Search for doubly resonant beyond the Standard Model $X \to SH \to bb\gamma\gamma$ in the ATLAS experiment at the LHC

Grigorii Tolkachev

JRJC 2024

28/11/2024

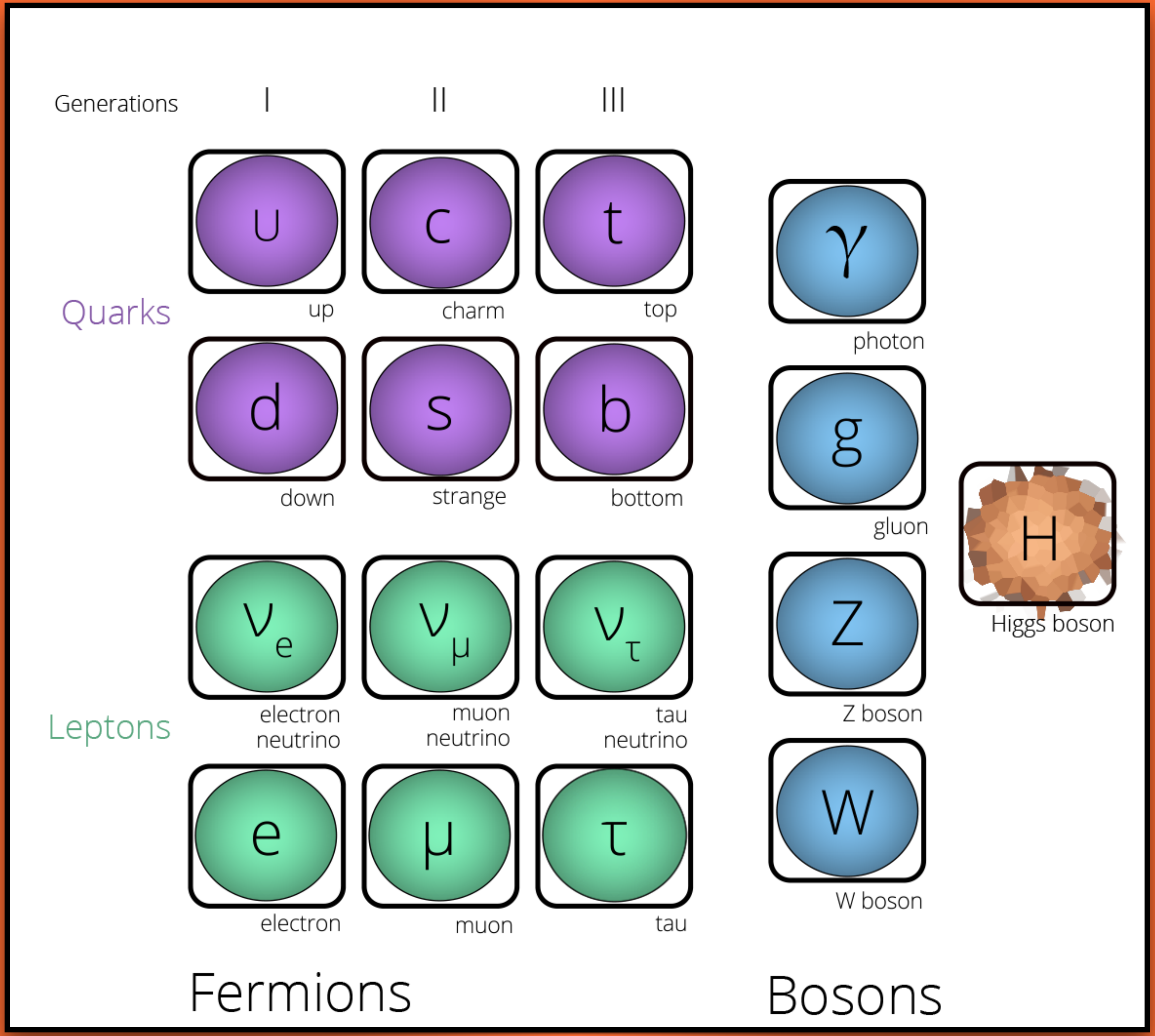
ATLAS EXPERIMENT

cnrs

Aix*Marseille université
Socialement engagée

CENTRE DE PHYSIQUE DES PARTICULES DE MARSEILLE
CPPM

The Standard Model



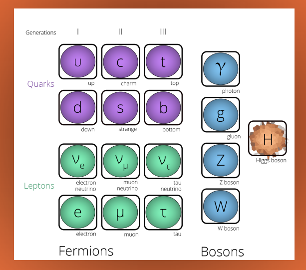
- The Standard Model explains the fundamental forces (except gravity) and particles
- Successfully predicts experimental results, including the Higgs boson discovery

BUT

- Fails to include gravity
- Provides no explanation for dark matter or dark energy
- Cannot fully account for neutrino oscillations and their nonzero masses

