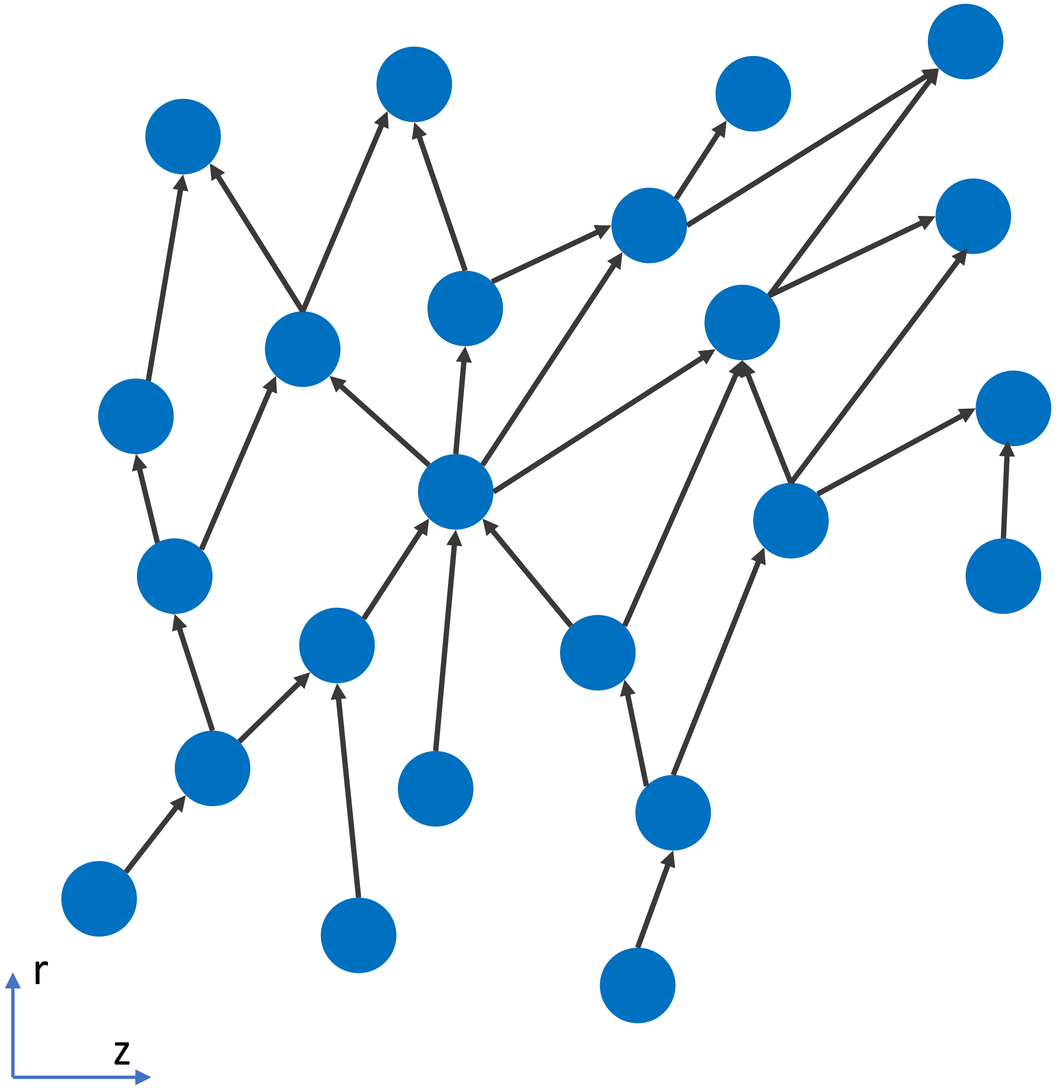
MLP

Node Encoder

$$N_{edges}\begin{bmatrix} \Delta\eta & \Delta\varphi & \Delta r & \Delta z \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

MLP

Edge Encoder



The interaction Network model

Tracking data represented as a graph

$$\mathrm{N_{nodes}}\begin{bmatrix} r & \varphi & z \\ \vdots & \vdots & \vdots \end{bmatrix}$$

$$N_{edges}\begin{bmatrix} \Delta\eta & \Delta\varphi & \Delta r & \Delta z \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

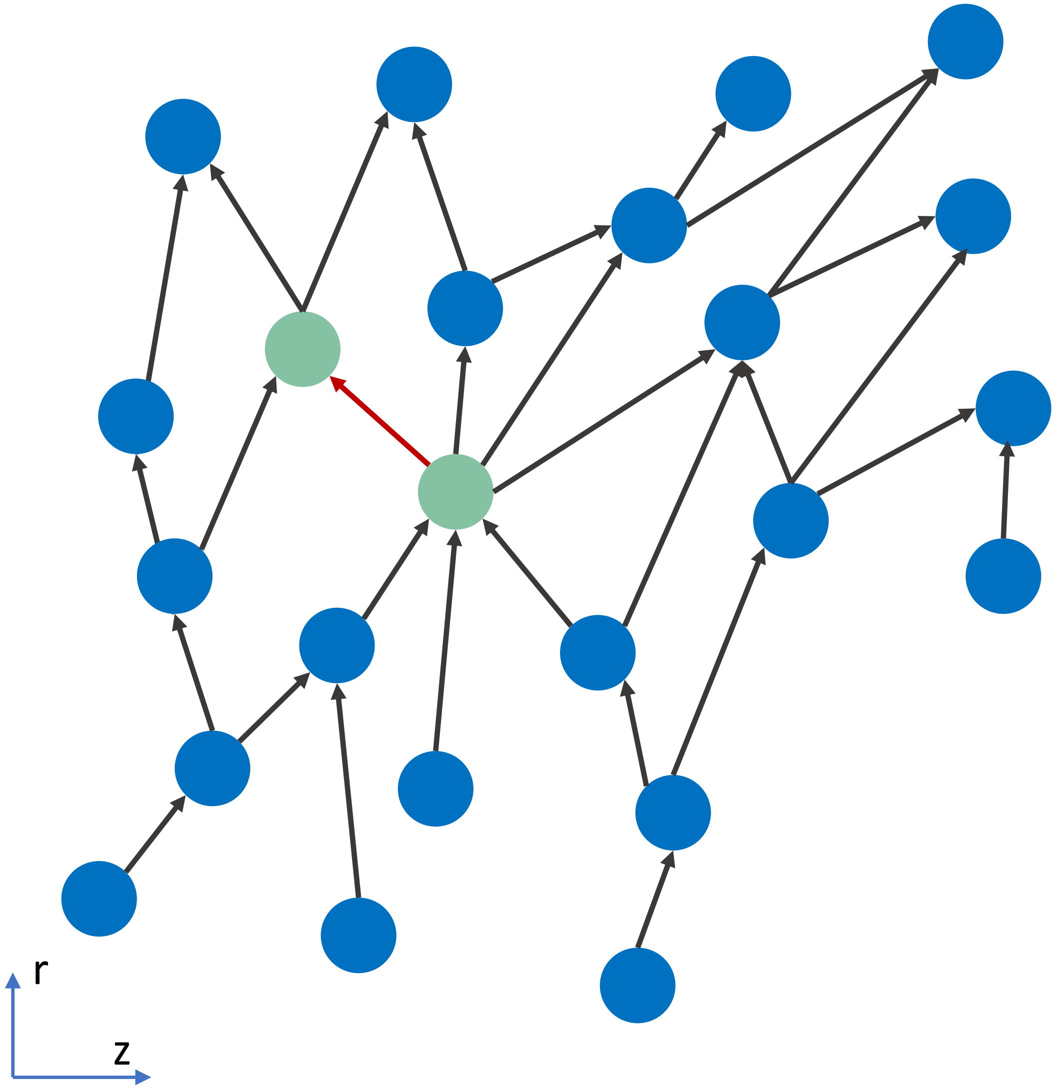
MLP

Node Encoder

MLP

Edge Encoder

Interaction Network



Edge update: $\mathbf{e}^{k}_{(u,v)} = \mathrm{MLP}^{k}_{\psi}\left([\mathbf{h}^{k-1}_{u}, \mathbf{h}^{k-1}_{v}, \mathbf{e}^{k\ -1}_{(u,v)}]\right)$

The interaction Network model

Tracking data represented as a graph

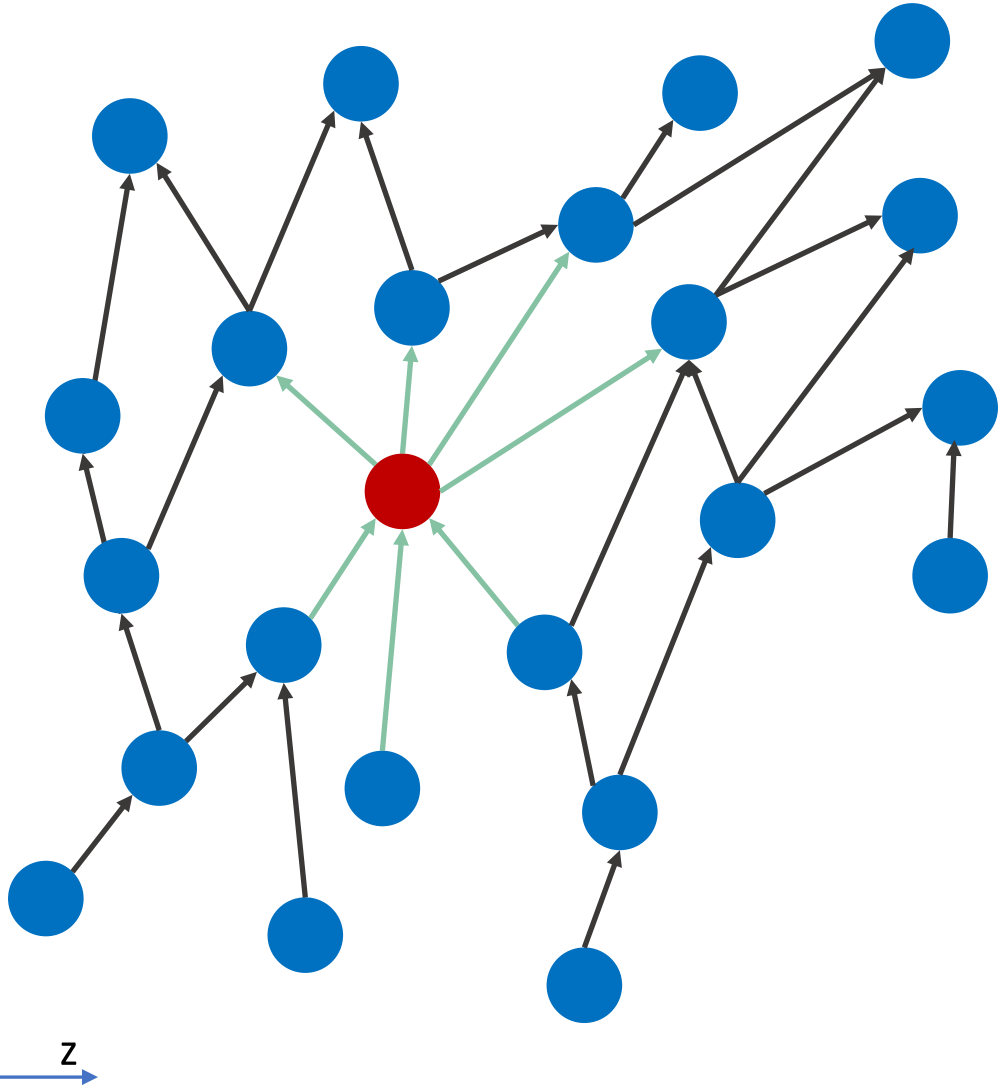
$$N_{\text{nodes}} \begin{array}{c} \\ \left[\begin{array}{ccc} r & \varphi & z \\ \vdots & \vdots & \vdots \end{array}\right] \end{array} \Rightarrow \text{MLP: Node Encoder}$$

$$N_{edges} \left[\begin{array}{cccc} \Delta\eta & \Delta\varphi & \Delta r & \Delta z \\ \vdots & \vdots & \vdots & \vdots \end{array}\right] \Rightarrow \text{MLP: Edge Encoder}$$

MLP
Node Encoder

MLP
Edge Encoder

⇒ Interaction Network



Node update: $\mathbf{h}_u^k = \mathrm{MLP}_\phi^k\left(\mathbf{h}_u^{k-1}, \sum_{v\in N_{\text{in}}(u)} \mathbf{e}_{(v,u)}^k, \sum_{v\in N_{\text{out}}(u)} \mathbf{e}_{(u,v)}^k\right)$

The interaction Network model

Tracking data represented as a graph

$$N_{\text{nodes}} \begin{array}{c} \begin{matrix} r & \varphi & z \end{matrix} \\ \begin{bmatrix} \vdots & \vdots & \vdots \end{bmatrix} \end{array}$$

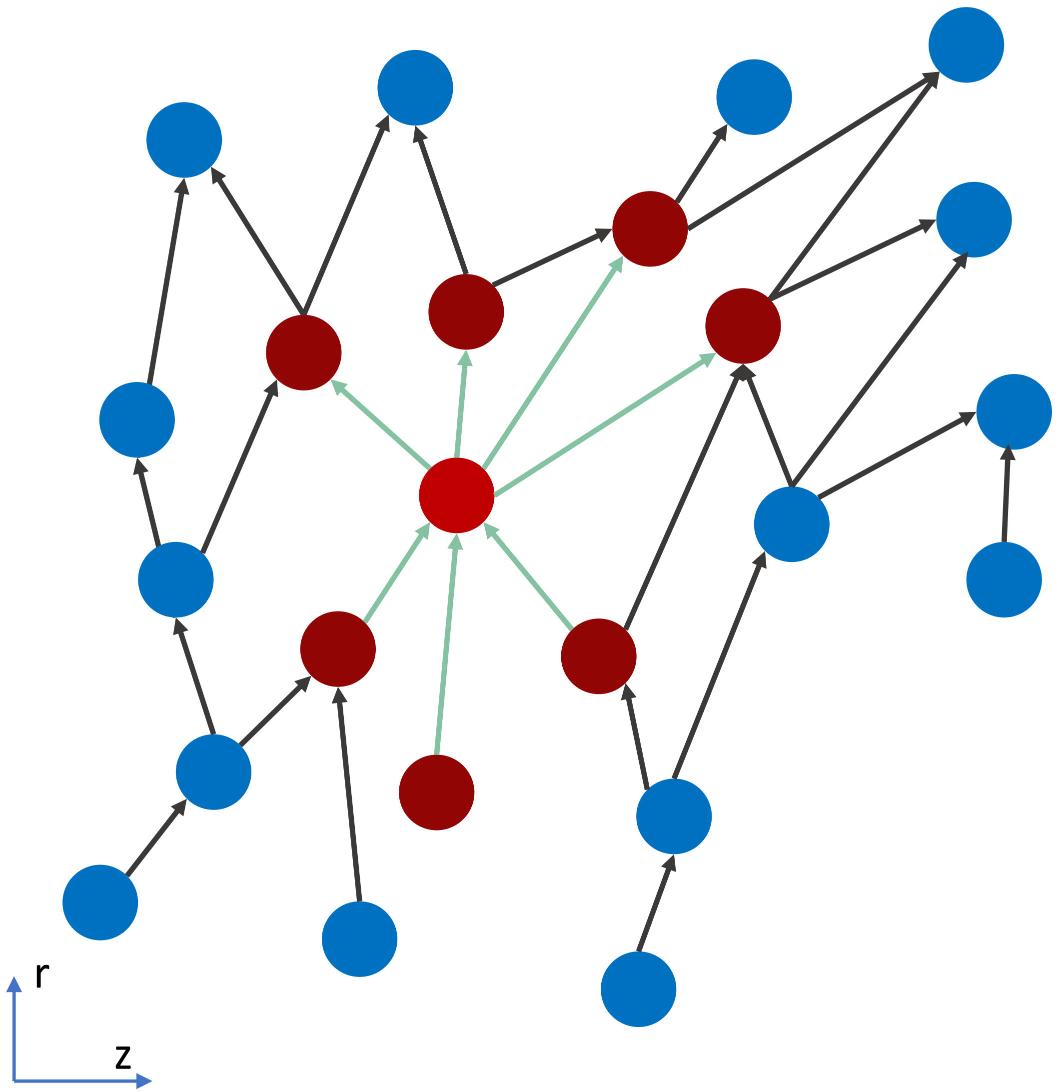
$$N_{edges} \begin{bmatrix} \Delta\eta & \Delta\varphi & \Delta r & \Delta z \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

⇨ MLP: Node Encoder

⇨ MLP: Edge Encoder

⇨ Interaction Network

Message-passing mechanism



The interaction Network model

Tracking data represented as a graph

$$N_{\mathrm{nodes}} \begin{bmatrix} r & \varphi & z \\ \vdots & \vdots & \vdots \end{bmatrix}$$

$$N_{edges} \begin{bmatrix} \Delta\eta & \Delta\varphi & \Delta r & \Delta z \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

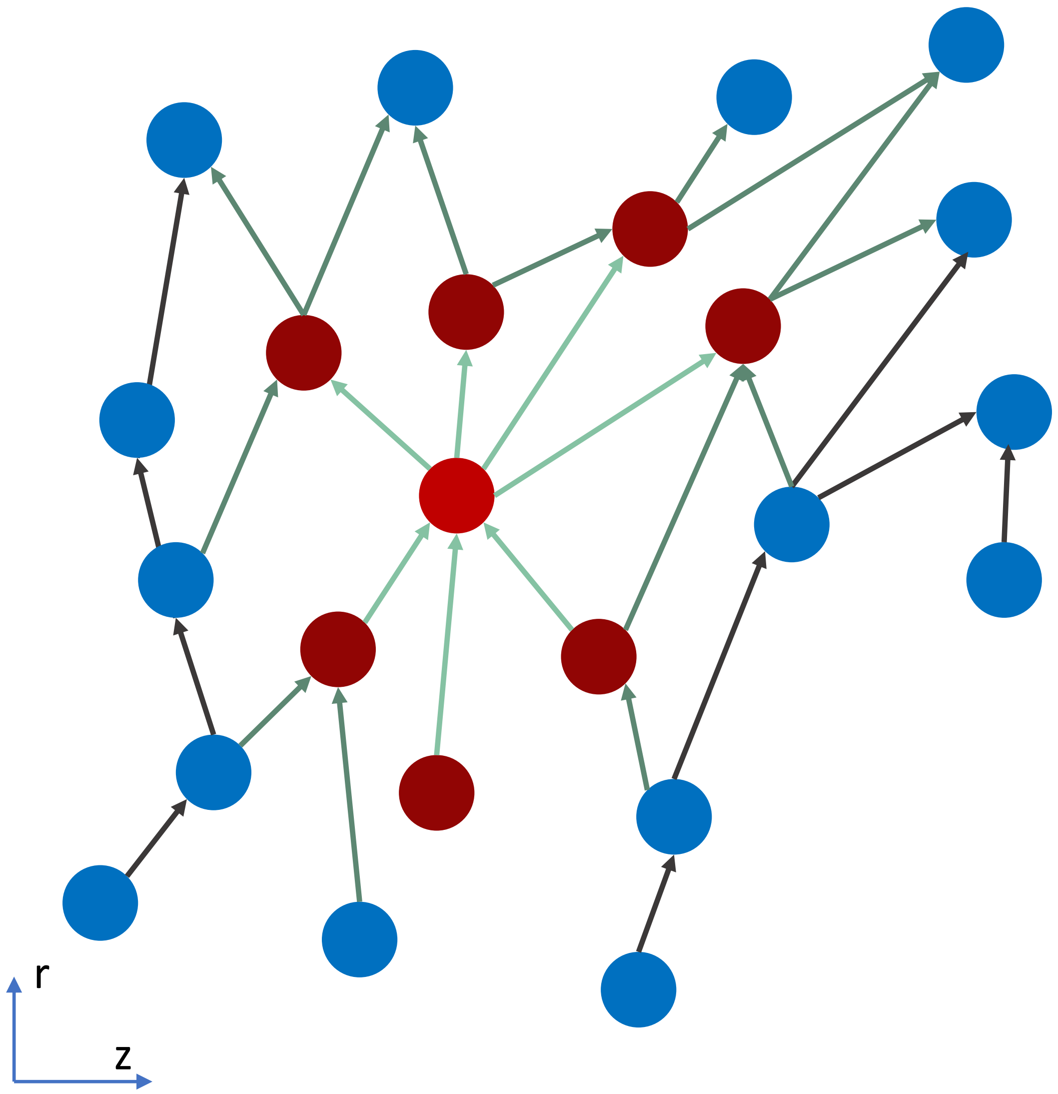
MLP

Node Encoder

MLP

Edge Encoder

Interaction Network



Message-passing mechanism

The interaction Network model

Tracking data represented as a graph

$$N_{\text{nodes}} \begin{bmatrix} r & \varphi & z \\ \vdots & \vdots & \vdots \end{bmatrix}$$

$$N_{edges} \begin{bmatrix} \Delta\eta & \Delta\varphi & \Delta r & \Delta z \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

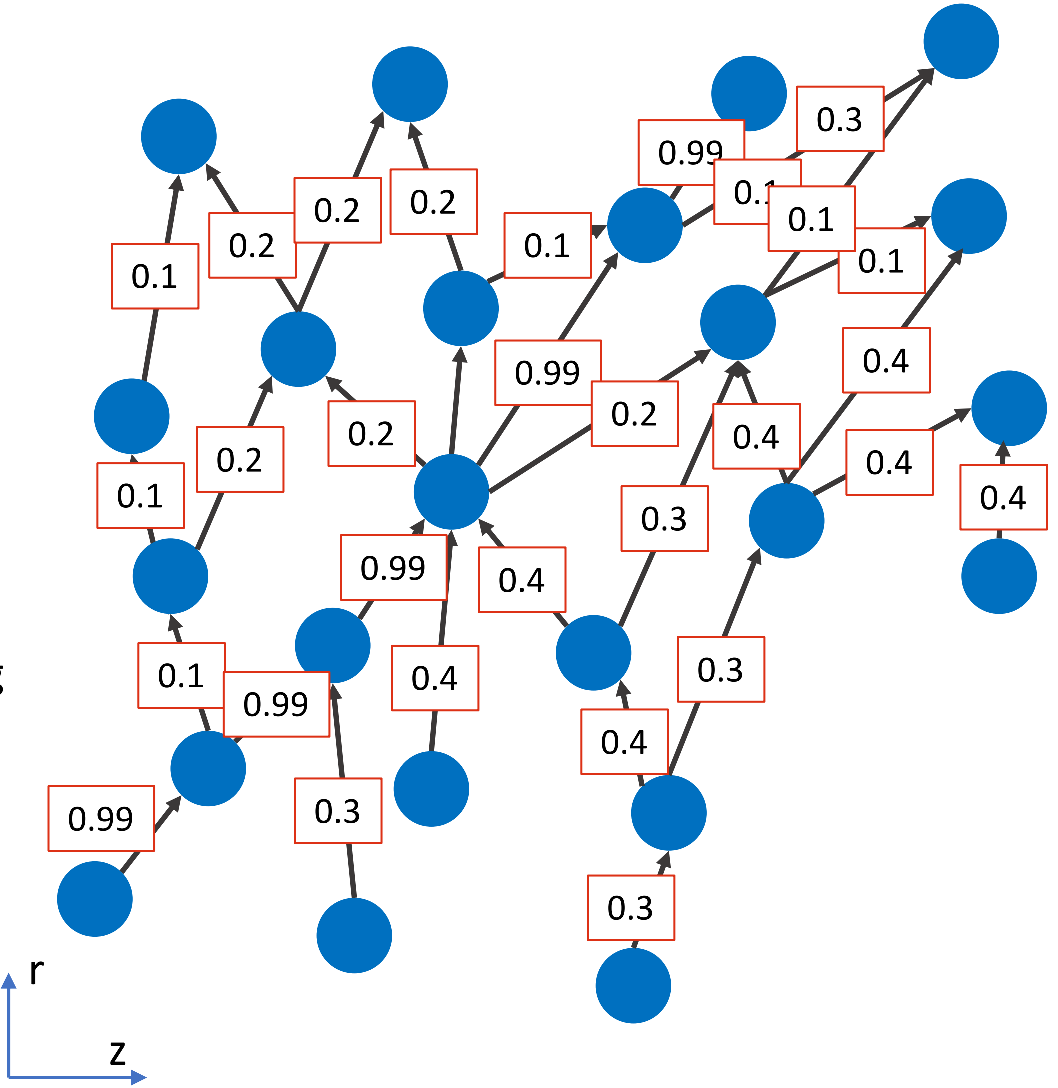
MLP
Node Encoder

MLP
Edge Encoder

Interaction Network

Message-passing

Classification score



Simulated sample

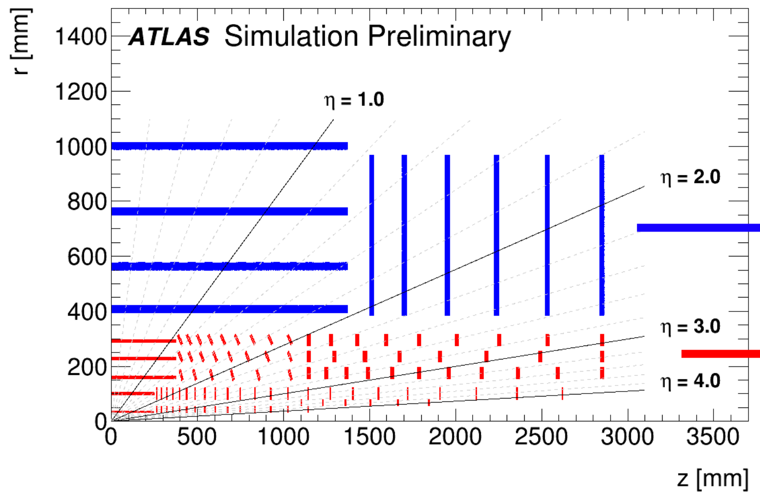
ATLAS simulated sample: $t\bar{t}$ with $\langle\mu\rangle = 200$ at $\sqrt{s} = 14\ TeV$

- About 10k charged particles per event
- About 300k space-points per event

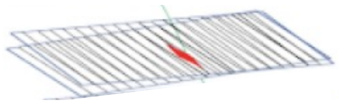
Define target particles

- $p_T > 1$ GeV
- No secondaries
- No electrons
- At least 3 space-points in the detector

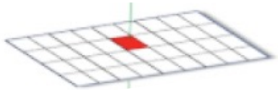
Select $\mathbf{t\bar{t}}$ and pile-up interactions



Strip subdetector: 1 space-point = 2 clusters



Pixel subdetector: 1 space-point = 1 cluster

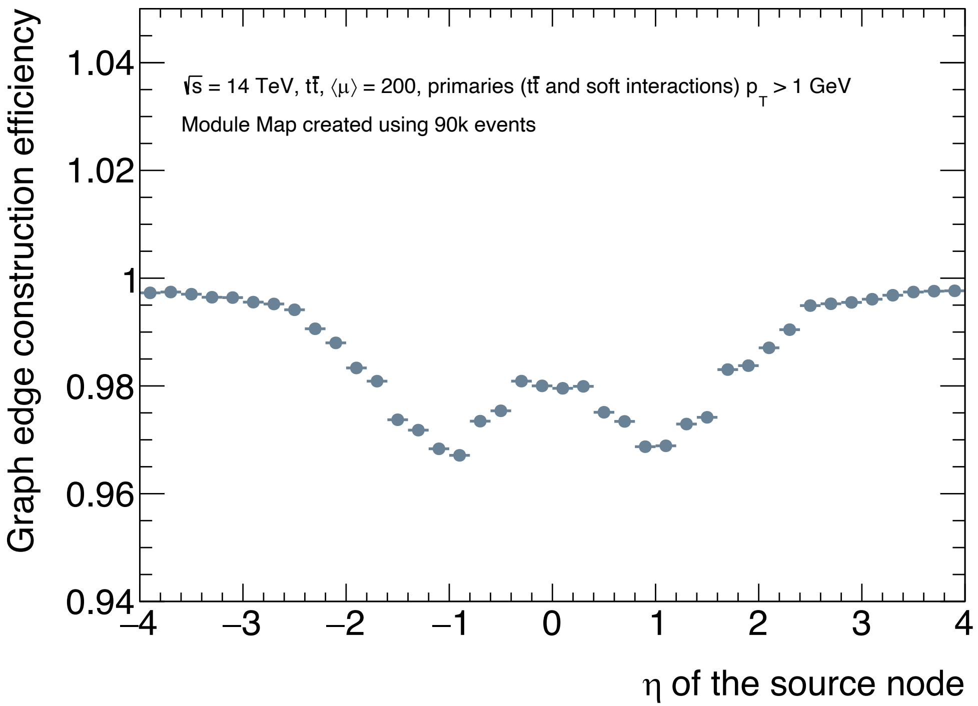


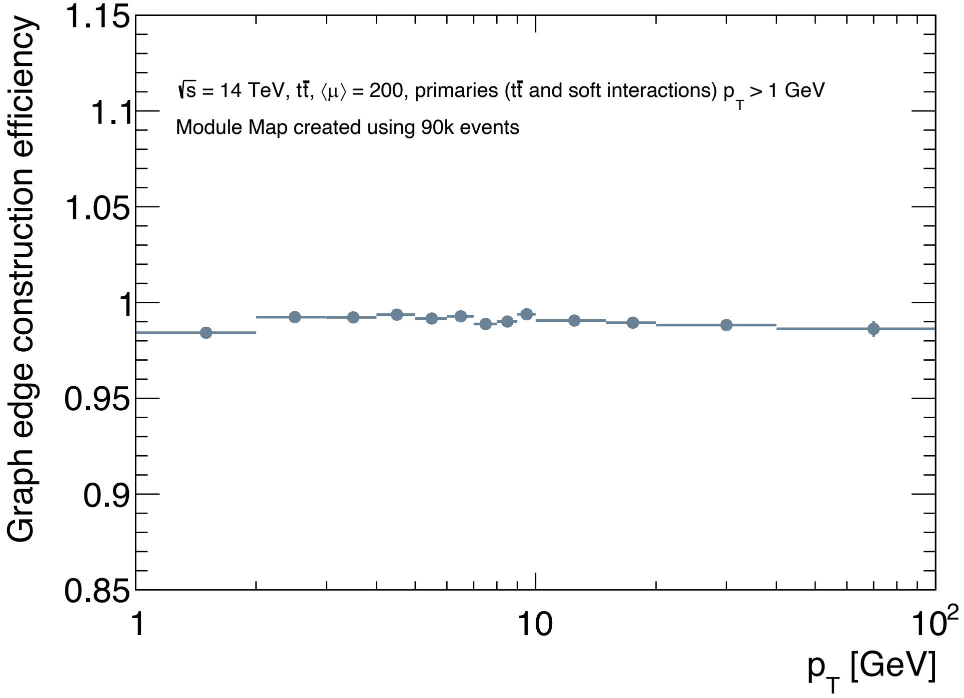
Graph creation: the *module map*

The path of a target particle is followed through ITk to record all possible **connections** between triplets of silicon **modules**:

- Built using 90 000 $t\bar{t}$ events with $\langle\mu\rangle = 200$
- It comprises **1 242 665** connections
- Direction "***inside-out***" is given to edges.

- On average, the graphs have O(270k) nodes and O(1.3 million) edges.





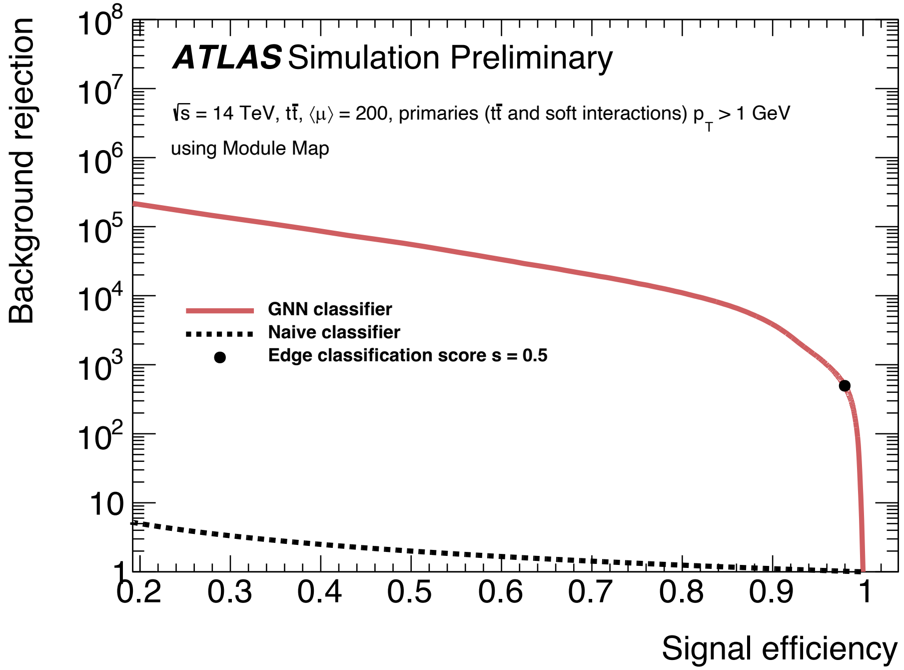
Training the GNN

A technical challenge

Simple GNN architecture configuration

Require 200 GB of GPU memory to train the model

Use of memory management technique.



Configuration of the GNN architecture

- 2 layers in each MLP
- 128−dimensional space parameters
- 8 message-passing iterations

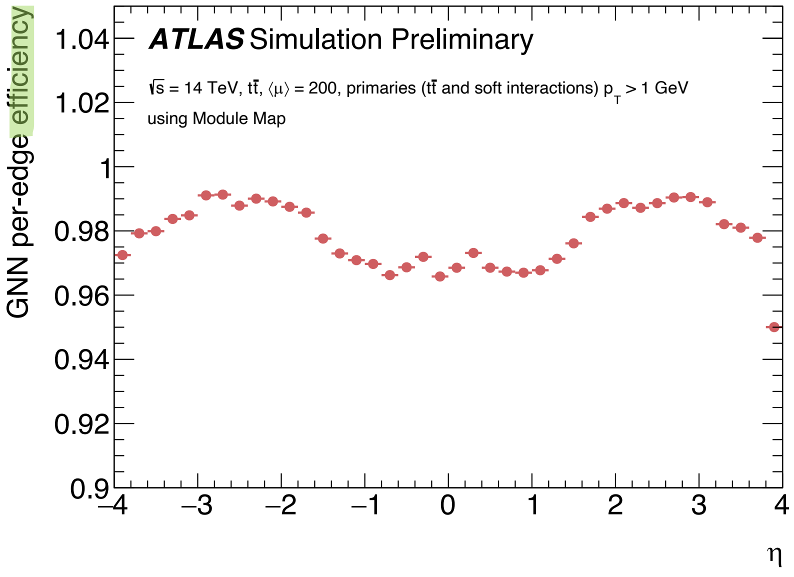
$$L_{BCE}(W,b) = -\frac{1}{N}\sum_{i=1}^{N}\Big[y_i \log(p_{w,b}(x_i)) + (1-y_i)\log(1-p_{w,b}(x_i))\Big] \times w_i$$

$$\mathrm{w_i} = \begin{cases} 1 \text{ for true edges} \\ 0.1 \text{ for fake edges} \\ 0 \text{ for non−target edges} \end{cases}$$

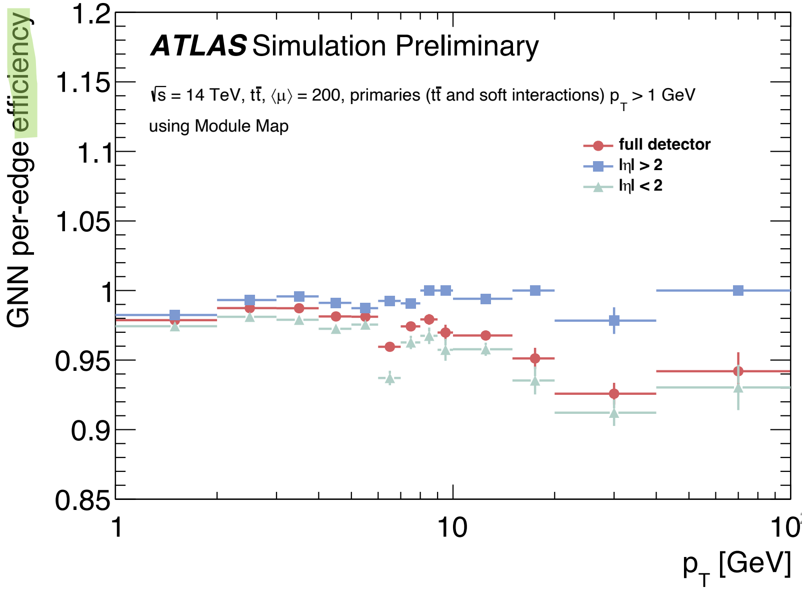
Edge classification performance

Cut at s = 0.5 on the edge classification score

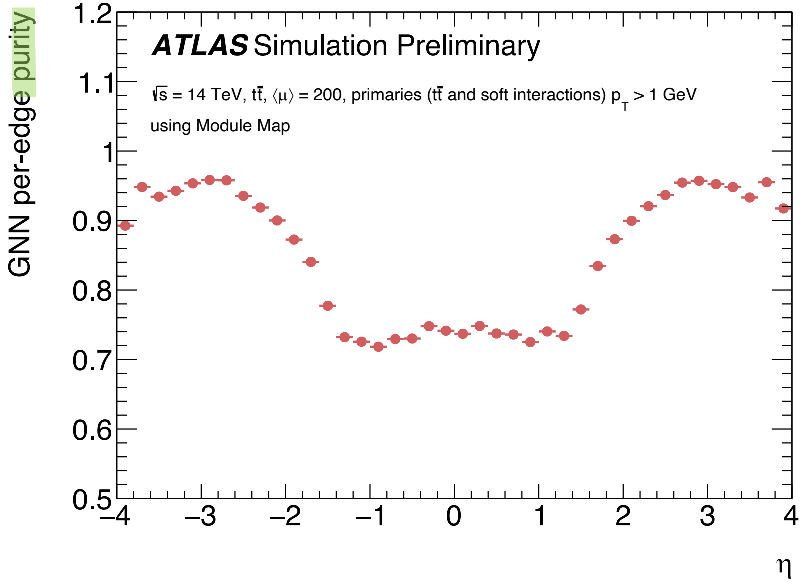
Efficiency vs. η



Efficiency vs. p_T



Purity vs. η

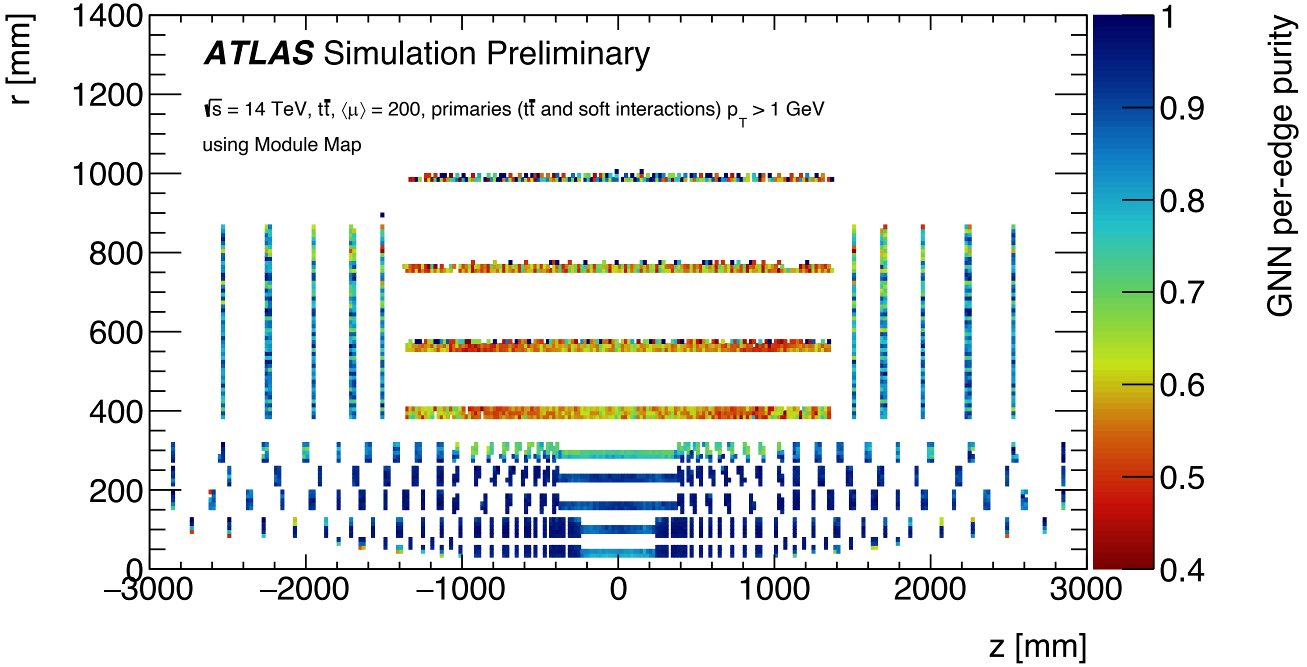


Efficiency and purity degradation in the central region.