…

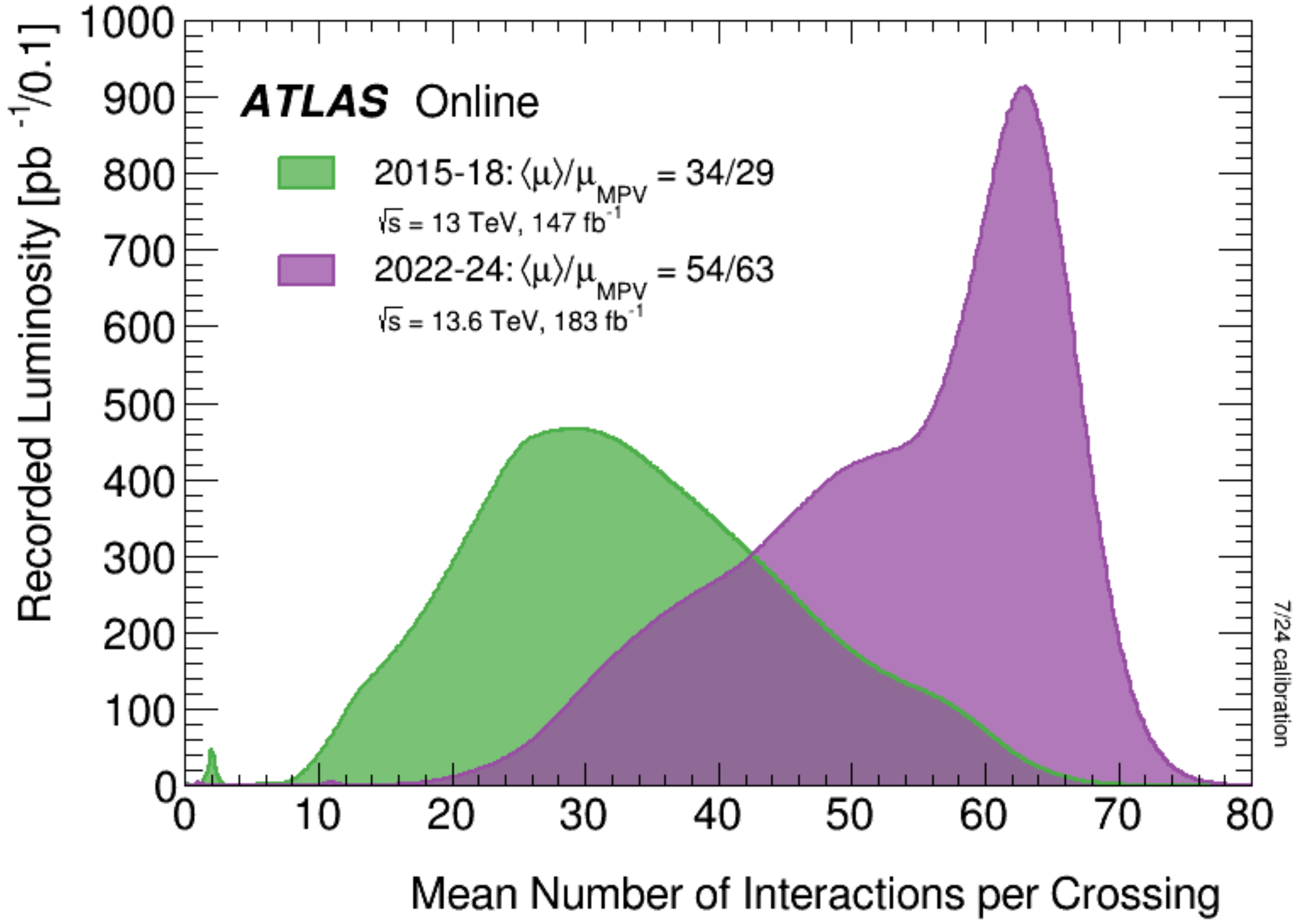
The Standard Model

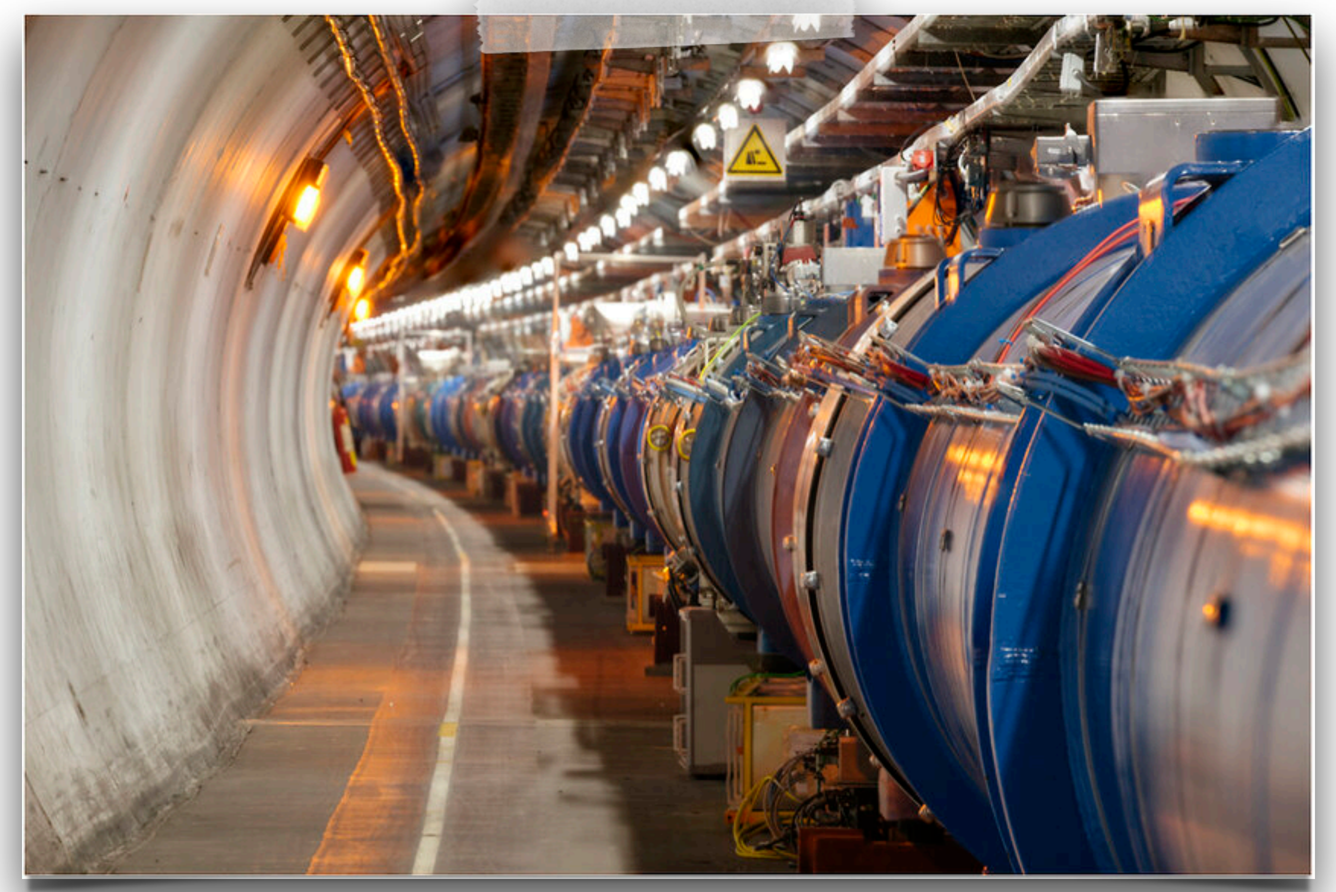


The Standard Model

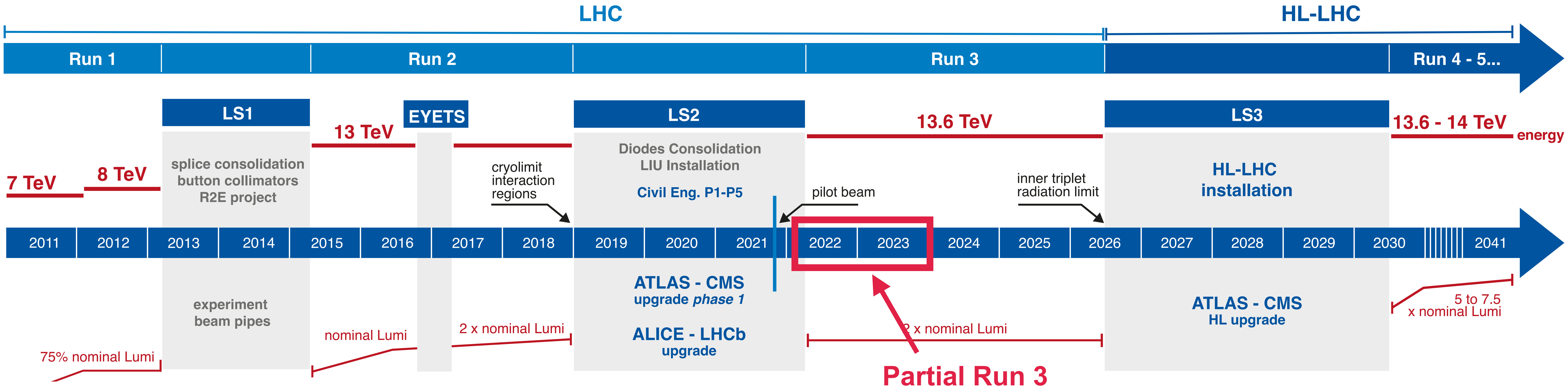
- Physics Beyond the Standard Model (BSM) is crucial to expand the existing theory and explain phenomena that the Standard Model cannot
- The search for BSM Physics can be pursued through experiments like ATLAS at the LHC accelerator complex

The Large Hadron Collider





- The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator
- Proton-proton and heavy ion collisionc, pp at centre-of-mass energies of 7, 8, 13, now 13.6 TeV