What is the source of the inefficiency ?

Investigation of the GNN edge-performance

Cut at s = 0.5 on the edge classification score



Building track candidates

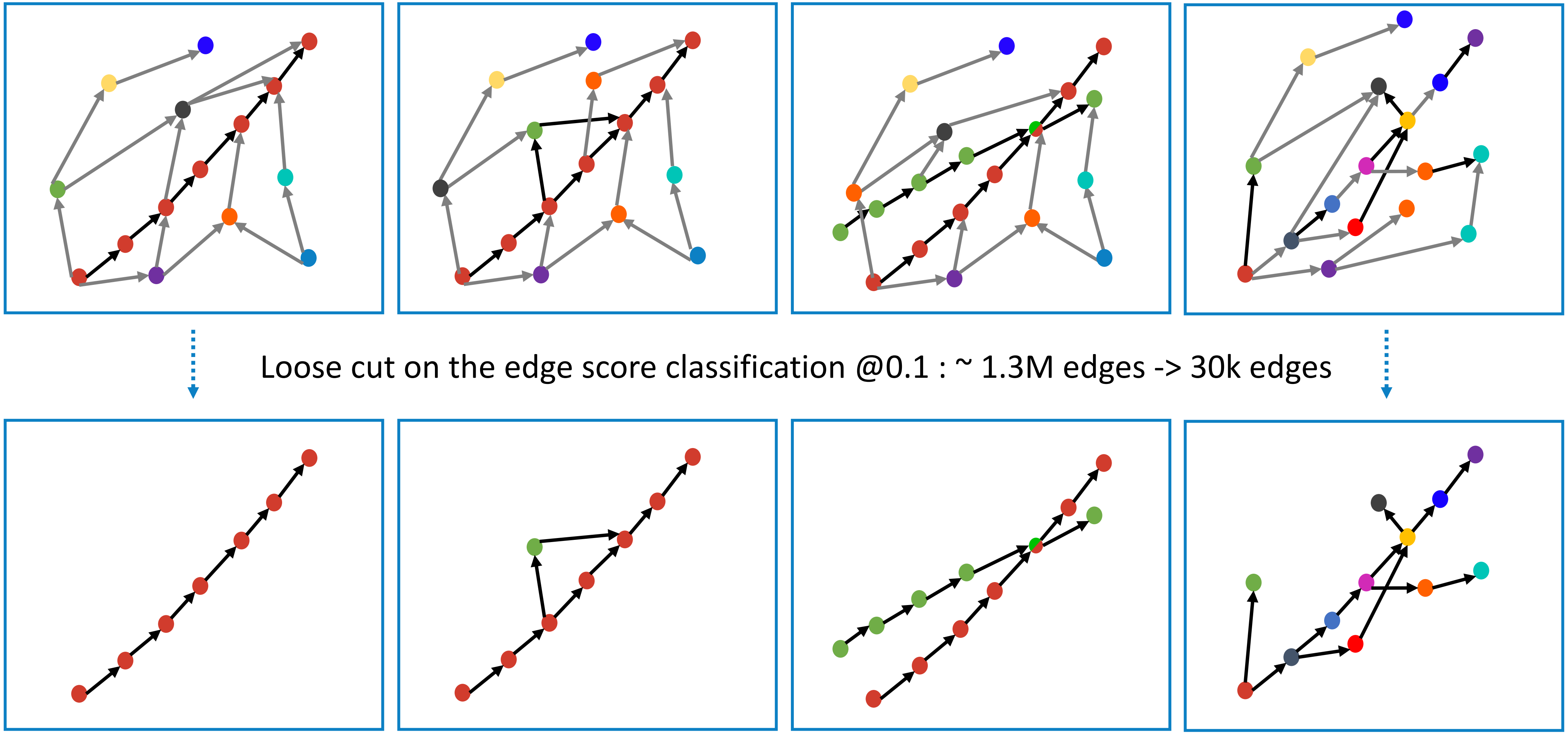
Legend

Edge below threshold

Edge above threshold

Nodes same color = Nodes same particles

Zoom on a graph



Loose cut on the edge score classification @0.1 : ~ 1.3M edges -> 30k edges

Connected components

Building track candidates

Slide 2
SH
Tracking
Back-up-content

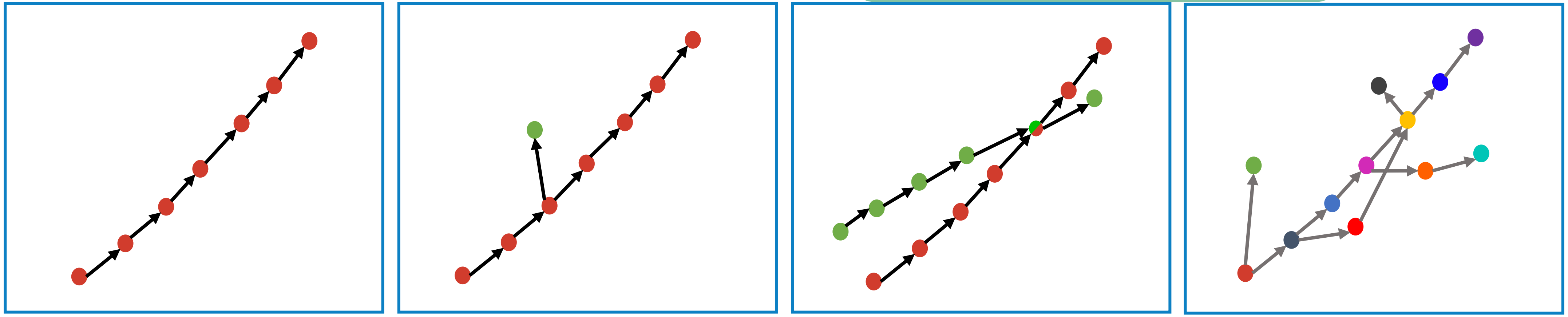
Legend

Edge below threshold

Edge above threshold

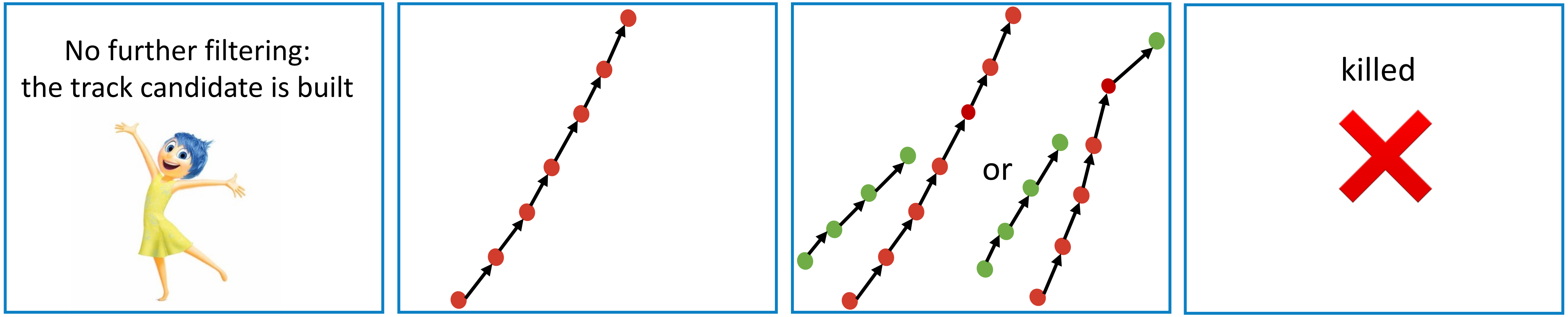
Nodes same color = Nodes same particles

Connected components



Iterative algorithm with tighter cuts on edge score classification

Walk-through algorithm (from TrackML)



GNN track reconstruction efficiency

Matching criteria:

Particle

Standard matching

$P_{match} > 0.5$

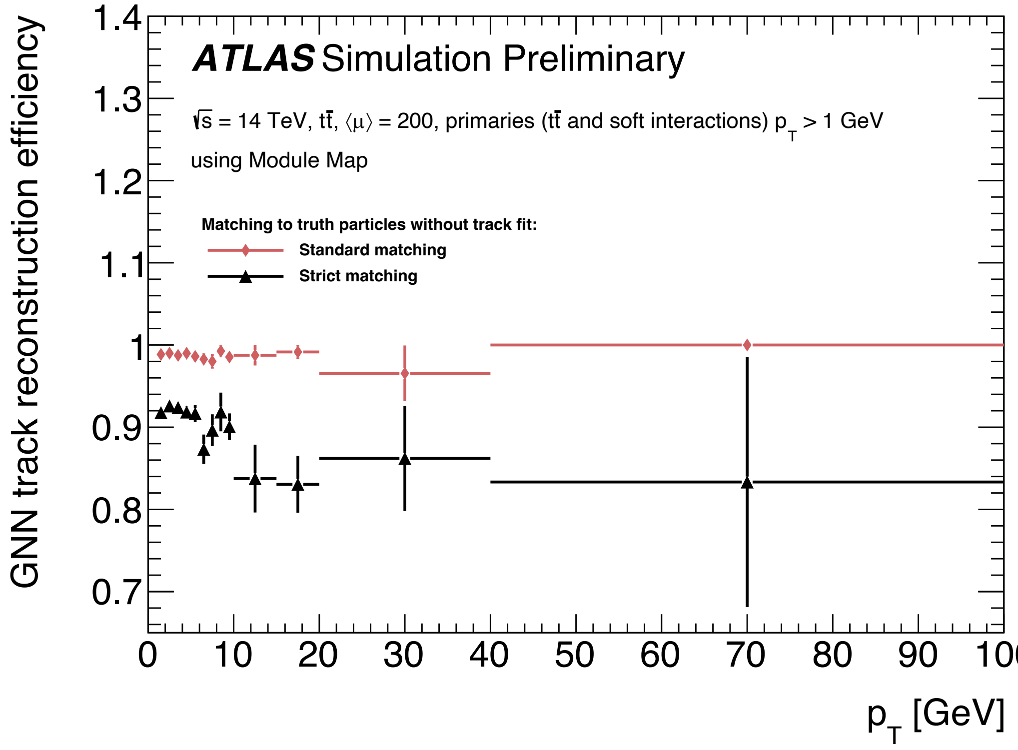
Strict Matching

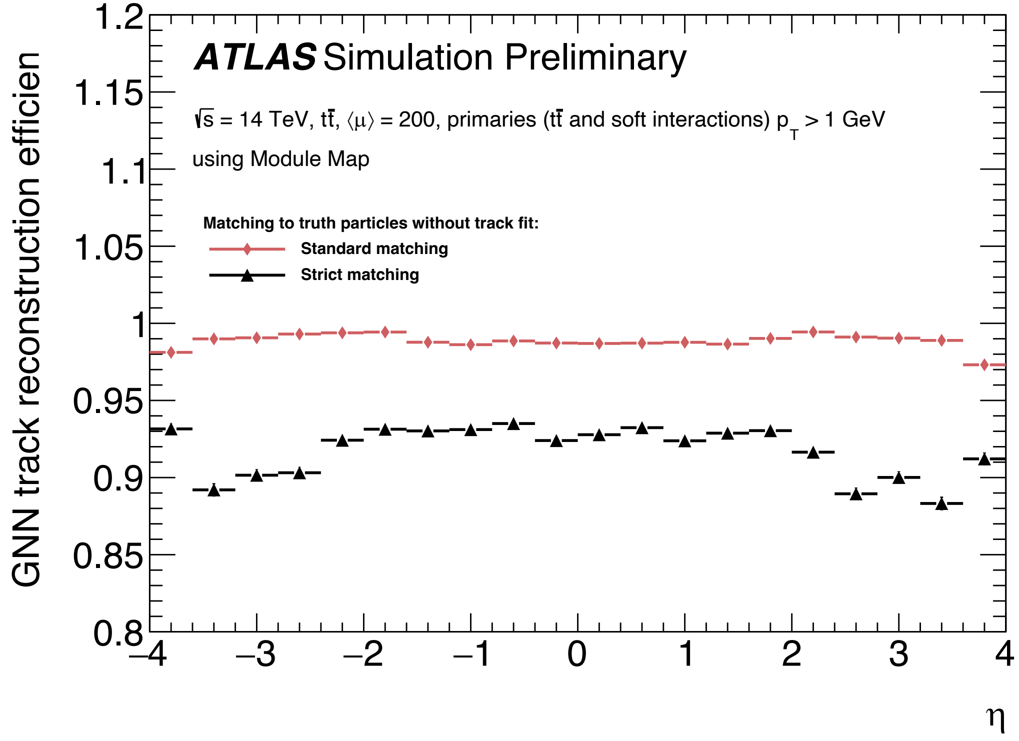
$P_{match} = 1$

Track purity = 1

Track candidate not matched to any particle = fake track

- ➢ found to be $O(10^{-3})$





A Ph.D. in 3 years

End Run 2 @LHC
2018

GNN tracking
Proof of principle
Exa.TrkX

TrackML

CORONAVIRUS

Start Run 3 @LHC
2022

Today

Run 2

2019
Creation of L2IT

Oct 2020:
Beginning of my Ph.D.

Oct 2023:
End of my Ph.D.

A Ph.D. in 3 years

End Run 2 @LHC

2018

GNN tracking
Proof of principle

TrkX

Start Run 3 @LHC

2022

Today

Run 2

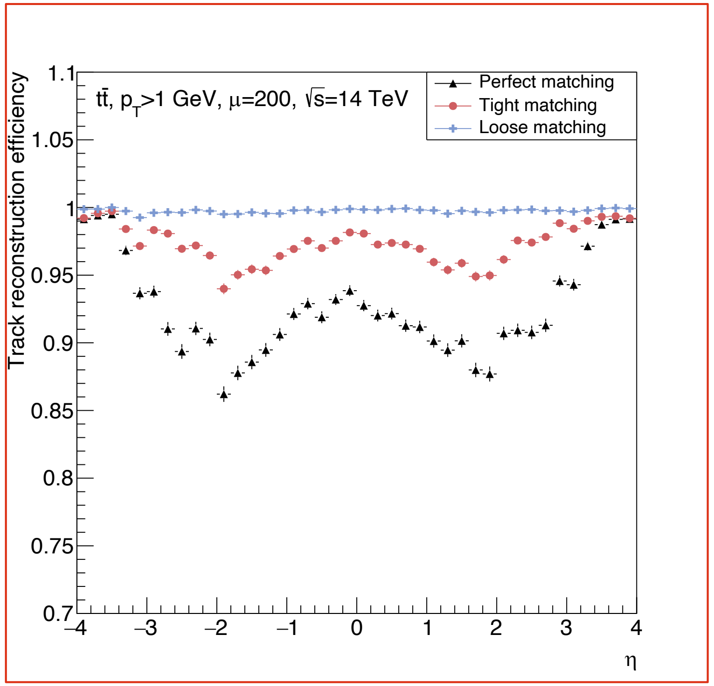
2019

Creation of L2IT

May 2021

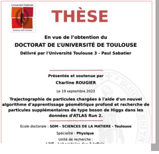
CHEP 2021

Plenary @CHEP 2021
First GNN tracking
performance



Oct 2020:
Beginning of my Ph.D.

Oct 2023:
End of my Ph.D.



A Ph.D. in 3 years

End Run 2 @LHC

2018

GNN tracking
Proof of principle

TrkX

TrackML

CORONAVIRUS

Start Run 3 @LHC

2022

Today

Run 2

QT: getting ATLAS authorship

2019

Creation of L2IT

May 2021

vCHEP 2021

Plenary @CHEP 2021
First GNN tracking performance

Oct 2020:
Beginning of my Ph.D.

Oct 2023:
End of my Ph.D.

A Ph.D. in 3 years

End Run 2 @LHC
2018

GNN tracking
Proof of principle

TrackML

Start Run 3 @LHC
2022

Today

Run 2

QT: getting ATLAS authorship

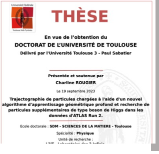
2019
Creation of L2IT

May 2021
vCHEP 2021
Plenary @CHEP 2021
First GNN tracking performance

End 2021:
Join the SH analysis

Oct 2020:
Beginning of my Ph.D.

Oct 2023:
End of my Ph.D.



A Ph.D. in 3 years

[Slide 2](#)
[SH](#)
[Tracking](#)
[Back-up-content](#)

End Run 2 @LHC

2018

GNN tracking
[Proof of principle](#)

TrkX

TrackML

Start Run 3 @LHC

2022

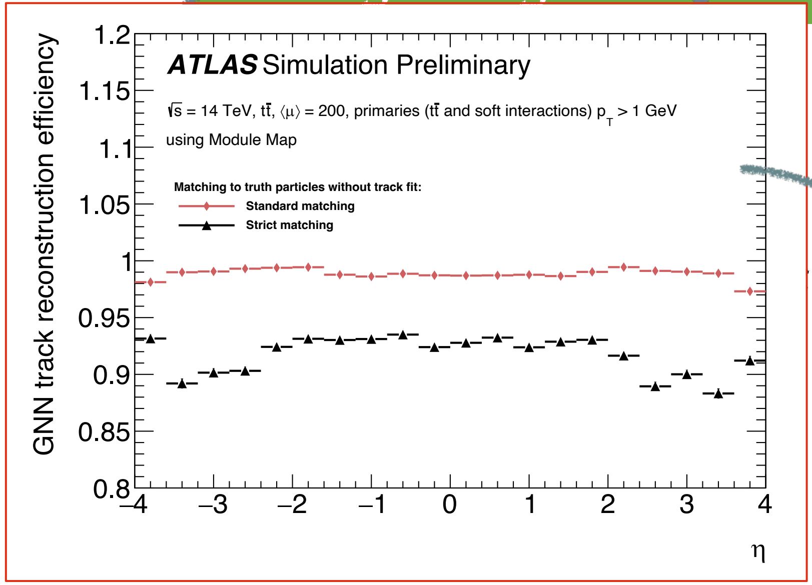
Today

Run 2

QT: getting ATLAS authorship

2019

Creation of L2IT



May 2022

Plenary @CTD 2022
First ITk GNN tracking performance

Oct 2020:
Beginning of my Ph.D.

Oct 2023:
End of my Ph.D.

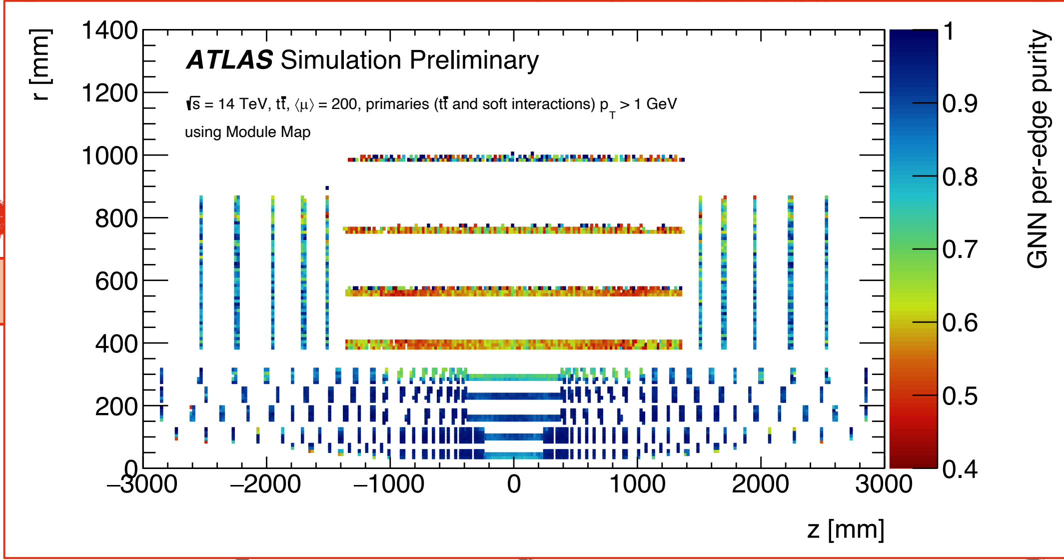
A Ph.D. in 3 years

End Run 2 @LHC

2018

GNN tracking

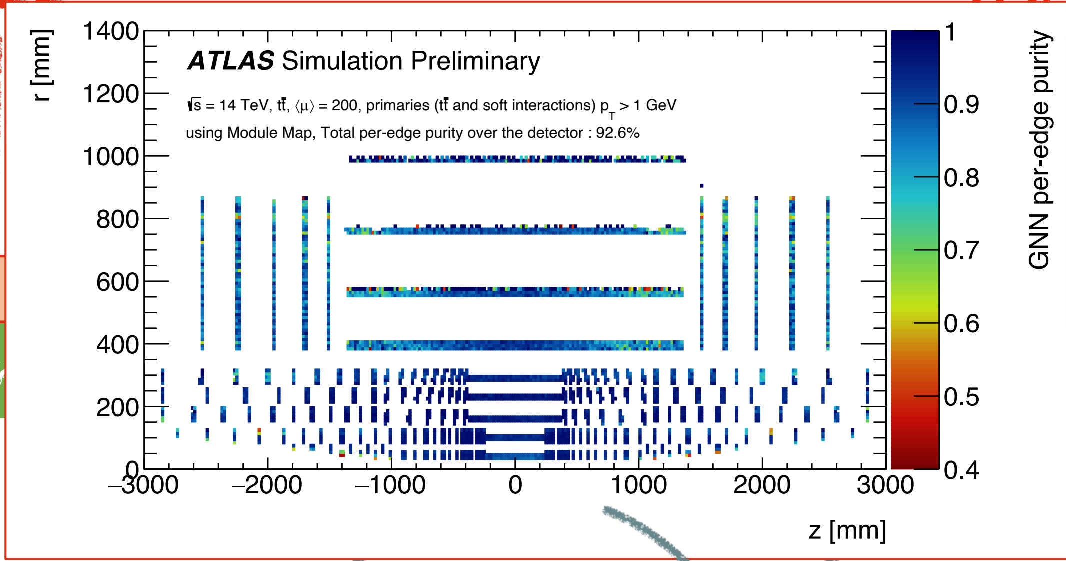
TrkX



Start Run 3 @LHC

2022

Today



Run 2

2019

Creation of L2IT

May 2021

CHEP 2021

Plenary @CHEP 2021
First GNN tracking performance

End 2021:
Join the SH analysis

May 2022

Plenary @CTD 2022
First ITk GNN tracking performance

May 2023

CHEP 2023

CHEP 2023

Oct 2020:
Beginning of my Ph.D.

Oct 2023:
End of my Ph.D.

A Ph.D. in 3 years

End Run 2 @LHC

2018

GNN tracking
[Proof of principle](#)

TrkX

TrackML

CORONAVIRUS

Start Run 3 @LHC

2022

Today

Run 2

QT: getting ATLAS authorship

2019
Creation of L2IT

Oct 2020:
Beginning of my Ph.D.

May 2021
CHEP 2021
Plenary @CHEP 2021
First GNN tracking performance

End 2021:
Join the SH analysis

May 2022
Connecting The Dots 2022
7th International CTD Workshop
Princeton University, Princeton, USA
May 31 - June 2, 2022
Plenary @CTD 2022
First ITk GNN tracking performance

May 2023
CHEP 2023
CHEP 2023

Sept 2023
HDBS approval
SH analysis

Oct 2023:
End of my Ph.D.

What next ?

End Run 3 @LHC

~2025

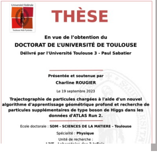
~2029

ATLAS & CMS Upgrades

Start Run 4 @LHC

First HiLumi collisions recorded

Oct 2023:
End of my Ph.D.



What next ?

End Run 3 @LHC

~2025

~2029

ATLAS & CMS Upgrades

Start Run 4 @LHC

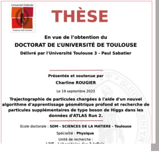
First HiLumi collisions recorded

Oct. 2023

CTD 2023
In Toulouse

Present first comparison between CKF and GNN

Oct 2023:
End of my Ph.D.



What next ?

End Run 3 @LHC

~2025

~2029

ATLAS & CMS Upgrades

Start Run 4 @LHC

First HiLumi collisions recorded

Oct. 2023

CTD 2023
In Toulouse

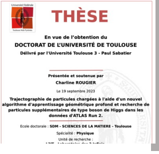
End 2023

SH paper

Search for a new resonance decaying to a scalar and a Higgs boson in the final state with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s}$=13 TeV

The ATLAS Collaboration

Oct 2023:
End of my Ph.D.



What next ?

End Run 3 @LHC

2024

~2025

~2029

ATLAS & CMS Upgrades

Start Run 4 @LHC

First HiLumi collisions recorded

Oct. 2023

CTD 2023
In Toulouse

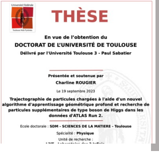
End 2023

SH paper

Completion of GNN tracking demonstrator for ATLAS ITk

Target papers on physics and technical performance

Oct 2023:
End of my Ph.D.



What next ?

End Run 3 @LHC

2024

~2025

~2029

ATLAS & CMS Upgrades

Start Run 4 @LHC

First HiLumi collisions recorded

Oct. 2023

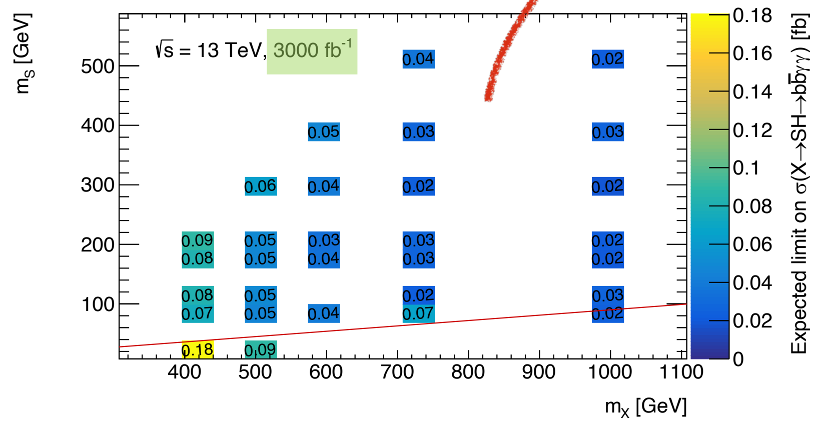
CTD 2023
In Toulouse

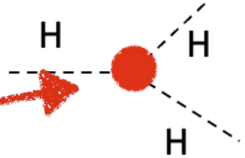
End 2023

SH paper

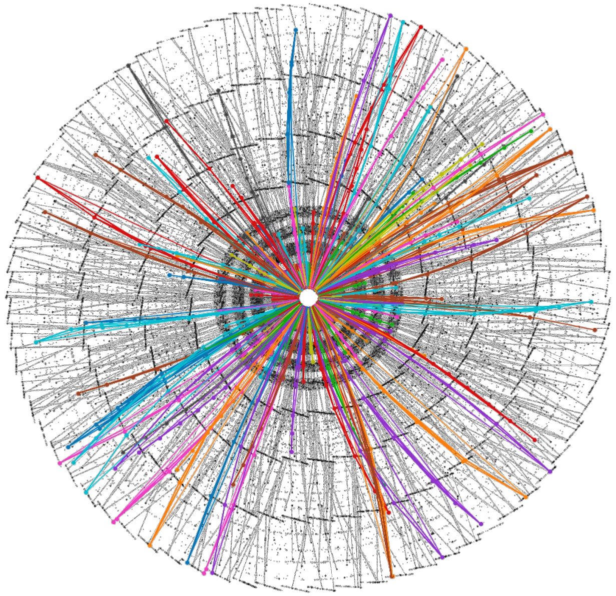
Completion of GNN tracking demonstrator for ATLAS ITk

Target papers on physics and technical performance

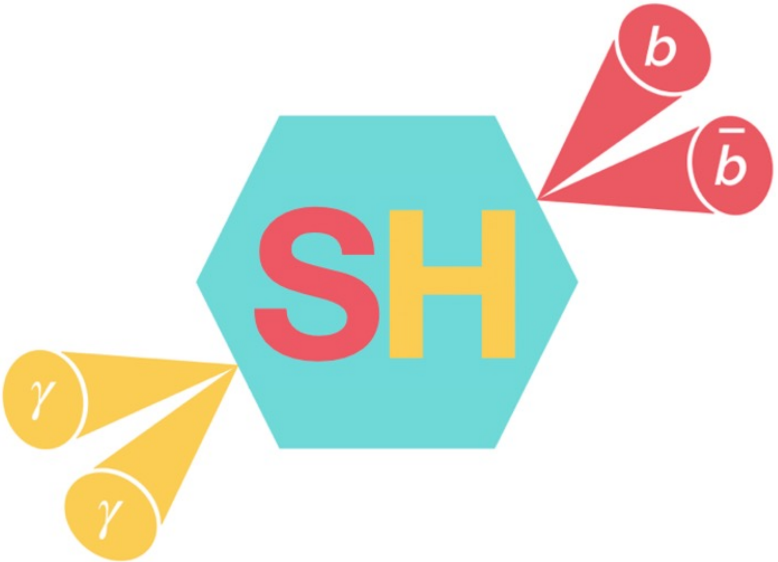




Oct 2023:
End of my Ph.D.

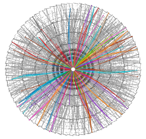


Thank you for your attention



Content back-up

Analysis



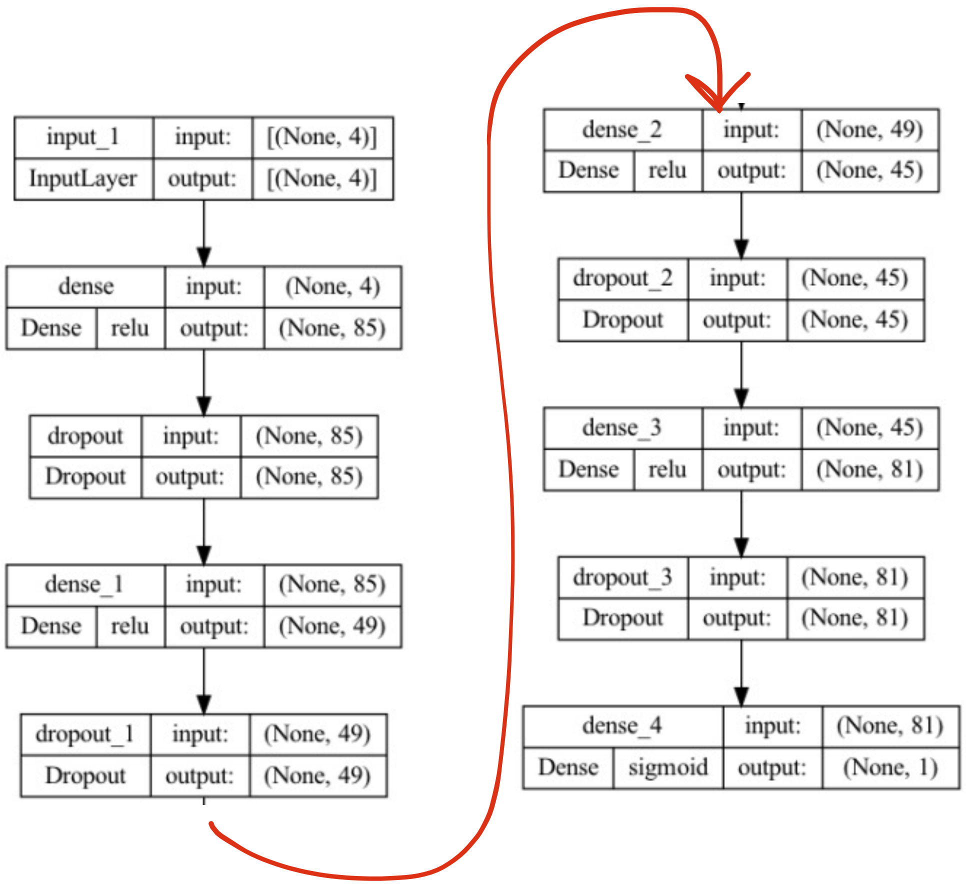
GNN Tracking

Tracking

Misc

Analysis: PNN architectures

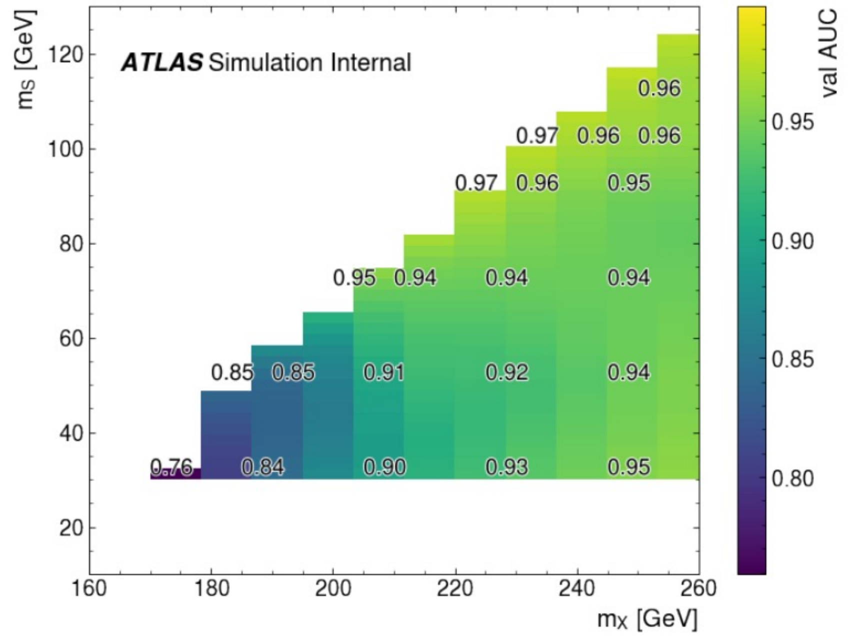
2 b-tagged category

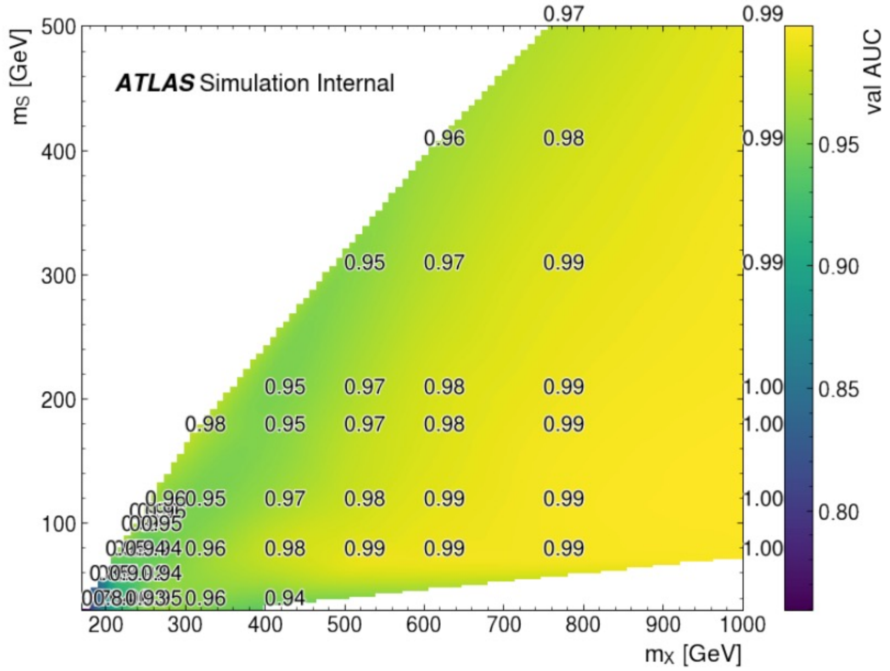


Number of hidden layers	4
Layer 1 dropout rate	0.05
Layer 2 droport rate	0.1
Layer 3 dropout rate	0.2
Layer 4 dropout rate	0.1
Learning rate	0.009137
Optimizer	Adam
Loss function	Binary Loss
Initial bias	0.118
Signal class weight	0.945
Background class weight	1.062
Batch size	212613
Number of batches	2
Number of epochs	2000

In both category, final activation function = sigmoid
Automatically tuned with Keras Tuner (maximize AUC)
Similar for 1 b-tagged category

Analysis: PNN architectures

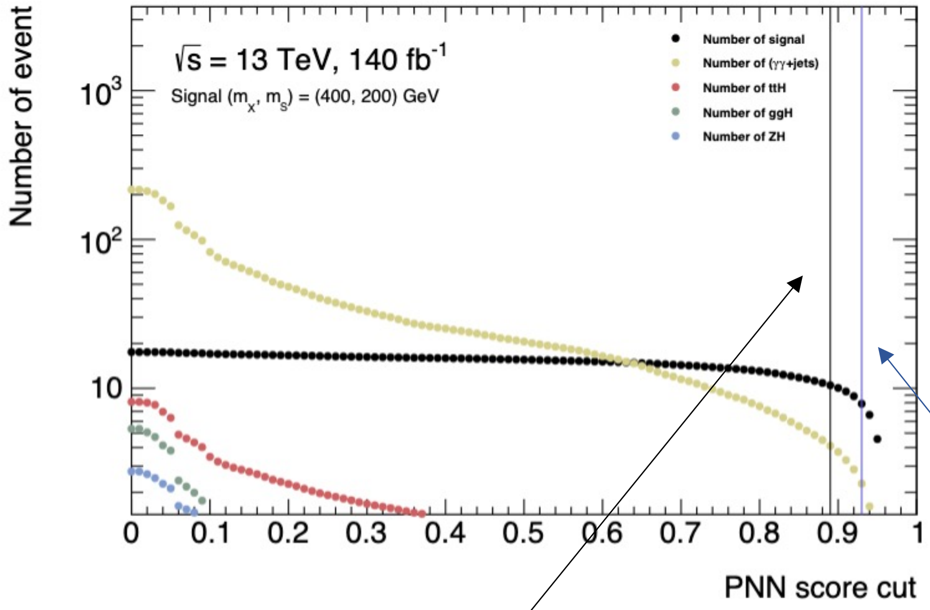




Optimization of the PNN classification cuts

We know from previous analysis, that we will be able to use the asymptotic approximation if we guarantee to have at least 9 event in the $m_{\gamma\gamma}$ data side-band. We want to maximize the significance:

$$Z_A = \sqrt{2\left[(S+B)\ln(1+S/B) - S\right]}\ ,$$



Black line: value of cut chosen, maximizing the significance.

Blue line: above this cut value, we do not have 9 events in the $m_{\gamma\gamma}$ data side-band.