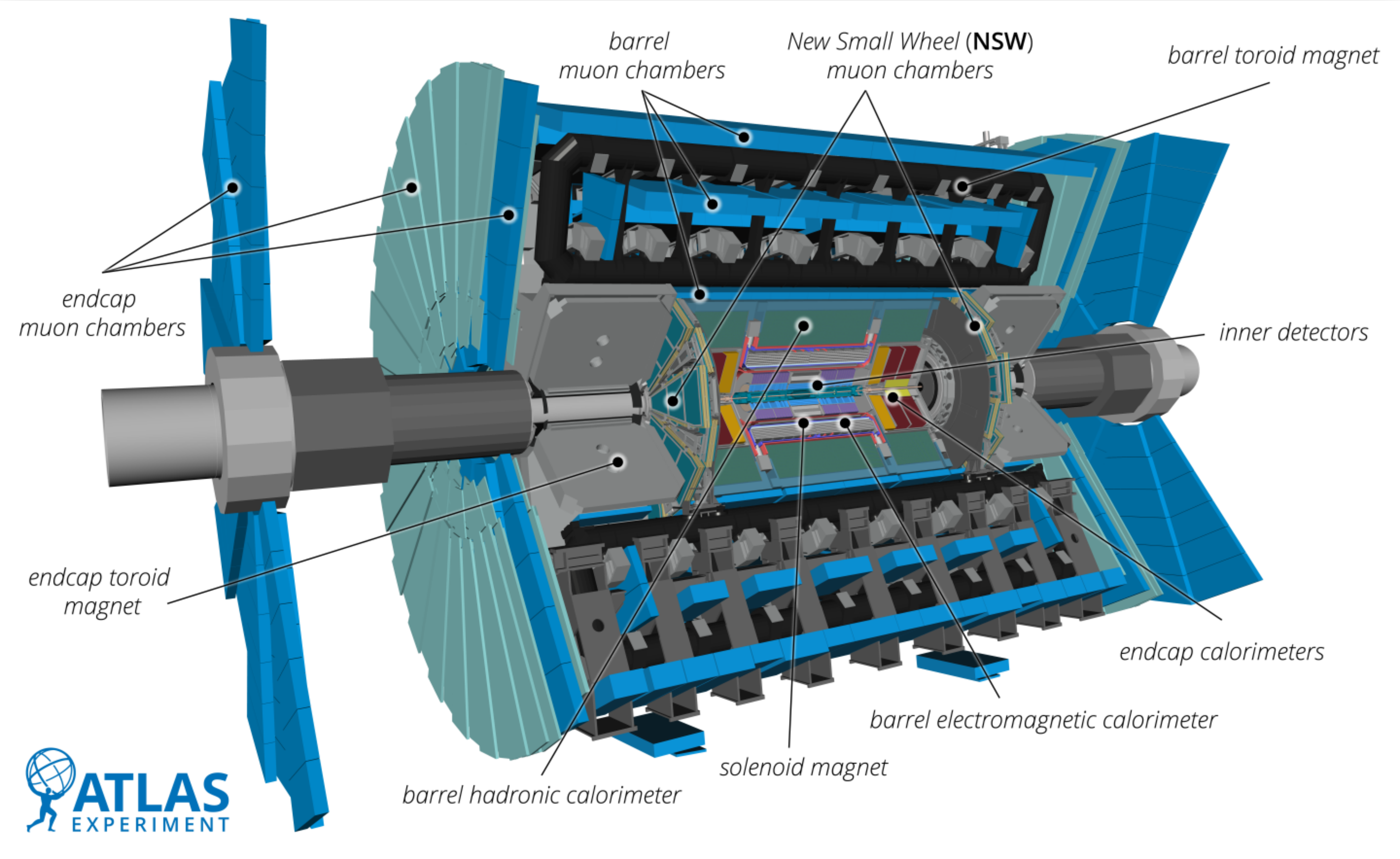


The ATLAS experiment

ATLAS is a multipurpose detector designed to study Higgs boson physics, make precision measurements for deviations from the Standard Model, and search for Beyond Standard Model particles, particularly in the Higgs sector

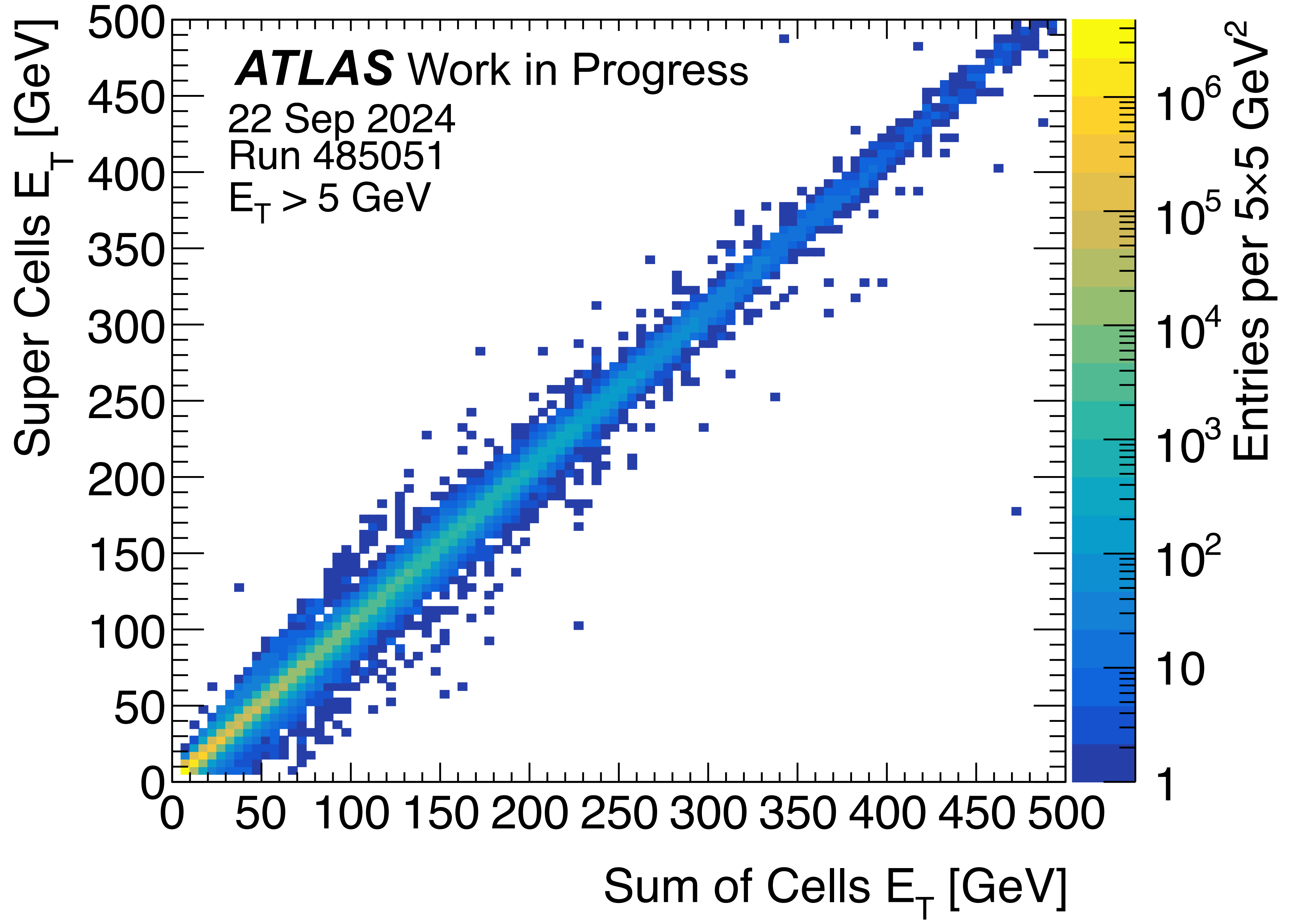
The ATLAS detector is composed by subsystems:

- Tracker
- Electromagnetic and hadronic calorimeters
- Muon spectrometer
- Magnet system
  (Central solenoid and toroid)

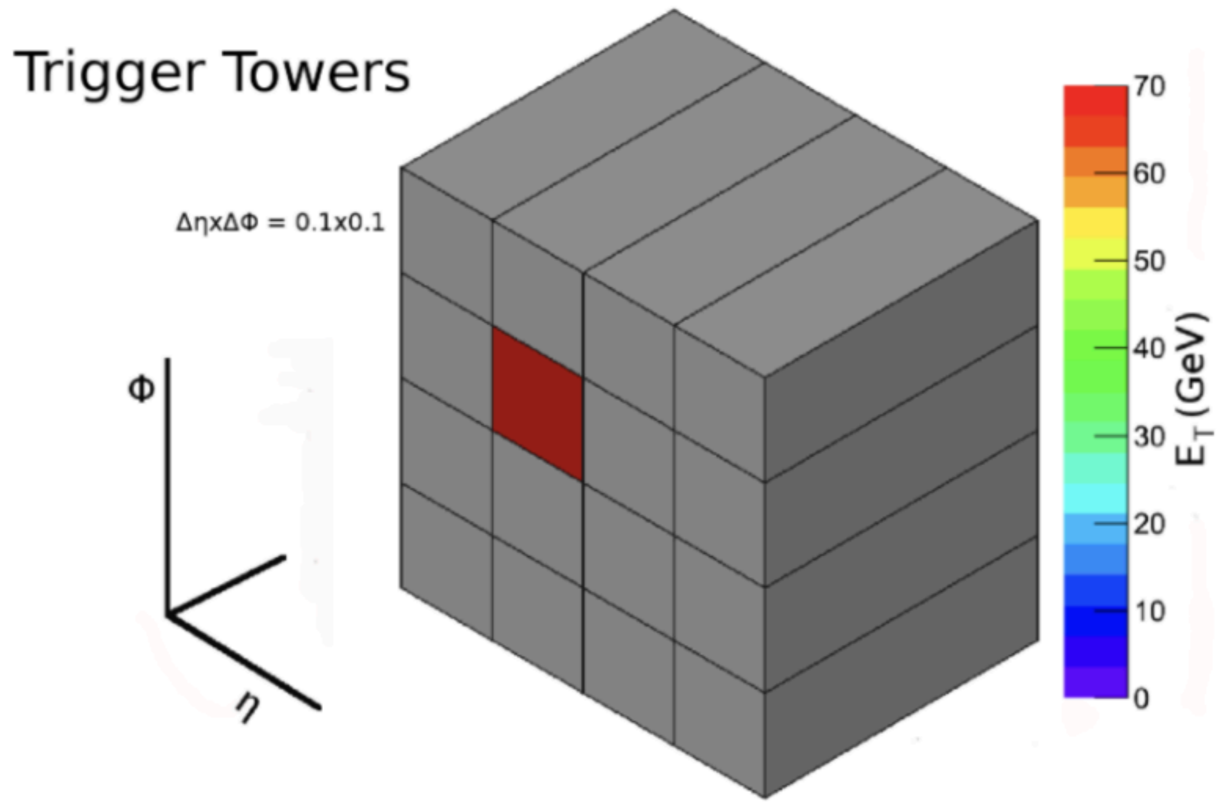


Liquid Argon Calorimeter

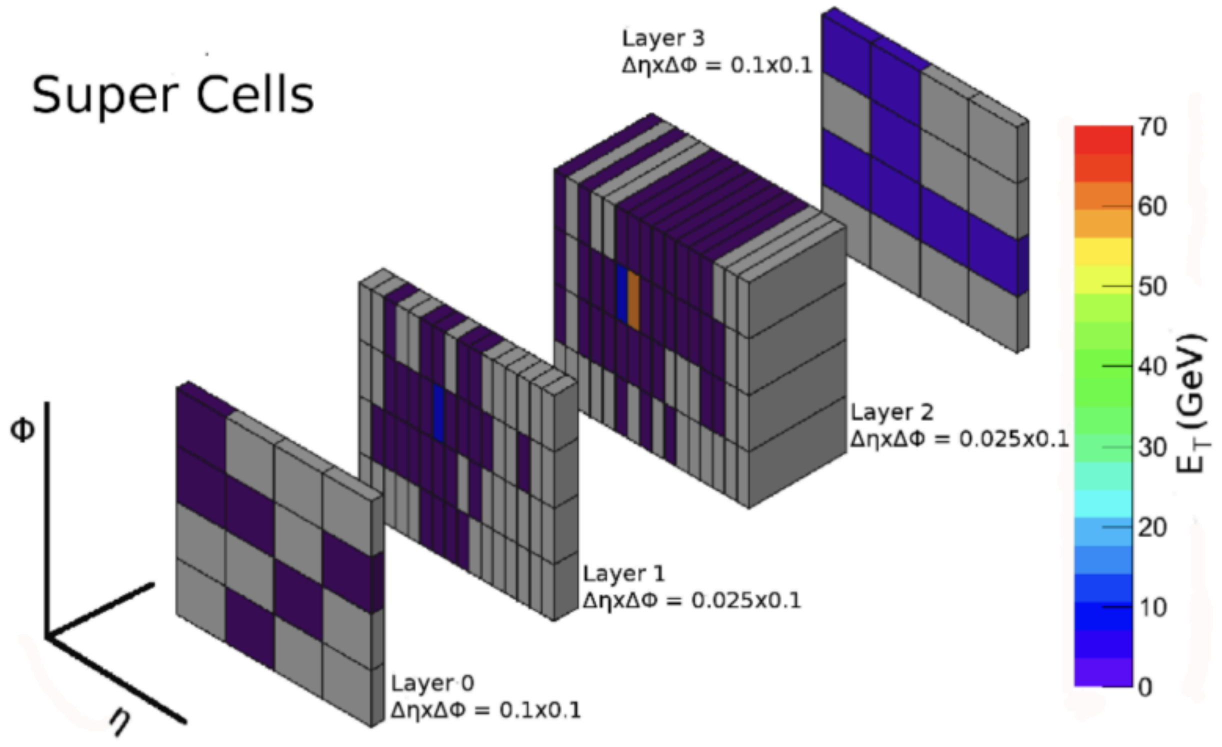
- Since we have photon in final state we need good calorimeter to detect them → Liquid Argon Calorimeter (LAr)
- For more detail see Christian's slides
- For Run 3 we have higher instantaneous luminosity and number of p-p collisions per bunch crossing
- Only a small fraction of events are relevant for physics studies
- Need to improve the background rejection at first level trigger to keep the same trigger rate (100 KHz)