(m_X, m_S) [GeV]	PNN cut value	(m_X, m_S) [GeV]	PNN cut value	(m_X, m_S) [GeV]	PNN cut value
(170, 30)	0.56	(245, 50)	0.92	(500, 300)	0.91
(180, 50)	0.81	(245, 70)	0.92	(600, 70)	0.86
(185, 30)	0.80	(245, 90)	0.91	(600, 170)	0.88
(190, 15)	0.70	(250, 15)	0.78	(600, 200)	0.87
(190, 50)	0.90	(250, 100)	0.91	(600, 300)	0.88
(200, 70)	0.92	(250, 110)	0.91	(600, 400)	0.91
(205, 30)	0.91	(300, 30)	0.92	(750, 70)	0.97
(205, 50)	0.91	(300, 70)	0.90	(750, 110)	0.93
(210, 15)	0.75	(300, 110)	0.92	(750, 170)	0.84
(210, 70)	0.92	(300, 170)	0.92	(750, 200)	0.81
(220, 90)	0.91	(400, 30)	0.88	(750, 300)	0.88
(225, 30)	0.92	(400, 70)	0.90	(750, 400)	0.81
(225, 50)	0.92	(400, 110)	0.89	(750, 500)	0.82
(225, 70)	0.92	(400, 170)	0.88	(1000, 70)	0.98
(230, 15)	0.74	(400, 200)	0.89	(1000, 110)	0.61
(230, 90)	0.91	(500, 30)	0.92	(1000, 170)	0.82
(230, 100)	0.91	(500, 70)	0.91	(1000, 200)	0.76
(240, 100)	0.91	(500, 110)	0.89	(1000, 300)	0.73
(240, 110)	0.90	(500, 170)	0.90	(1000, 400)	0.55
(245, 30)	0.92	(500, 200)	0.87	(1000, 500)	0.67

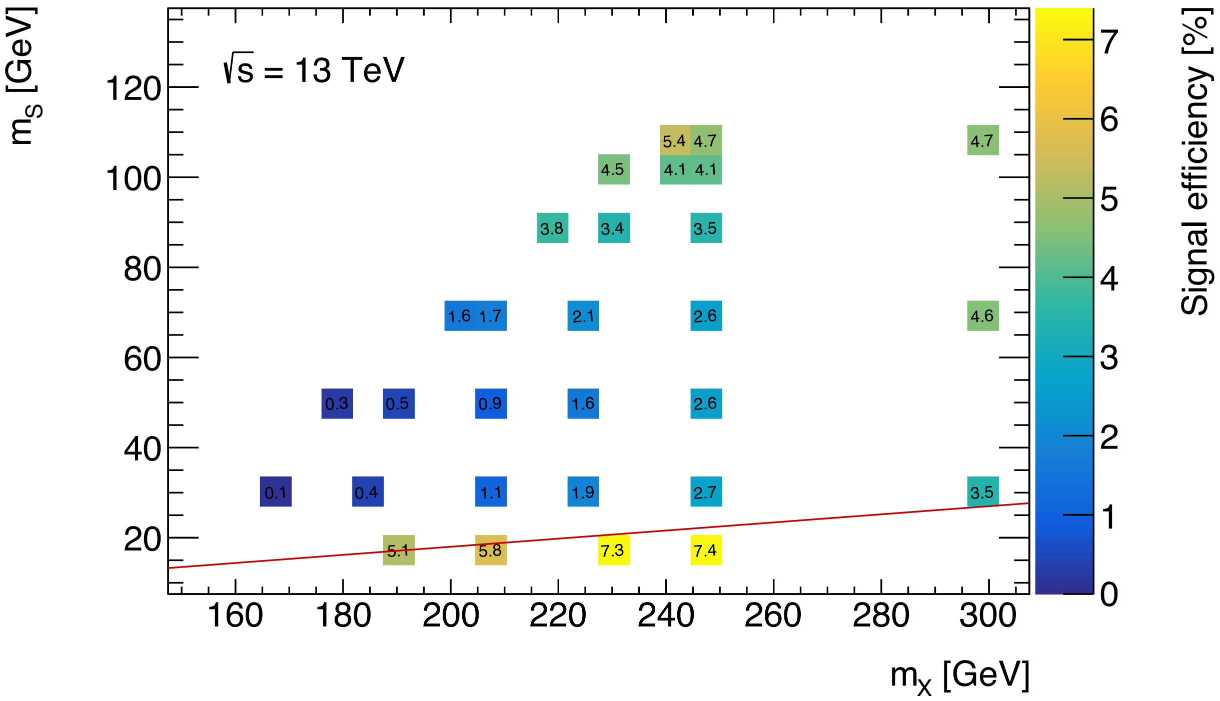
Blue cell = 1 b-tagged category

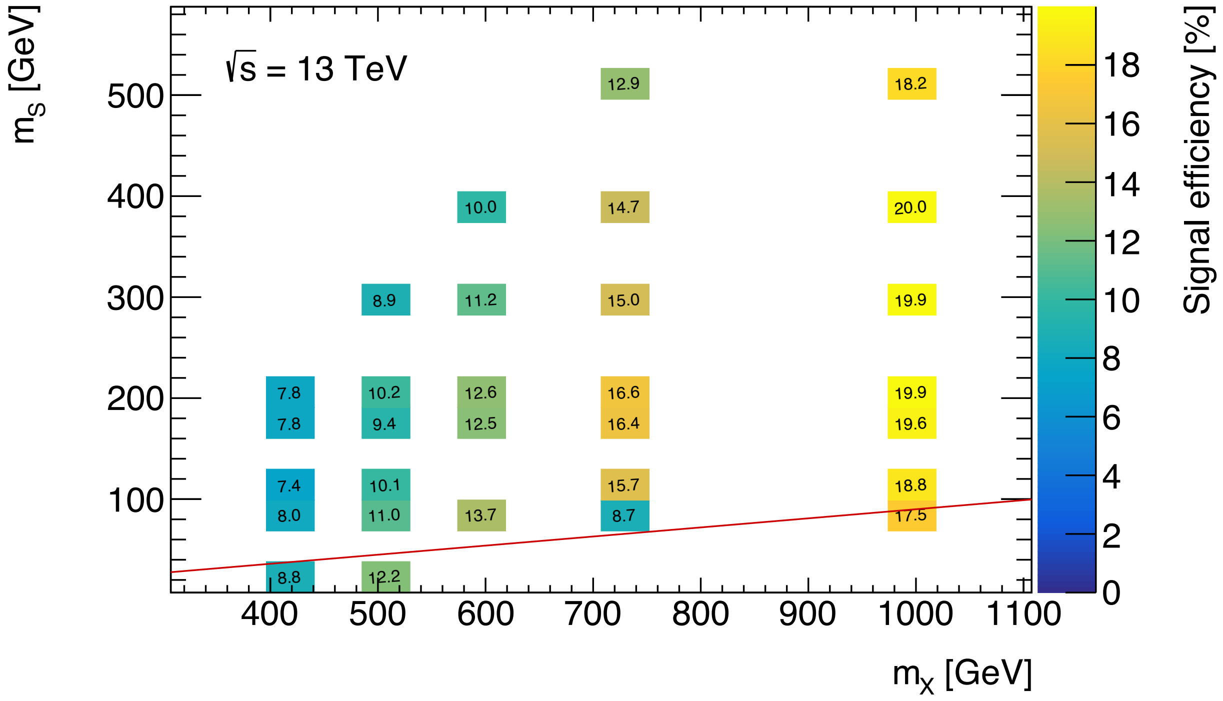
Analysis: cutflows before PNN cuts

Table 6: Cutflows for some generated MC signal sample. A signal cross-section of 1 fb is considered.

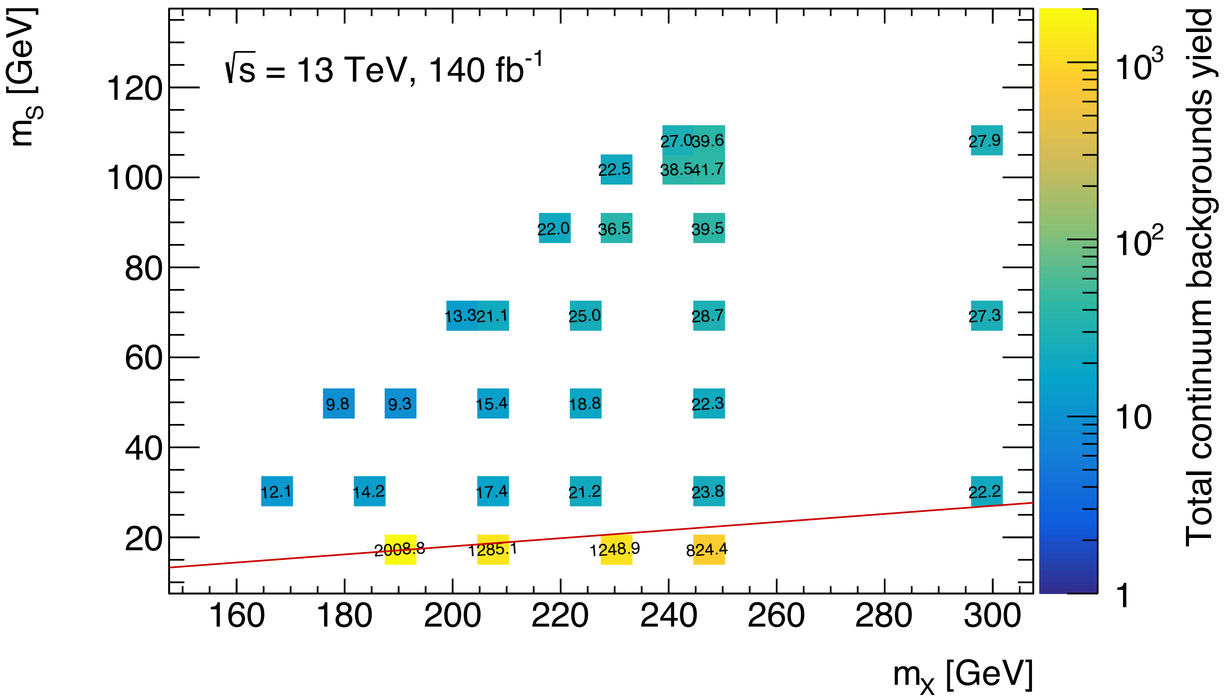
Selection	(190,15)		(250,110)		(600, 170)		(1000, 300)	
	Yield	Efficiency	Yield	Efficiency	Yield	Efficiency	Yield	Efficiency
All events	148.18	100.0%	148.78	100.0%	144.19	100.0 %	143.64	100.0 %
Pass trigger	93.41	63.0%	97.40	65.5%	126.71	87.9%	133.82	93.2 %
Has primary vertex	93.41	63.0%	97.40	65.5%	126.71	87.9%	133.82	93.2%
2 loose photons	73.44	49.6%	77.38	52.0%	86.76	60.2%	95.58	66.5%
$e-\gamma$ ambiguity	73.40	49.5%	77.35	52.0%	86.72	60.1%	95.55	66.5%
Trigger match	66.94	45.2%	70.35	47.3%	84.48	58.6%	94.85	66.0%
Photon tight ID cut	57.19	38.6%	59.72	40.1%	71.82	49.8%	80.21	55.8 %
Photon isolation cut	49.60	33.5%	49.80	33.5%	65.62	45.5%	74.99	52.2%
Rel. p_{T} cuts	45.59	30.8%	46.01	30.9%	60.46	41.9%	71.61	49.9%
$m_{\gamma\gamma} \in [105, 160]$ GeV	45.56	30.7%	45.96	30.9%	60.37	41.9%	71.48	49.8%
$\mathrm{N_{lepton}} == 0$	45.53	30.7%	45.9	30.9%	60.30	41.8%	71.34	49.7%
$\mathrm{N_{centraljets}} \in [2, 5]$	17.03	11.5%	30.42	20.4%	54.53	37.8%	68.62	47.8%
1 b-tagged selection	10.71	7.2%	14.67	9.9%	21.95	15.2%	24.31	16.9%
2 b-tagged selection	0.59	0.4%	9.94	6.7%	23.60	16.4%	30.83	21.5 %