Legacy trigger



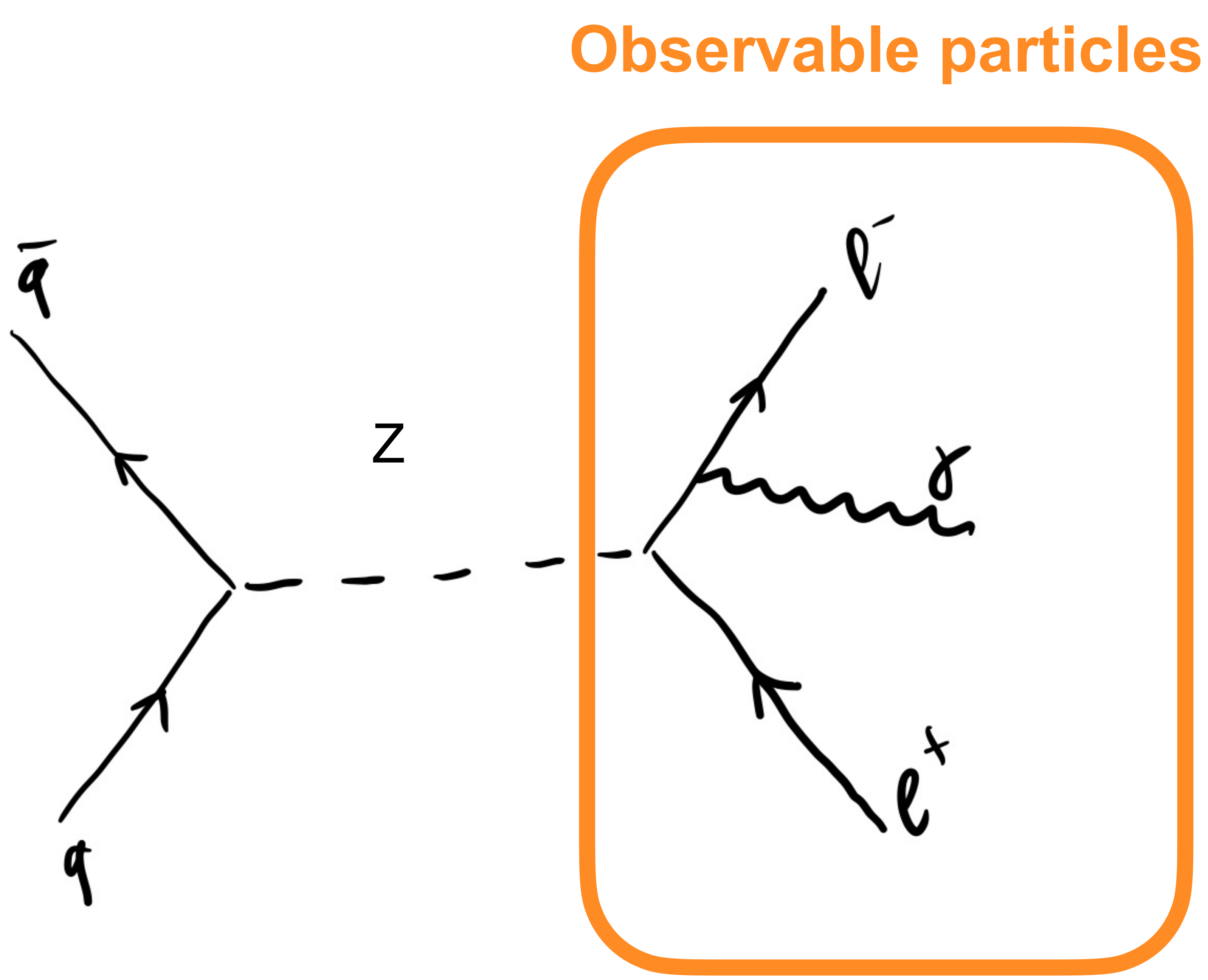
Digital Trigger

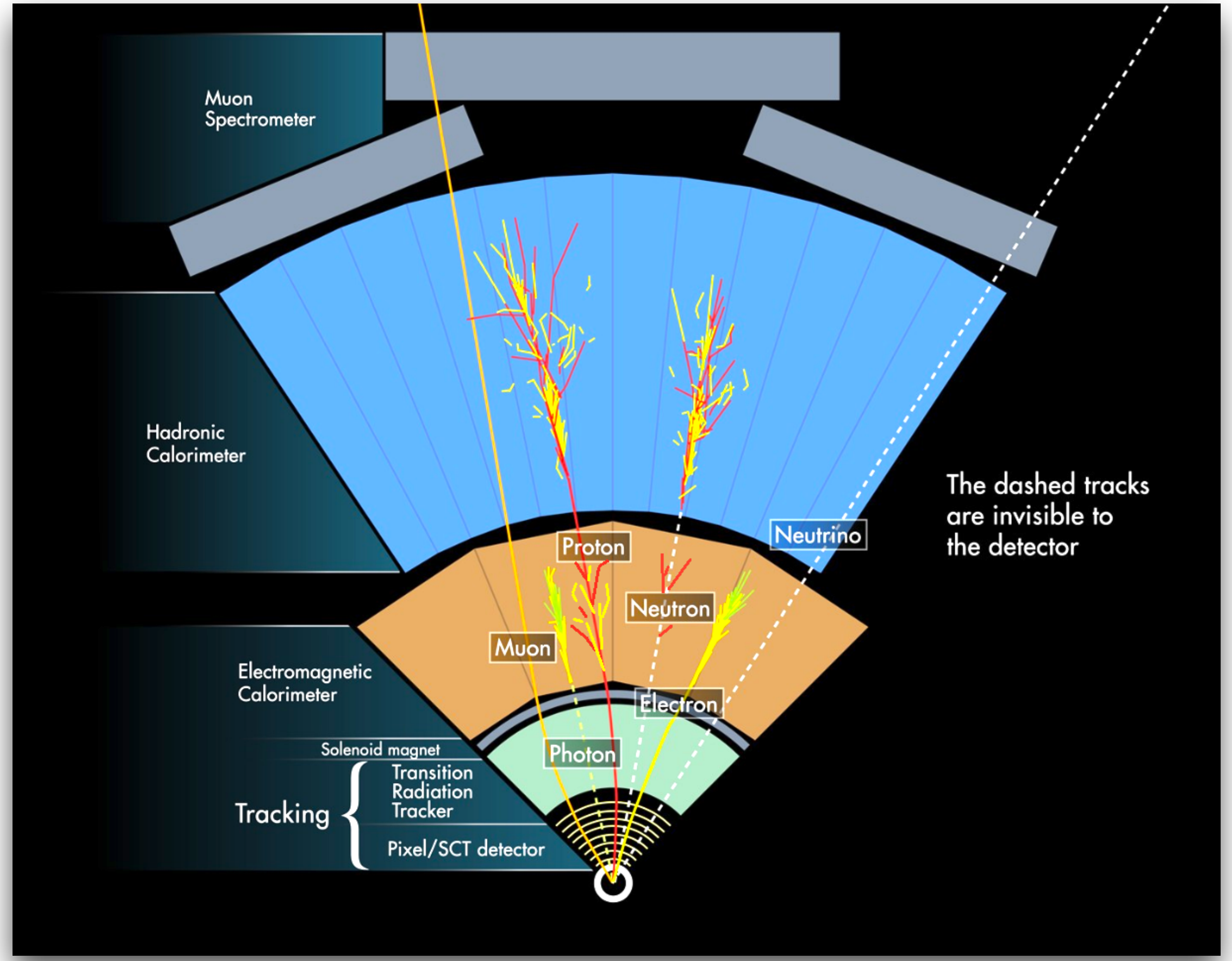


- Digital trigger system offers four-layer information and 10×granularity
- But we would like to make sure that new system works as expected. Need to compare with main readout
- Main readout path: Front-End Boards send cell by cell information to Read Out Drivers

During my qualification period I contributed to development of the needed software tools for LAr Data Quality, data taking and analysis with a focus on problematic super cell channels

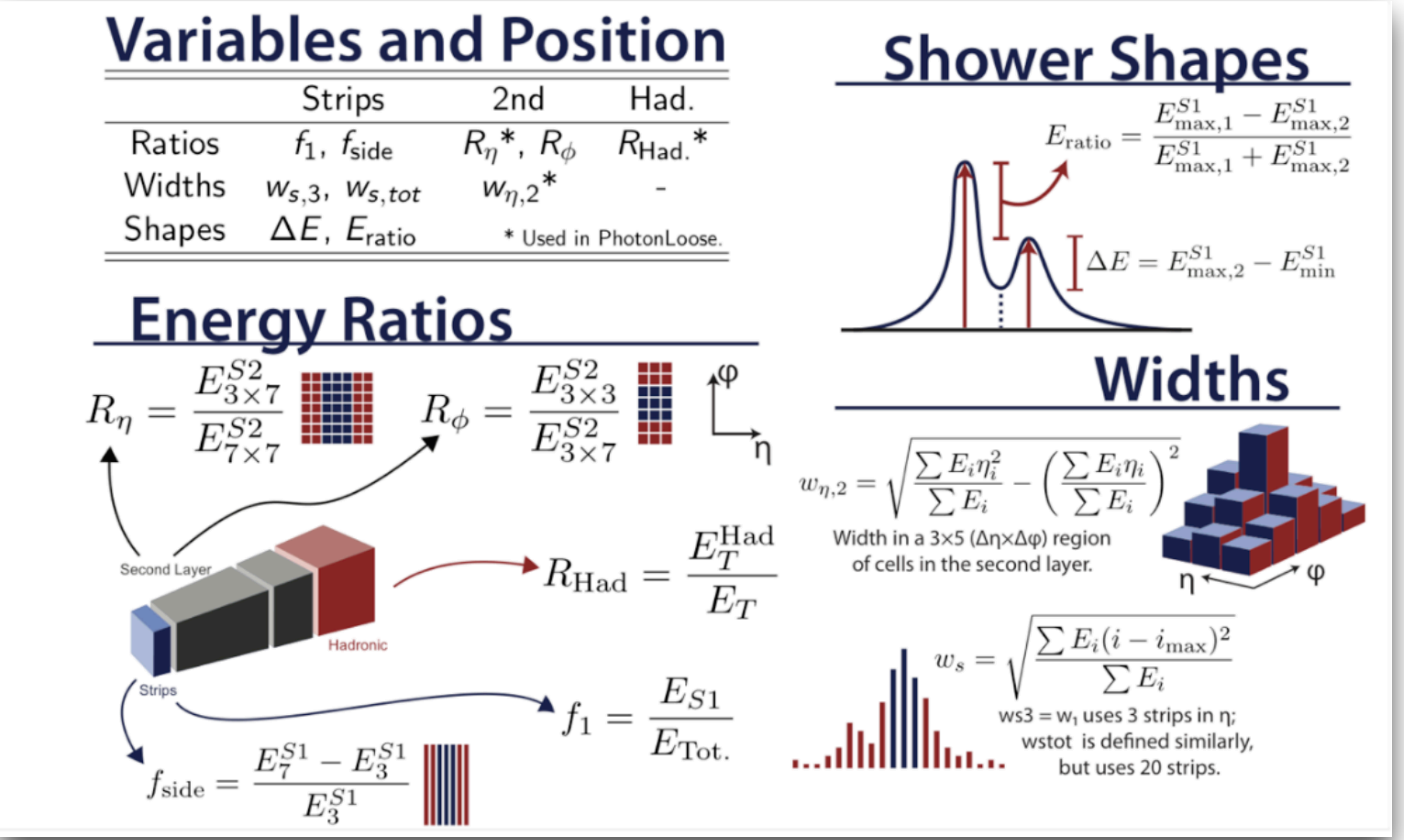
Particle identification





- Energy and momentum of particles are recored as a hits in the tracker and in the muon spectrometer, and energy deposits in calorimeters are reconstructed from these informations.
- Physics objects - photon, electron, muon, hadrons, jets
- Do not detected neutrino, but can calculate $\overrightarrow{E}_T^{miss} = -\sum_{i=e,\mu,..} \vec{p}_T^{\,i}$

Photon Identification



- Photon identification in ATLAS uses cuts on calorimetric variables to effectively separate photons from fake signatures
- Shower shapes - discriminating variables derived from the particle shower shapes in the electromagnetic and hadronic calorimeters