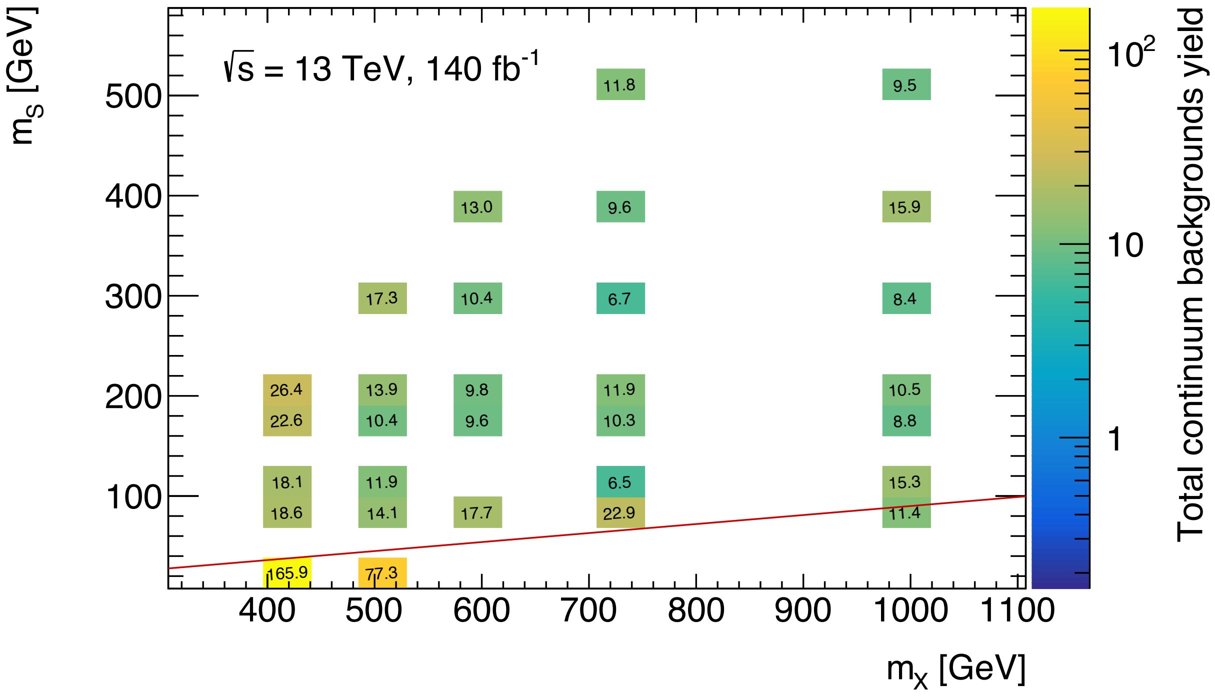
Analysis: signal efficiency



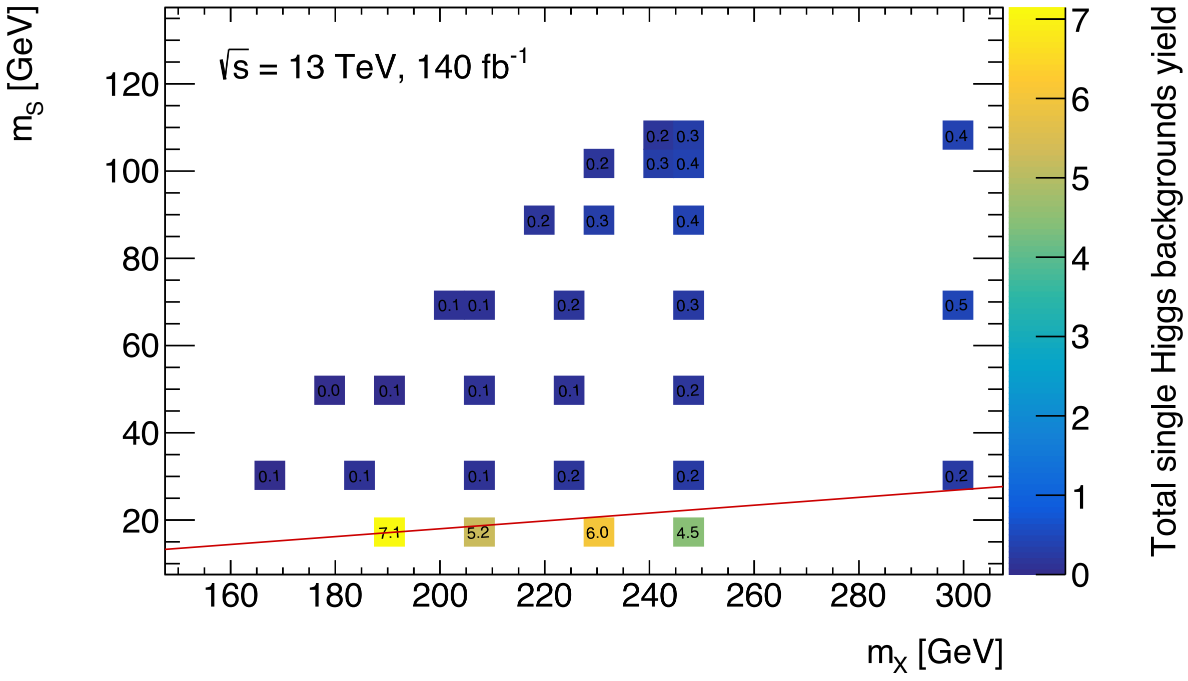


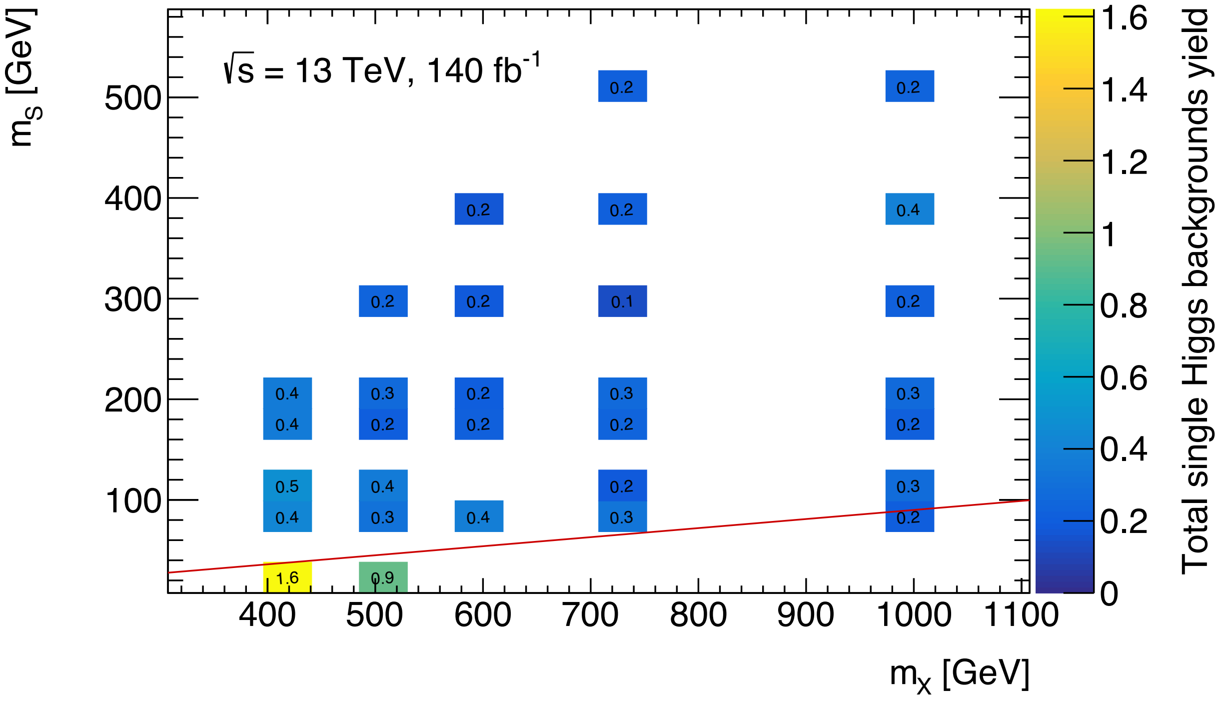
Analysis: continuum



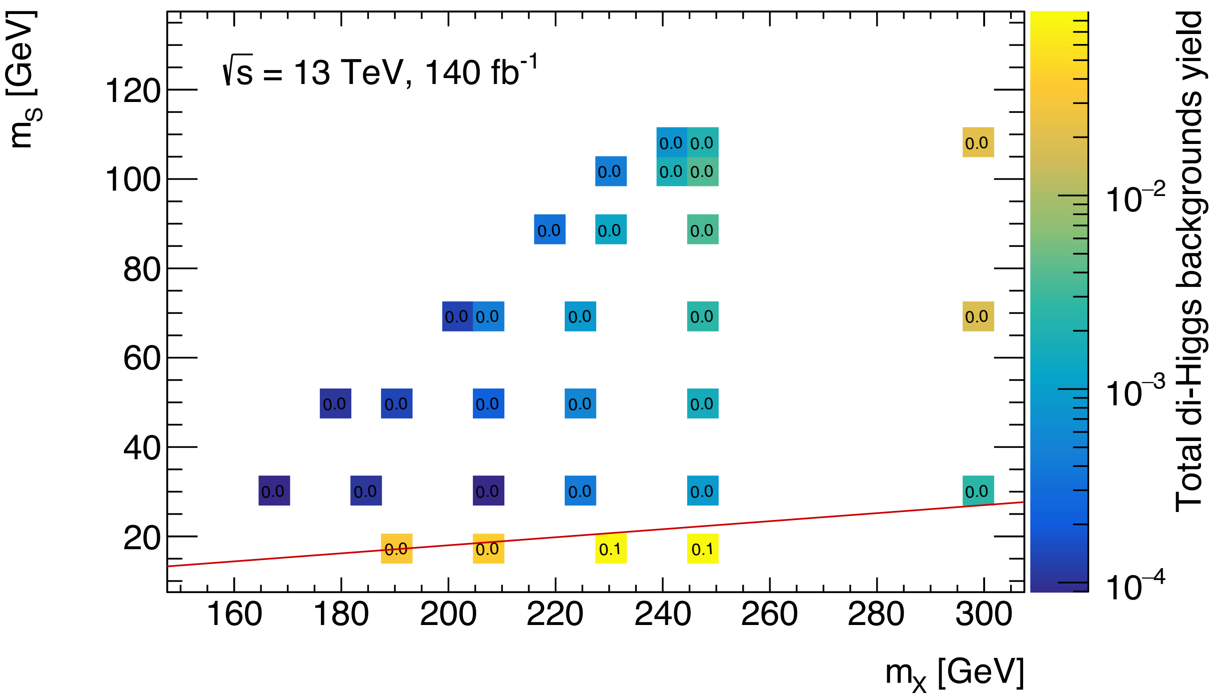


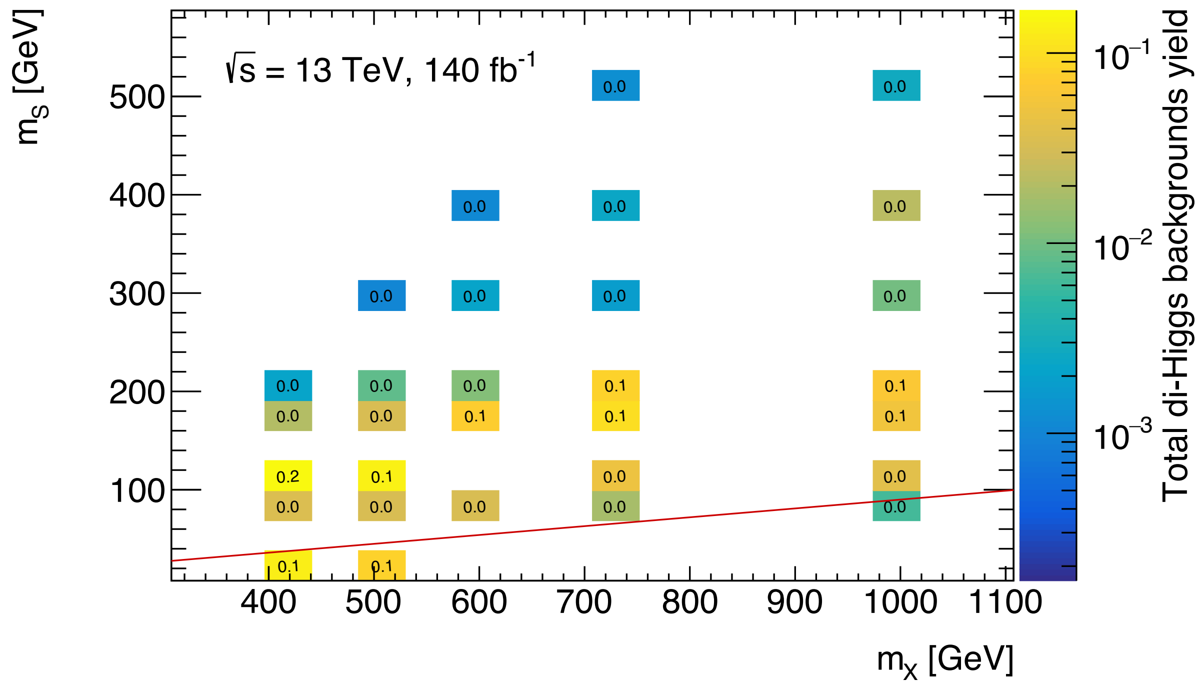
Analysis: single Higgs





Analysis: single Higgs





Tracking: timing

Timing consideration

- The target is to run the full pipeline in **< 1 second.**
- Need to be fully run on GPU.

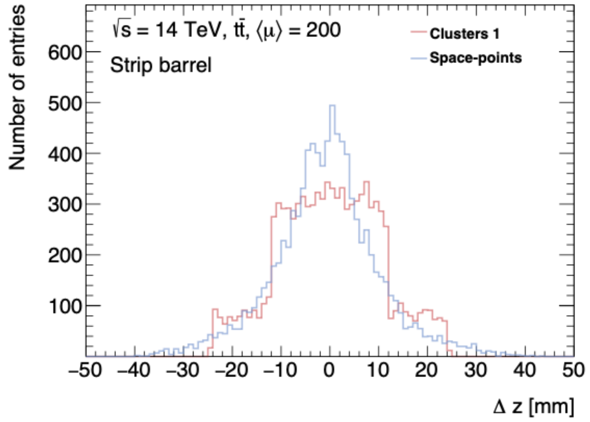
TrackML timing (Similar graph size as for ITk):

Pipeline step	V100 GPU
Graph construction (metric learning)	~ 500 ms
Graph construction (module map)	In progress target ~ 100 ms
GNN	~170 ms
Connecting component	~100 ms

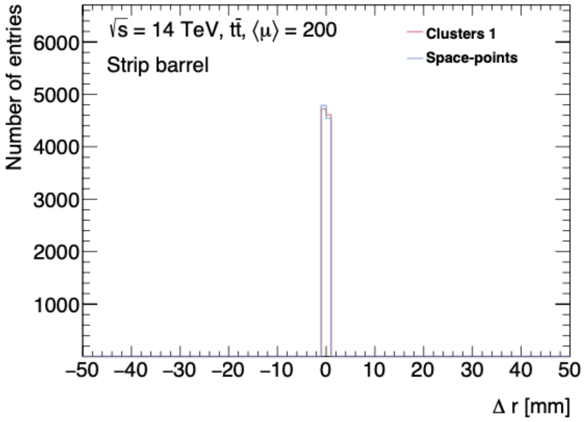
(See this paper)

How to improve: GPU kernels have dedicated operation for NN. But the GNN model is much complex with its 8 message-passing operations and the way the memory is therefore handle.
Using dedicated GNN kernels could only improve the timing, the memory consumption and the energy cost.

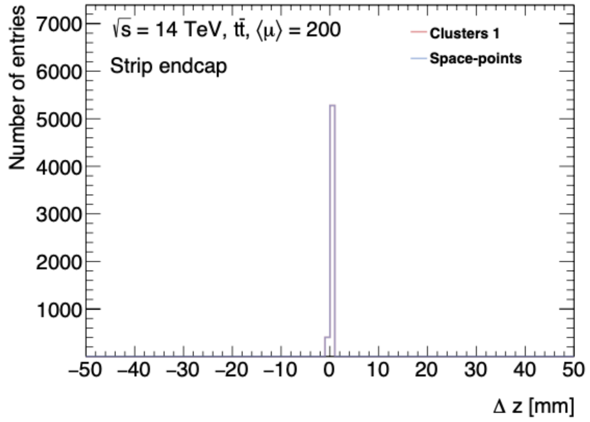
Tracking: resolution of clusters and SPs in strip



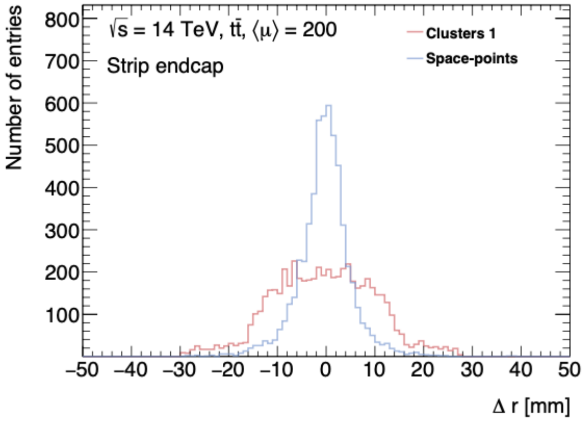
(a) Cluster and space-point resolution vs z.



(b) Cluster and space-point resolution vs r.



(c) Cluster and space-point resolution vs z.



(d) Cluster and space-point resolution vs r.

Figure 12.4: Cluster and space-point resolution in the ITk strip subdetector. The resolution is defined as the difference in z or r between the simulated Geant4 hit positions and the reconstructed cluster and space-point ones.

Tracking QT: validation of the seeder

- **Motivation**

Porting and validation of the new ITk seed maker algorithm of r22 to r21.9. This back-porting allows us to study the new seed maker with well-understood simulated samples produced using r21.9.

Jira ticket here

Creation of a Seed Maker compatible with multi-threading:

ITk Seed Maker
of release 22
ITkSiSpacePointsSeedMaker

backward-ported into release 21.9

Named **New ITK seeder** later

ITk Seed Maker
of release 22
backported into release 21.9
SiSpacePointsSeedMaker_ITKNew

Default ITk Seed Maker
of release 21.9
SiSpacePointsSeedMaker_ITK

Named **default seeder** later

Test the release 22 seed maker and upgrade it if needed

Two seed makers: validation of the release 22 seeder is now possible by comparing to the default ITk seeder of release 21.9.

Tracking QT: validation of the seeder

21.9

- **Comparison of the new ITk seeder with the default one**

Using the 400 $t\bar{t}$ events with ⟨μ⟩=200 :

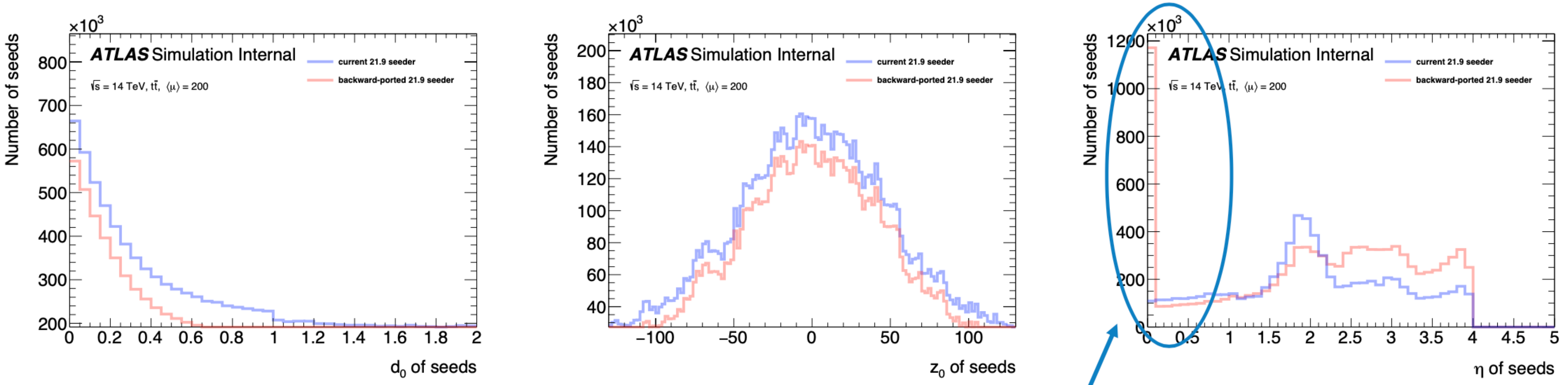
	Default seeder	New seeder	Ratio new / default
Total number of seeds	12 419 289	9 866 806	**79.4%**
PPP seeds	10 234 093	7 702 037	**62.0%**
SSS seeds	2 185 196	2 164 769	**99.0%**
Total number of seeds giving a track	1 383 446	1 336 765	**96.6%**
PPP seeds giving a track	932 797	888 978	**95.3%**
SSS seeds giving a track	450 649	447 787	**99.4%**
Tracking efficiency	90.396%	90.282%	**99.99%**
Fake tracks rate for 9 clusters	$6.68{\times}10^{-3}$	$6.51{\times}10^{-3}$	**97.5%**

The New ITk seeder find less seeds than the default one, especially in the pixel. The tracking efficiency is slightly degraded as well.

Tracking QT: validation of the seeder

21.9

- **Comparison of distribution of seeds**



- The New ITk seeder produces no seeds for $d_0 > 0.7$
- η distribution is buggy, but it is only use in the validation Ntuple.

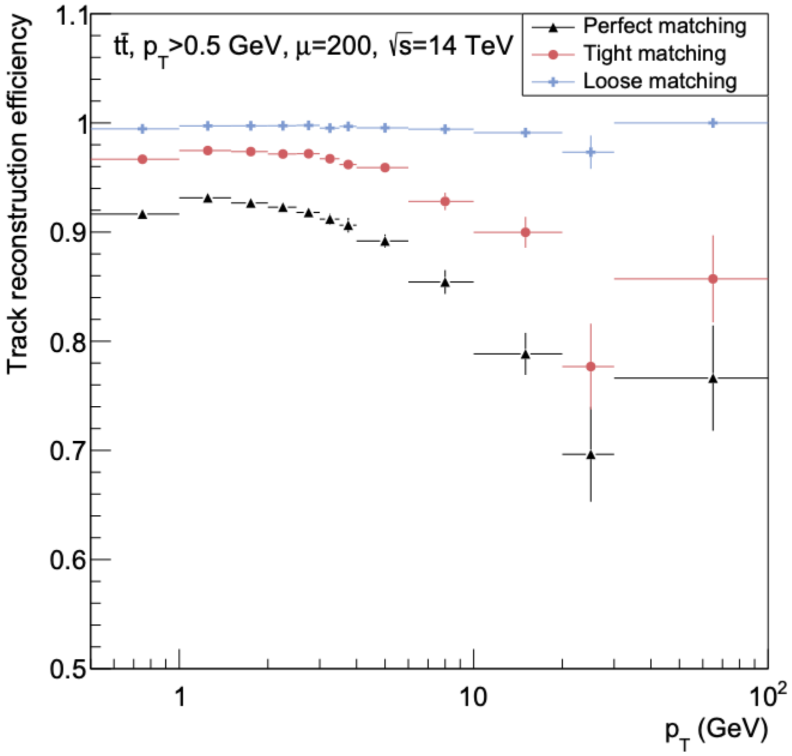
➢ Let's hunt bugs

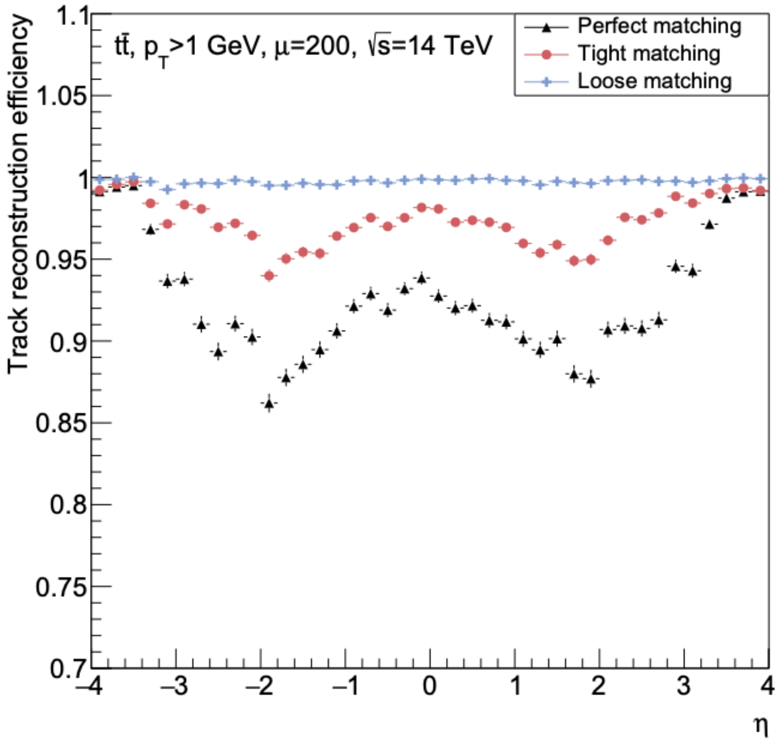
Tracking QT: validation of the seeder

	Default seeder	New seeder	Ratio new / default
Total number of seeds	12 419 289	12 563 203	101.2%
PPP seeds	10 234 093	10 378 008	101.4%
SSS seeds	2 185 196	2 185 195	100.0%
Total number of seeds giving a track	1 383 446	1 389 585	100.4%
PPP seeds giving a track	932 797	938 936	100.6%
SSS seeds giving a track	450 649	450 649	100.0%
Tracking efficiency	90.396%	90.398%	100.0%
Fake tracks rate for 9 clusters	6.68×10^{-3}	6.69×10^{-3}	100.0%

This does not degrade the tracking efficiency. More seeds are produced, so the tracking efficiency is slightly improved.

Tracking: TrackML GNN results



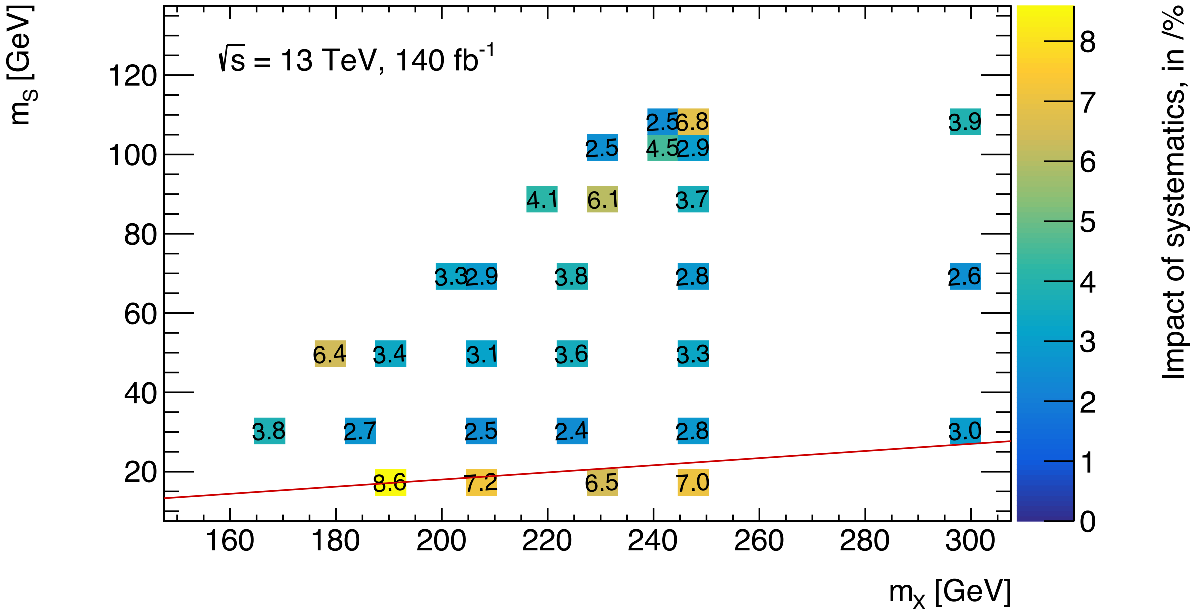


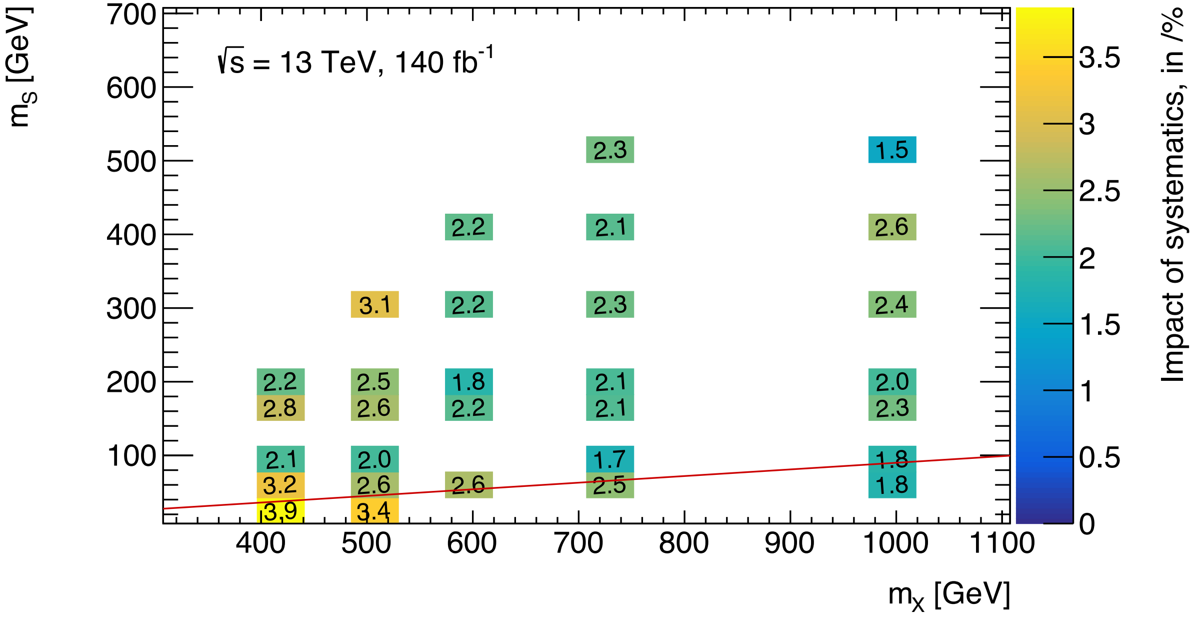
https://arxiv.org/abs/2103.00916

Summary of included systematics

<table>
<tr><th></th><th></th><th>Signal</th><th>ttH</th><th>ggH</th><th>ZH</th><th>Others single Higgs</th><th>Di-Higgs</th><th>Continuum</th></tr>
<tr><td>Theory</td><td>Yield</td><td>Parton shower</td><td>QCD scale
PDF+α_s</td><td>QCD scale
PDF+α_s</td><td>QCD scale
PDF+α_s</td><td>QCD scale
PDF+α_s</td><td>QCD scale
PDF+α_s</td><td></td></tr>
<tr><td rowspan="2">Experimental</td><td>Yield</td><td colspan="4">Luminosity
Photon trigger efficiency
Pile-up scale factors
Photon energy resolution and scale
Photon identification efficiency
Photon isolation efficiency
Jet energy resolution and scale
Jet η calibration
Flavor tagging efficiencies
Jet Vertex tagger
Jet pile-up offset</td><td>Negligeable</td><td>Negligeable</td><td>Spurious signal</td></tr>
<tr><td>Shape</td><td colspan="4">Photon energy resolution and scale
Photon identification efficiency
Photon isolation efficiency
Photon trigger efficiency</td><td>Negligeable</td><td>Negligeable</td><td></td></tr>
</table>