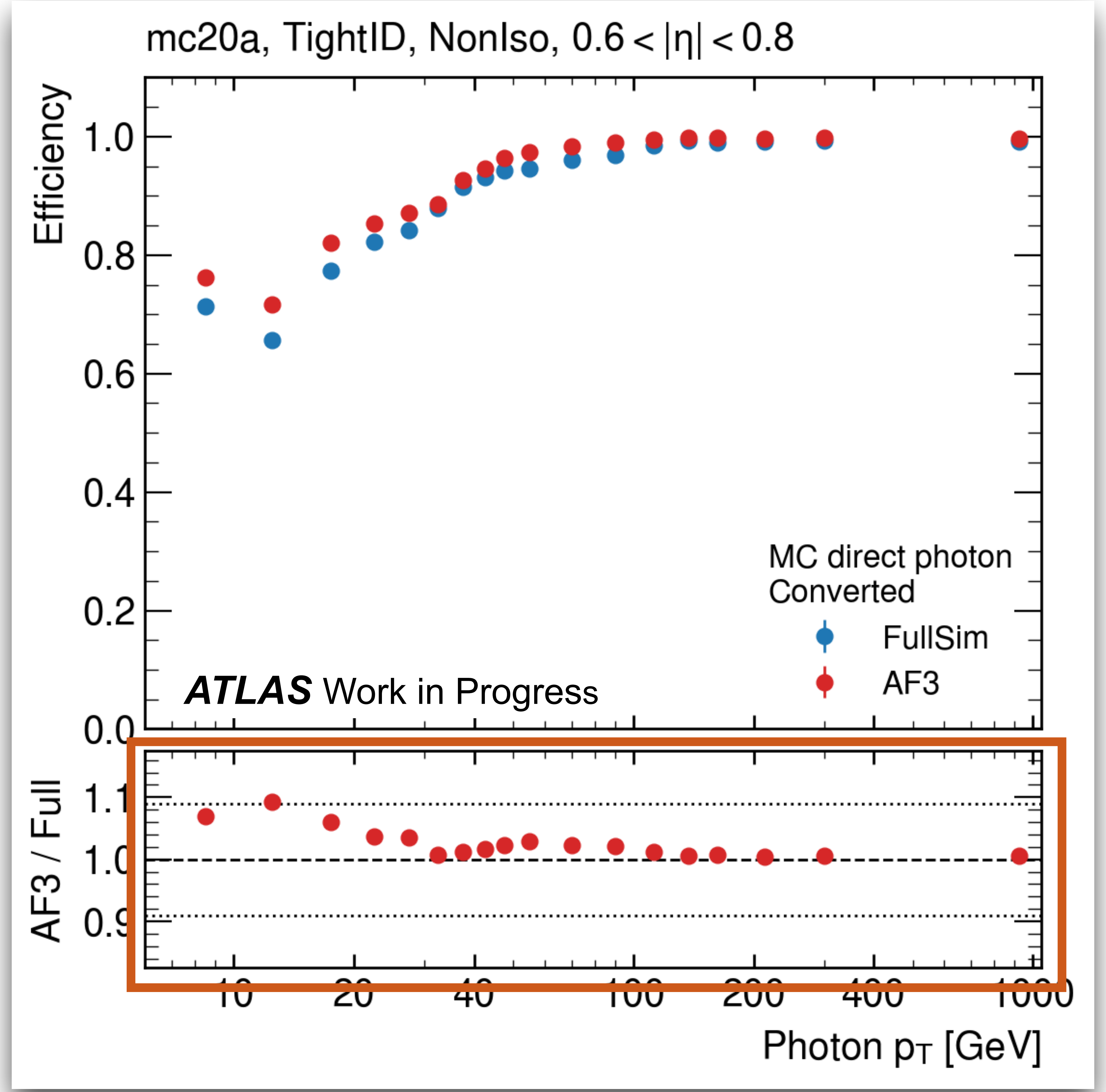
Photon Identification

- In order to have good agreement between data and simulations we should have same efficiency of identification
- Needs to apply Scale Factor (SF) defined as:

$$SF = SF\left(\frac{DATA}{FullSim}\right) * SF\left(\frac{FullSim}{FastSim}\right)$$

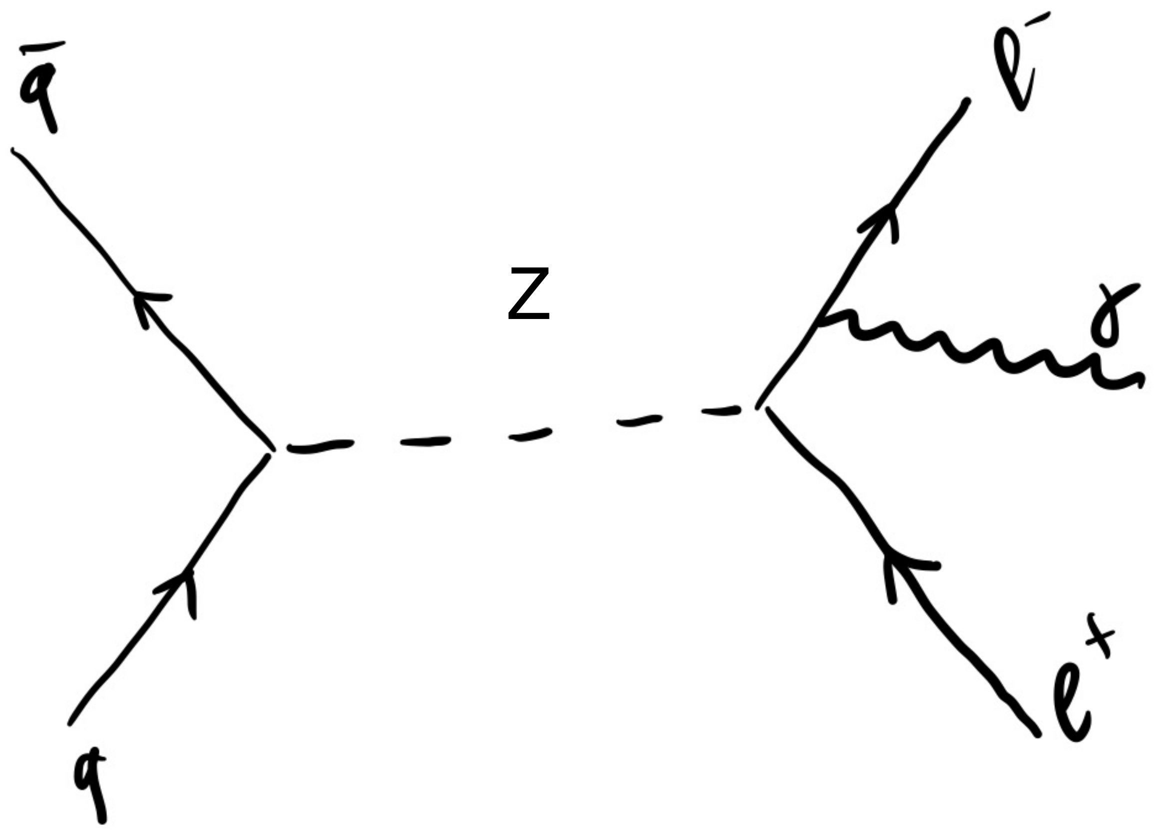
Efficiency of identification

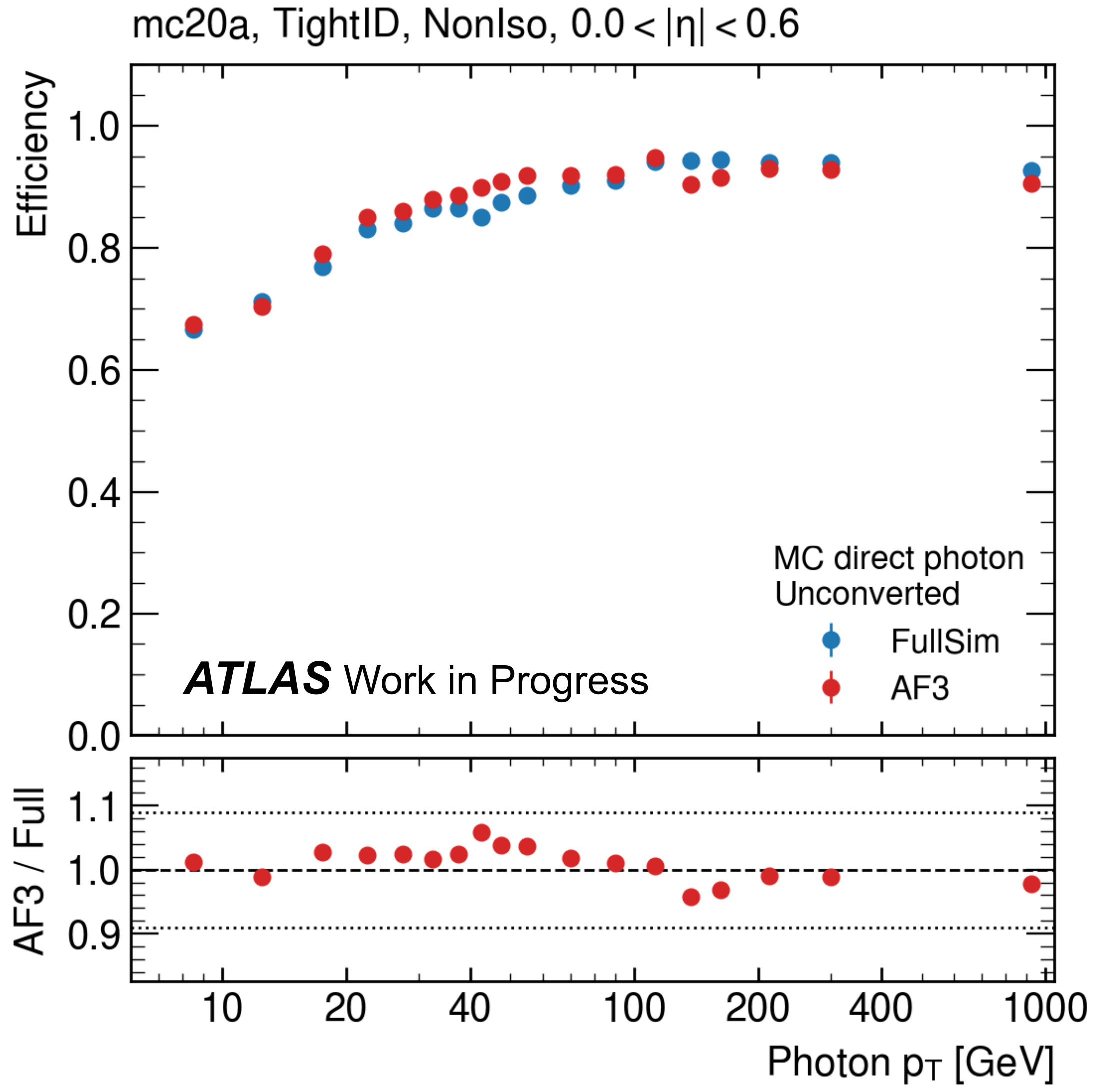
- Full simulation (FS) based on GEANT4
- Meany analyses rely on ATLAS Fast simulation (AF3). Consumes less CPU time
- ATLAS fast simulation uses a parameterized response of the calorimeters