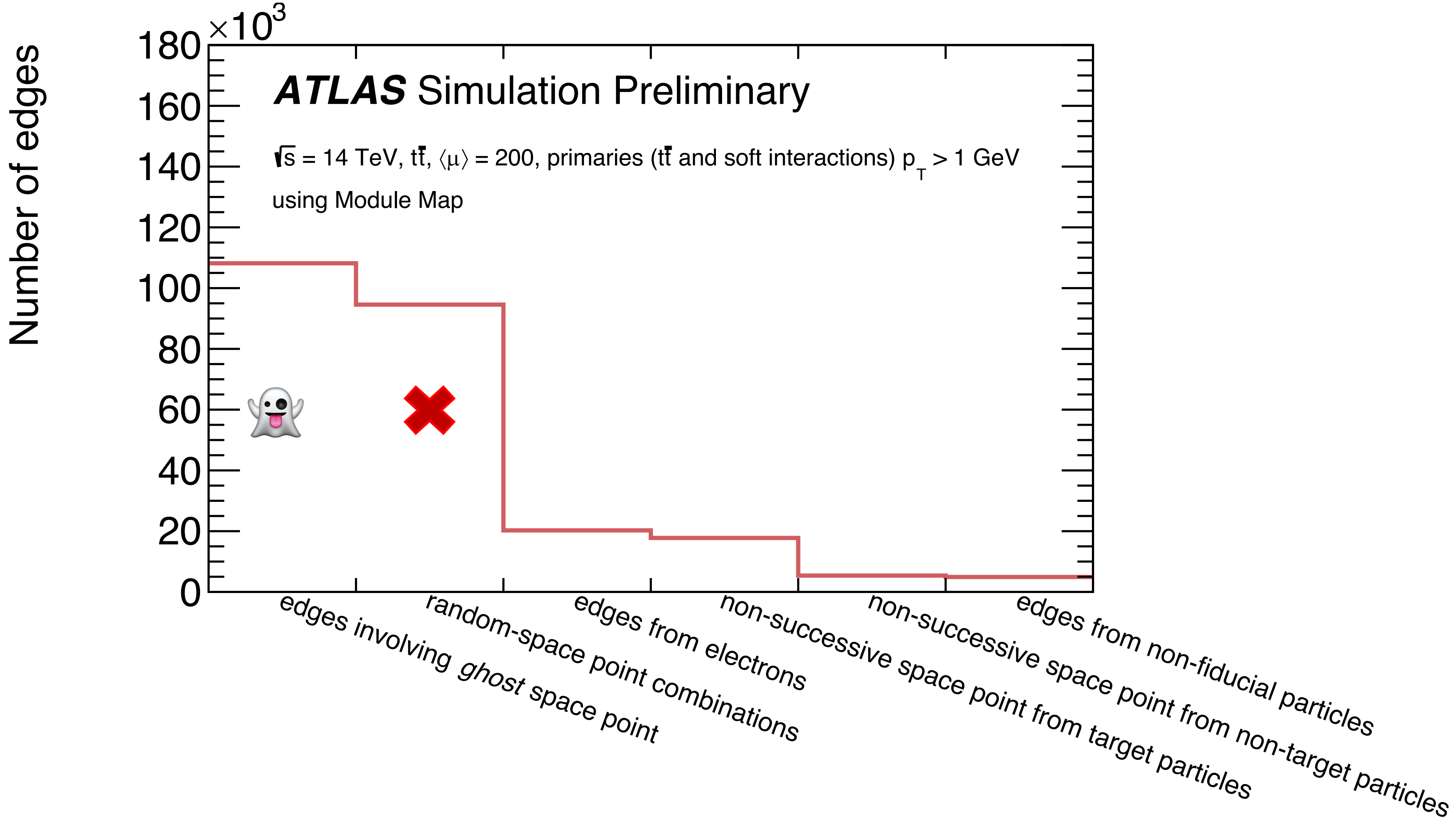
Analysis: impact systematics





Investigation of the GNN edge-performance

Non-fiducial particles = particles with $|\eta| > 4$ or $r_{vertex} > 26$ cm

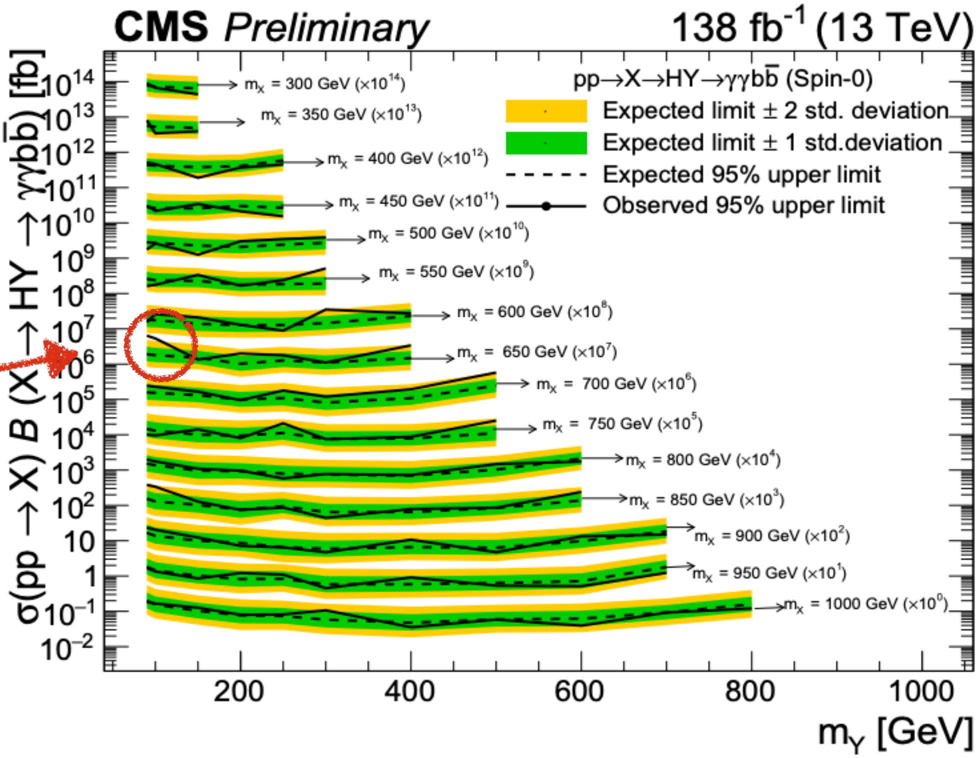


Previous CMS results

CMS already released public results on this analysis.

No significant excesses are seen.

Although, local (global) significance of 3.8 (2.8) standard deviations is observed for $\left(m_X, m_{S(Y)}\right) = (650, 90)$ GeV.



Discovery and upper limits

Testing for discovery

→ Background only hypothesis H_0

Number of expected events

$$\mathcal{L} = \frac{e^{-N}}{n!} \prod_{i=1}^{n} p(m_{\gamma\gamma}^{i}) \prod_{s=1}^{n_{syst}} C_s(\theta_s; \alpha_s, \delta_s)$$

Systematic uncertainties

PDF describing the shape of the distribution

	Significance Z	$1 - p_0$ [%]
Evidence	3σ	99.7%
Discovery	5σ	99.9997%

Probability to reject the background only hypothesis.

Exclusion limits

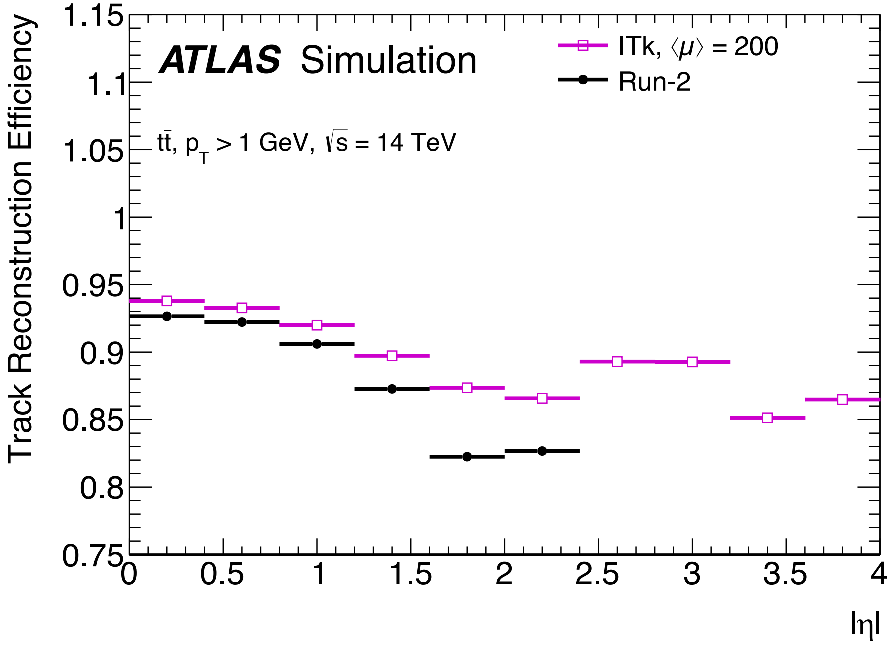
→ Signal + background hypothesis H_1

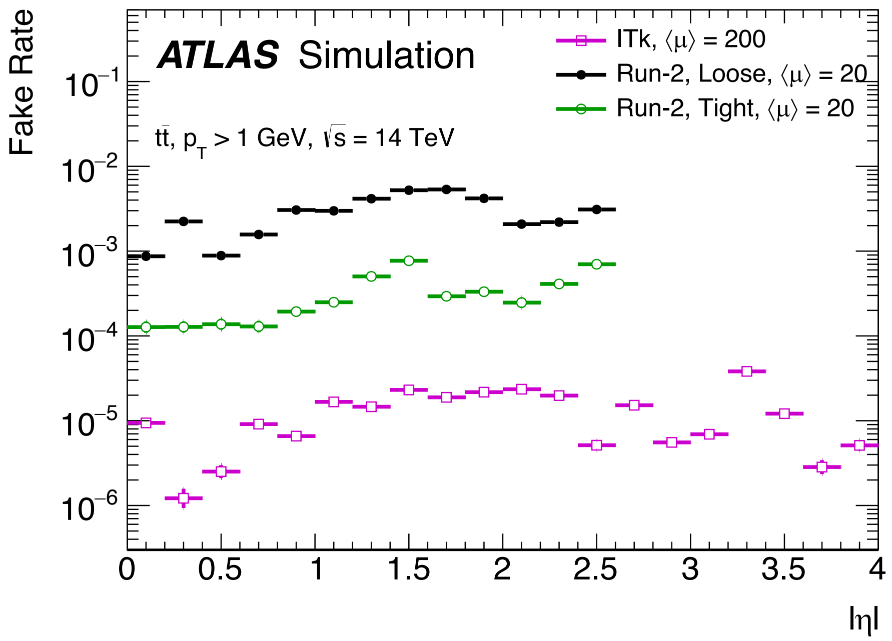
We set upper limits on the signal cross-section $\sigma(pp \to X \to SH \to b\bar{b}\gamma\gamma)$.

Track selection and CKF perf

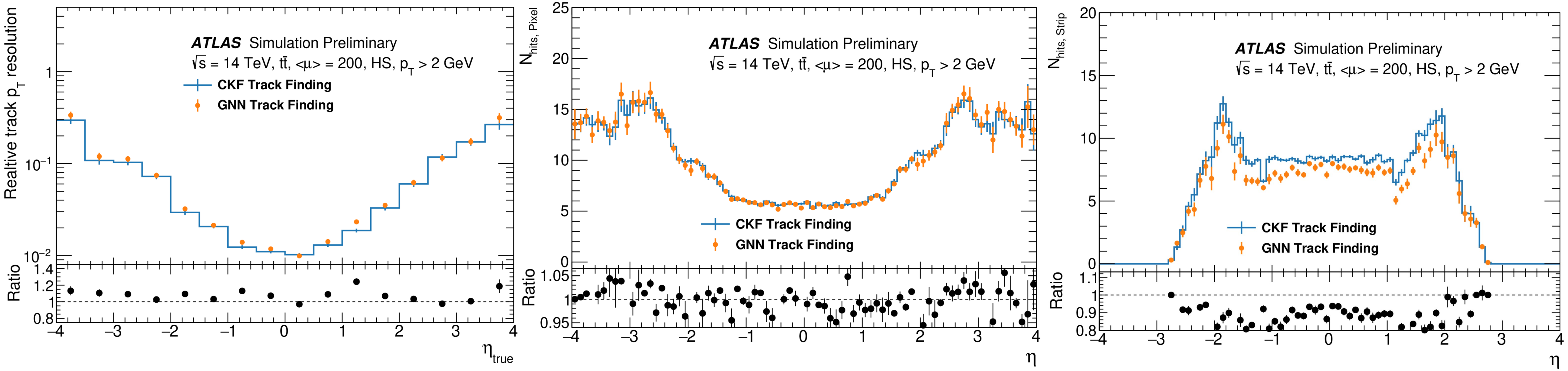
Requirements	$\lvert\eta\rvert < \mathbf{2.0}$	$\mathbf{2.0} < \lvert\eta\rvert < \mathbf{2.6}$	$\mathbf{2.6} < \lvert\eta\rvert < \mathbf{4.0}$
Number of pixel + strip hits	≥ 9	≥ 8	≥ 7
Number of pixel hits	≥ 1	≥ 1	≥ 1
Number of holes	< 2	< 2	< 2
Number of double holes	≤ 1	≤ 1	≤ 1
Number of pixel holes	< 2	< 2	< 2
Number of strip holes	< 2	< 2	< 2
p_T [MeV]	> 900	> 400	> 400
d_0 [mm]	≤ 2	≤ 2	≤ 10
z_0 [cm]	≤ 20	≤ 20	≤ 20

from ITk pixel TDR



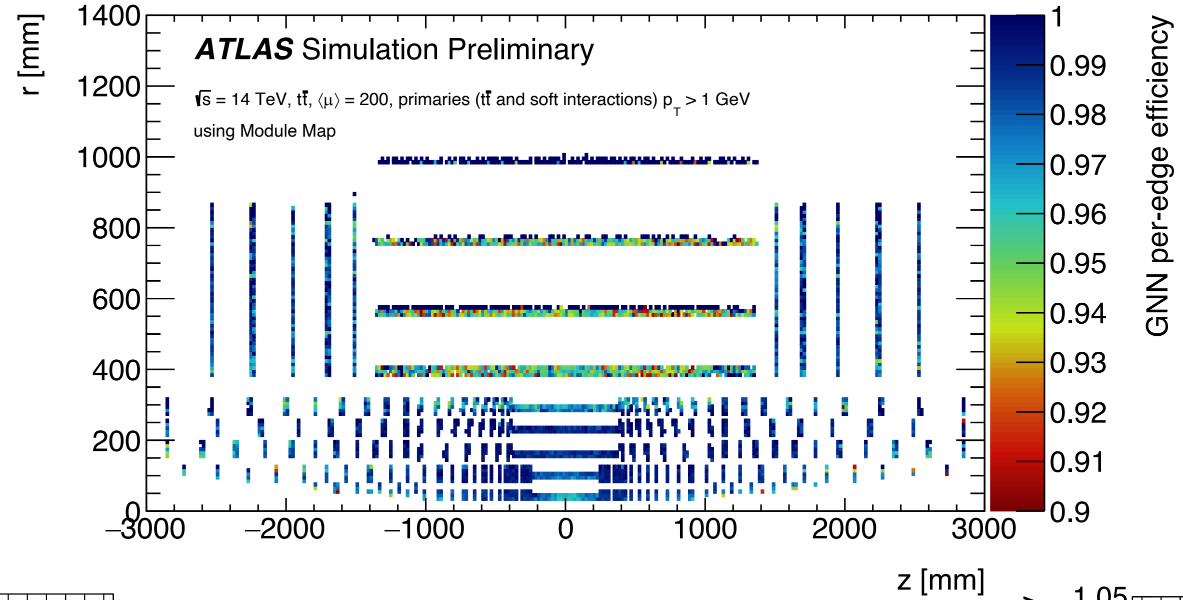


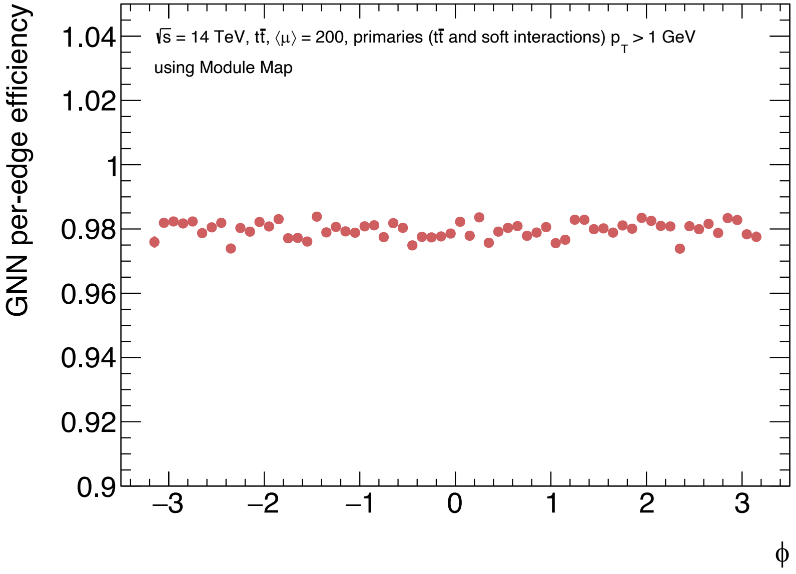
Comparison with CKF

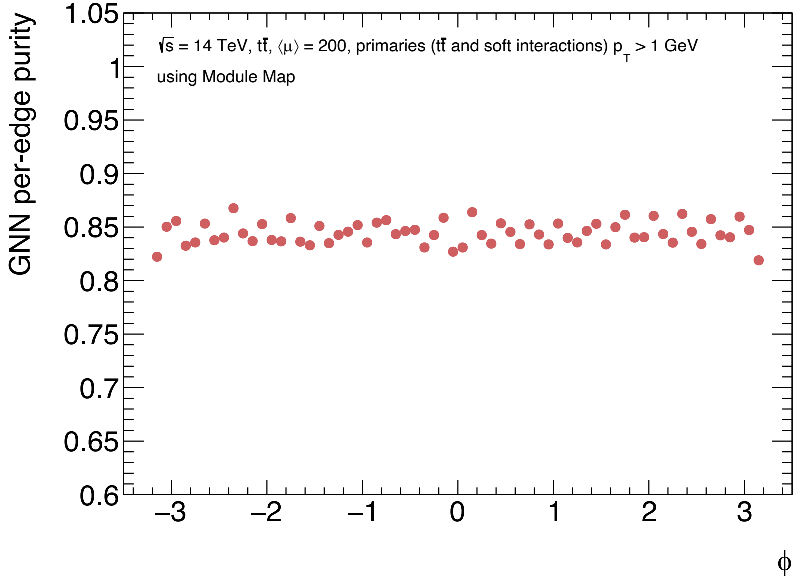


Average number of strip hits on a track as a function of the true pseudorapidity η_{true} for hard scattering primary particles from $t\bar{t}$ decays at $\langle\mu\rangle = 200$. The tracks, i.e. sets of hits, from the two track finding methods are fit using the method that is part of the standard reconstruction (ATL-PHYS-PUB-2019-014). Tracks found by the CKF and the subsequent ambiguity solution (ATLAS-TDR-025) are selected using the standard quality requirements (cf. Table 7 of ATL-PHYS-PUB-2019-014) and are required to satisfy $p_T > 1$ GeV. Tracks found by the GNN are required to satisfy criteria that are slightly less strict: at least 6 silicon hits, transverse impact parameter $|d_0| < 20$ mm, longitudinal impact parameter $|z_0| < 25$ cm and $p_T > 1$ GeV. The simulated particles matched to reconstructed tracks are required to satisfy $p_T > 2$ GeV to avoid turn-on effects. The ratio is defined as the number of hits for the GNN track finding divided by the number of hits for the CKF track finding. The same set of 100 simulated events is used to evaluate the track hit content obtained for tracks from CKF and from GNN track finding. The differences in the number of strip hits per track from the two track finding methods is largely due to the use of space points (built from hits on both sides of a given strip module) in the GNN track finding and the use of individual hits in the CKF track finding.

Edge classification performance







Inference benchmark

Prediction sample

- 100 events from TrackML
- Graphs have similar size as those obtained with ITK

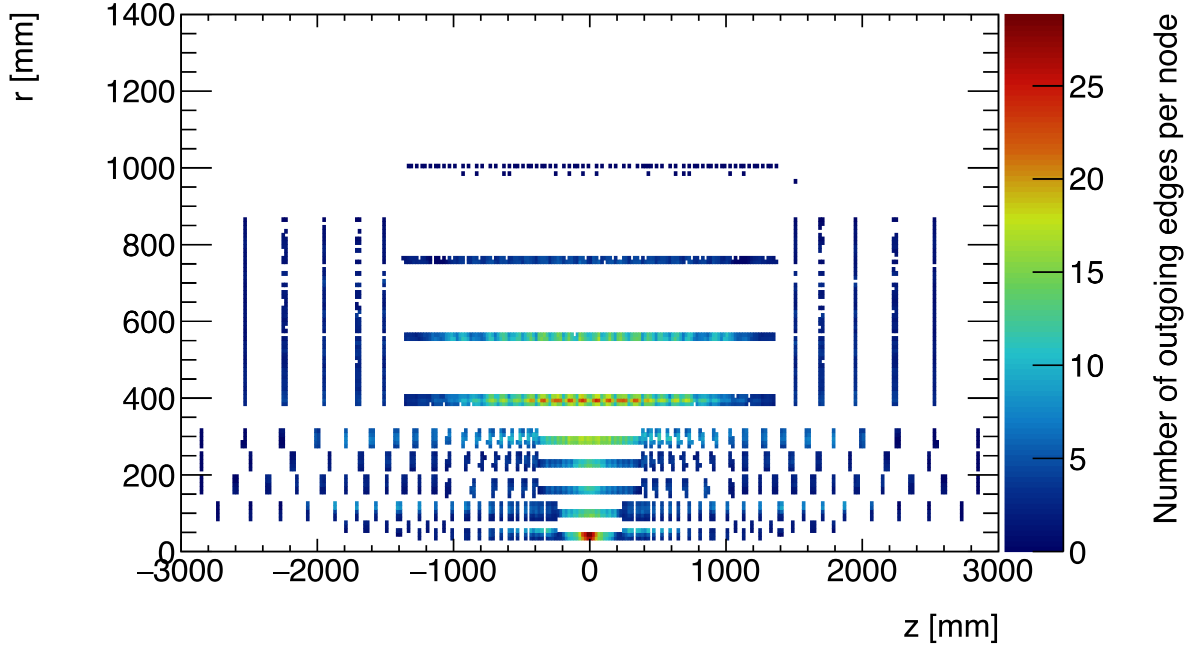
	Quadro RTX 8000	GeForce RTX 2080 Ti Gaming GPU
GPU memory capacity (GB)	48	11
Runtime mixed precision (16/32)	350 ms / event	
Memory peak consumption	5.4 GB	

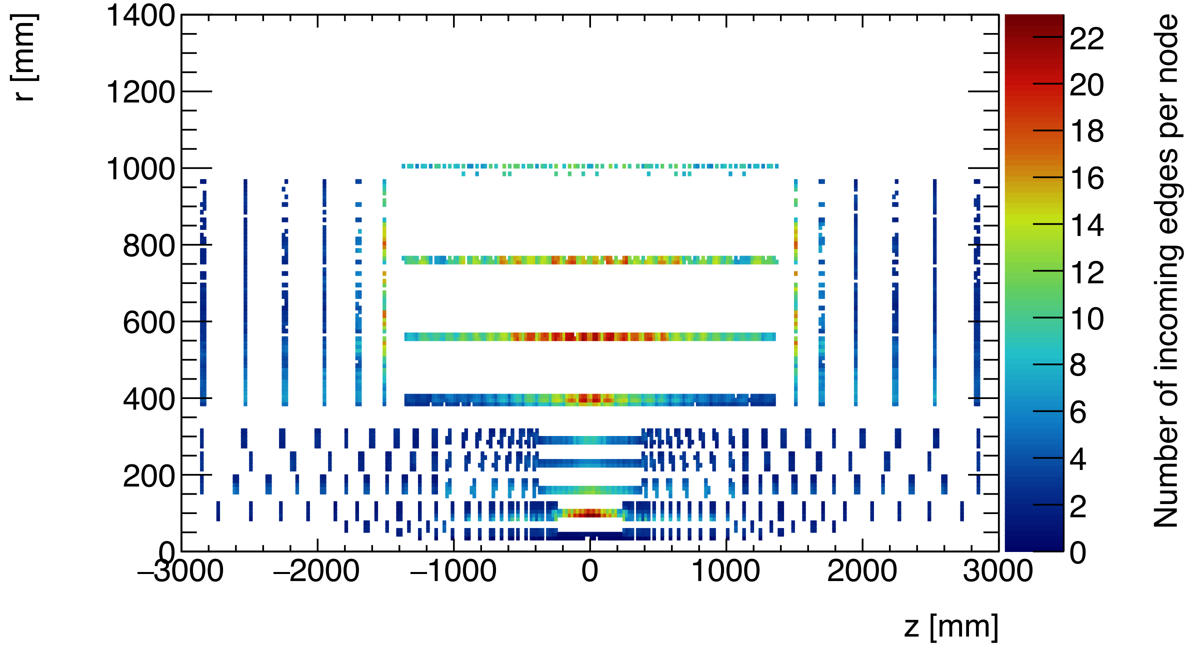
Study on TrackML sample, without optimization.

No memory issue during prediction

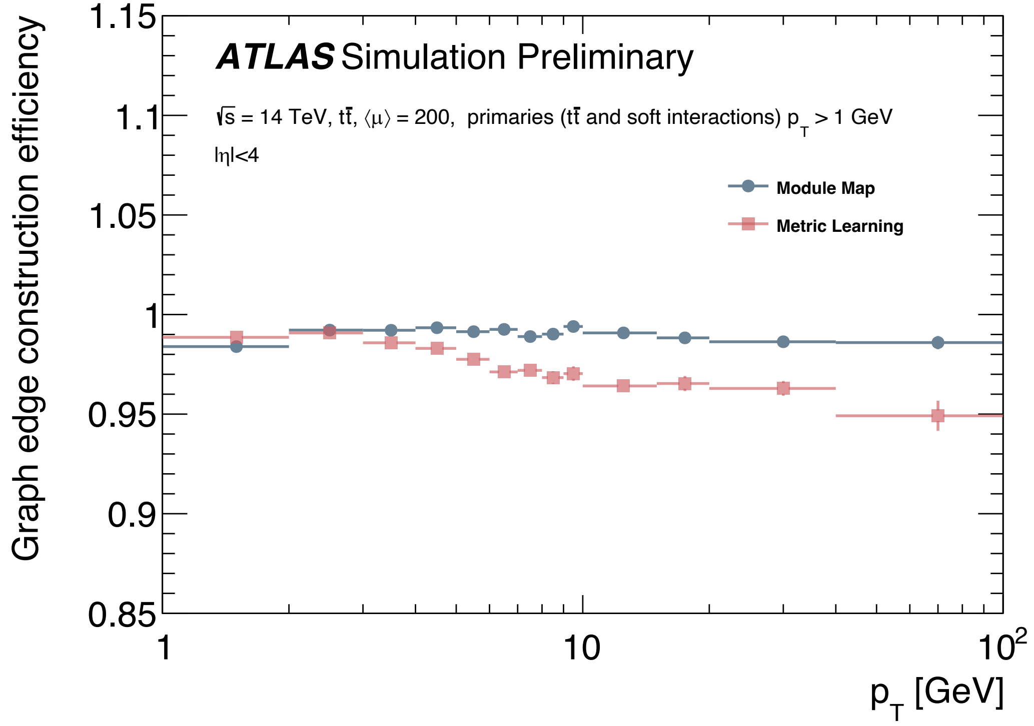
Will also benefit from dedicated kernel => large factor of improvement expected from dedicated CUDA kernel

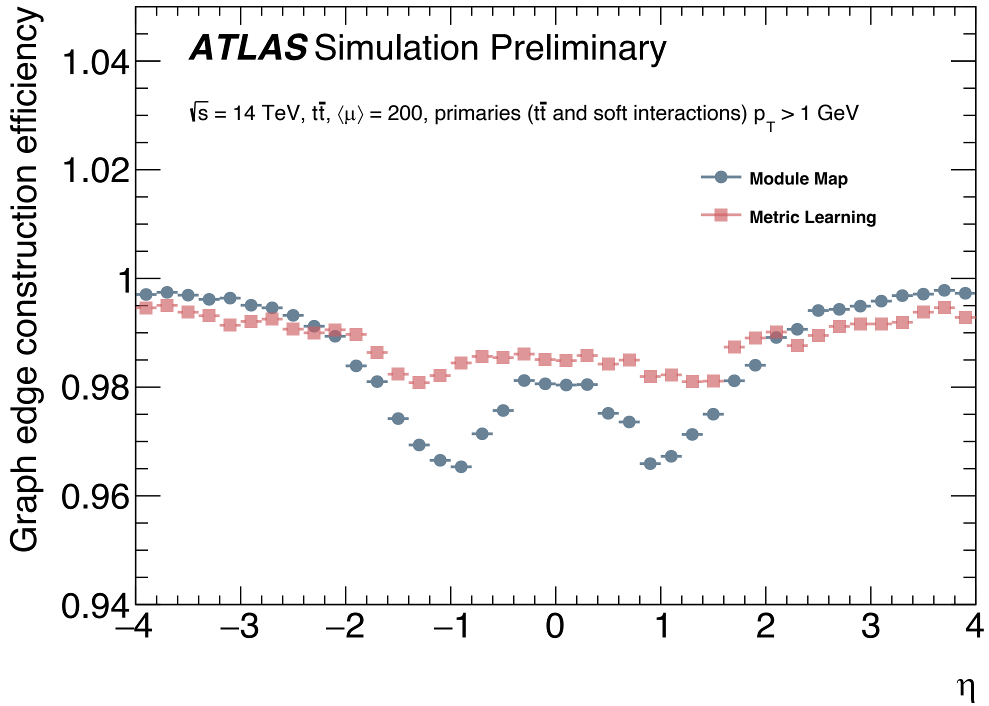
Graph topology



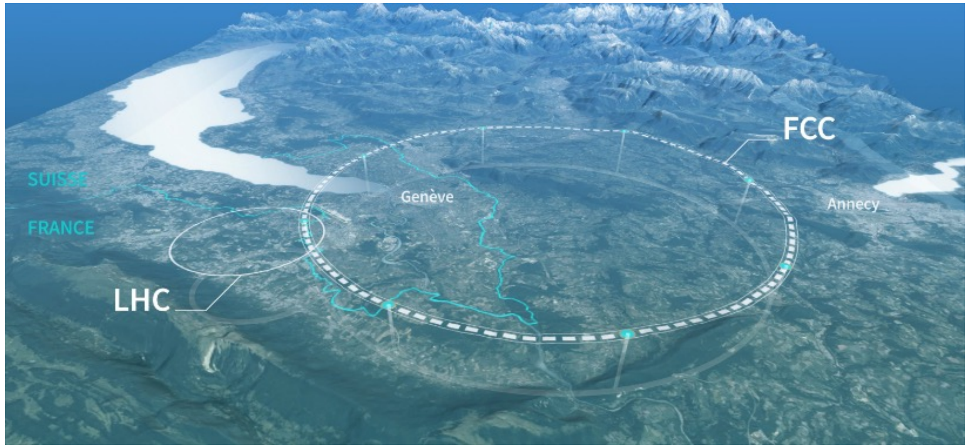


MM vs ML


ATLAS Simulation Preliminary
√s = 14 TeV, tt̄, ⟨μ⟩ = 200, primaries (tt̄ and soft interactions) p_T > 1 GeV
|η|<4
Module Map
Metric Learning
Graph edge construction efficiency
p_T [GeV]


ATLAS Simulation Preliminary
√s = 14 TeV, tt̄, ⟨μ⟩ = 200, primaries (tt̄ and soft interactions) p_T > 1 GeV
Module Map
Metric Learning
Graph edge construction efficiency
η

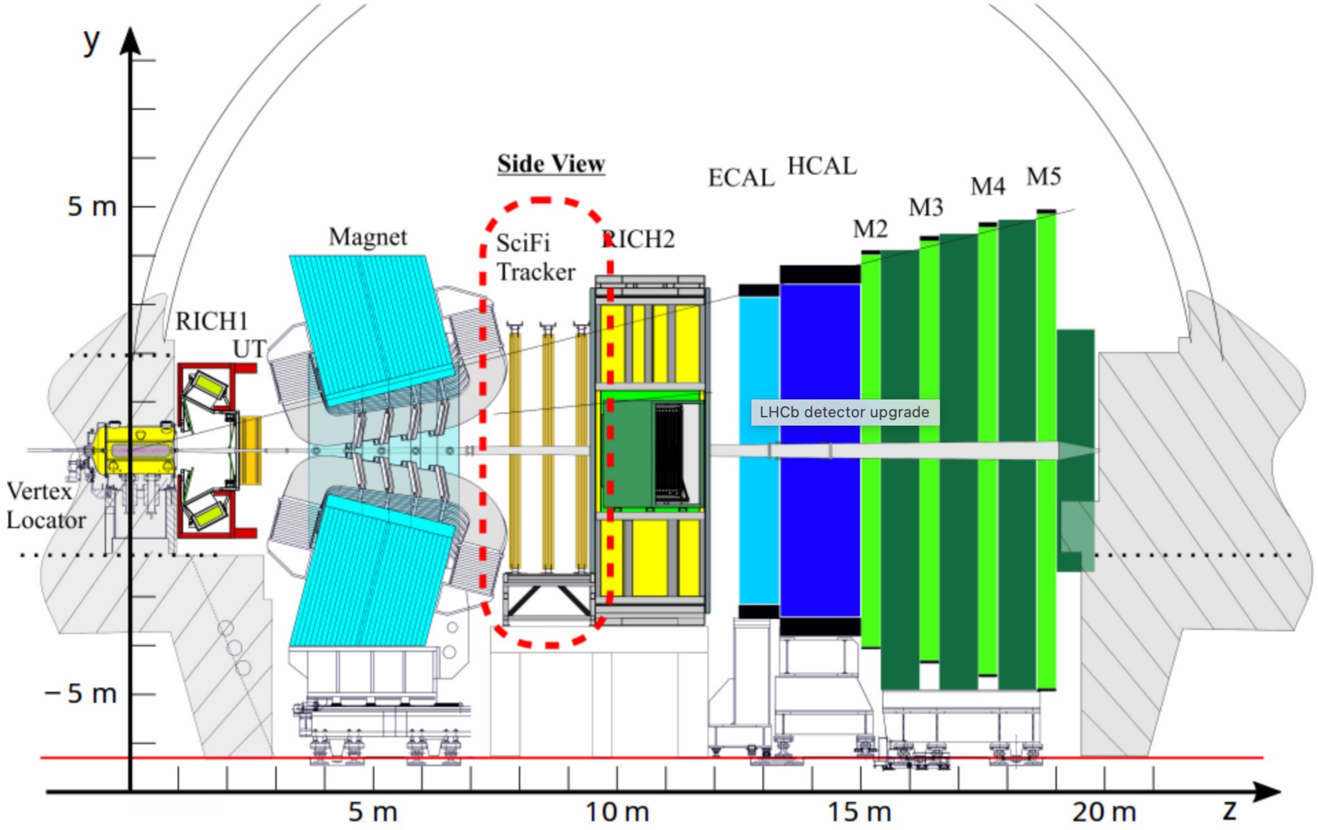
Future circular collider

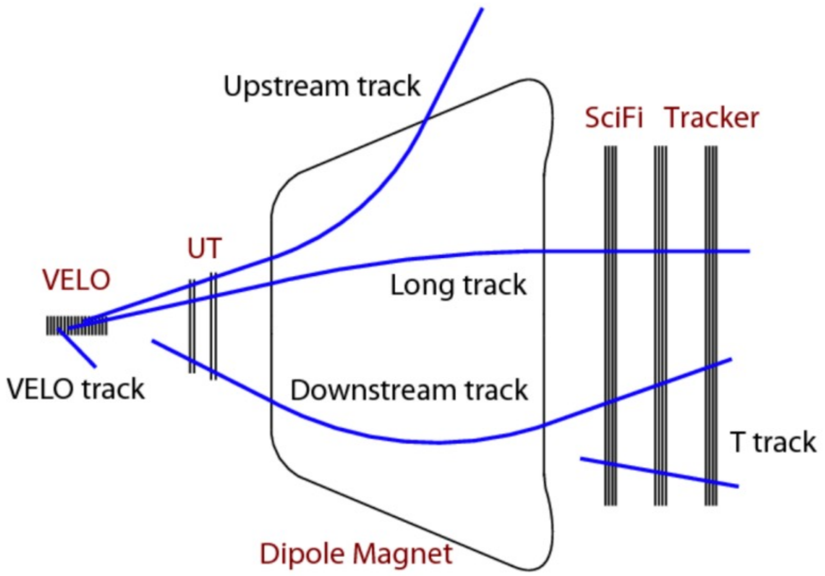


- ~ 91 km ring-shaped underground tunnel
- Start of construction = 2030
- Proto-collaboration are being formed
- at least 4 detectors

	Start	Operation	
e-e	~2040 After HL-LHC	15 years	Energies from 80 to 400 GeV
p-p	~2080	25 years	100 TeV

LHCb tracker





Single electron: event display