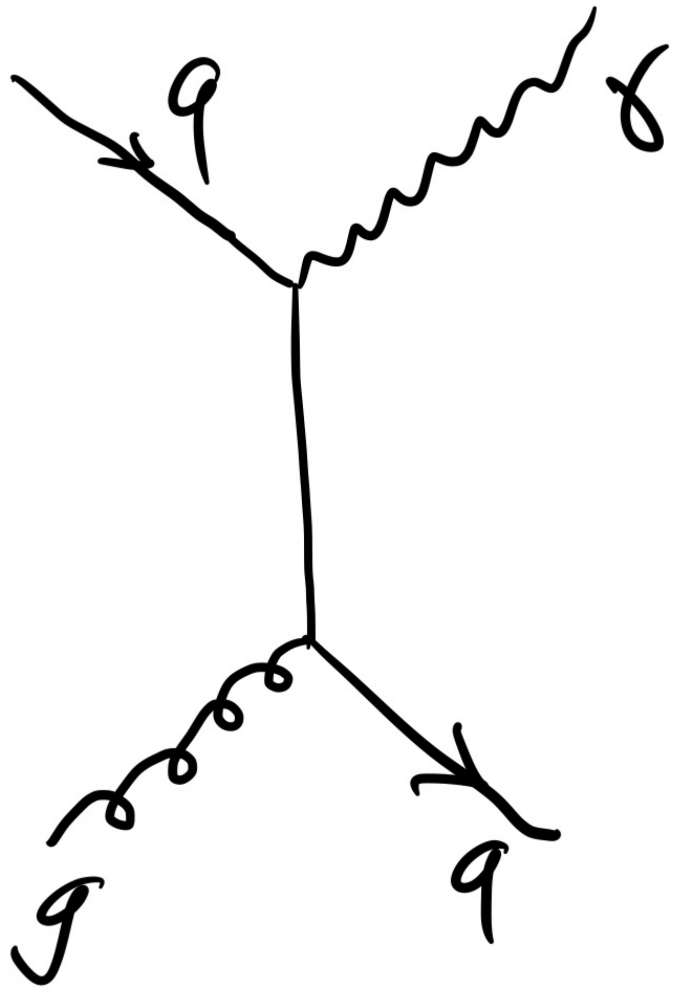
Photon Identification

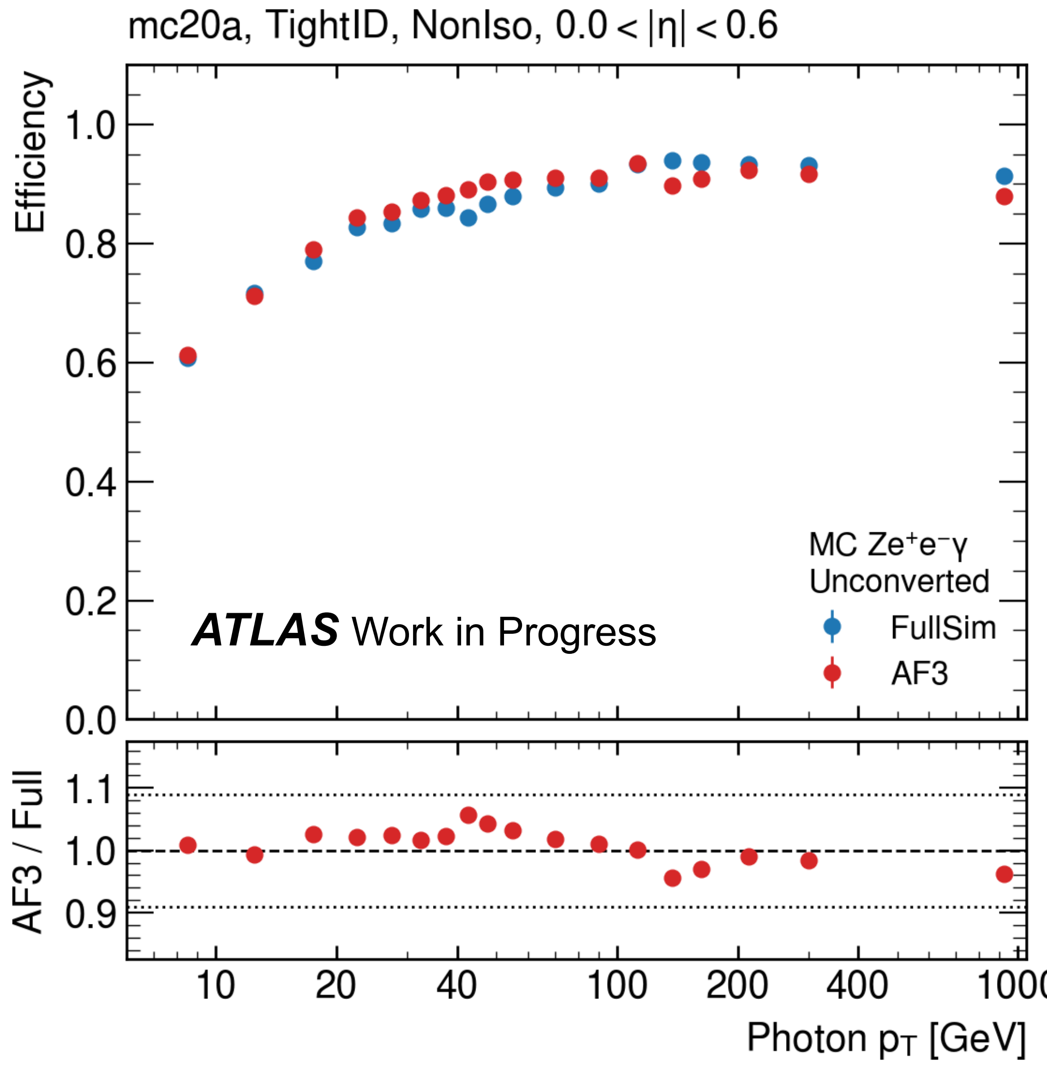
$Zee\gamma$ **and** $Z\mu\mu\gamma$





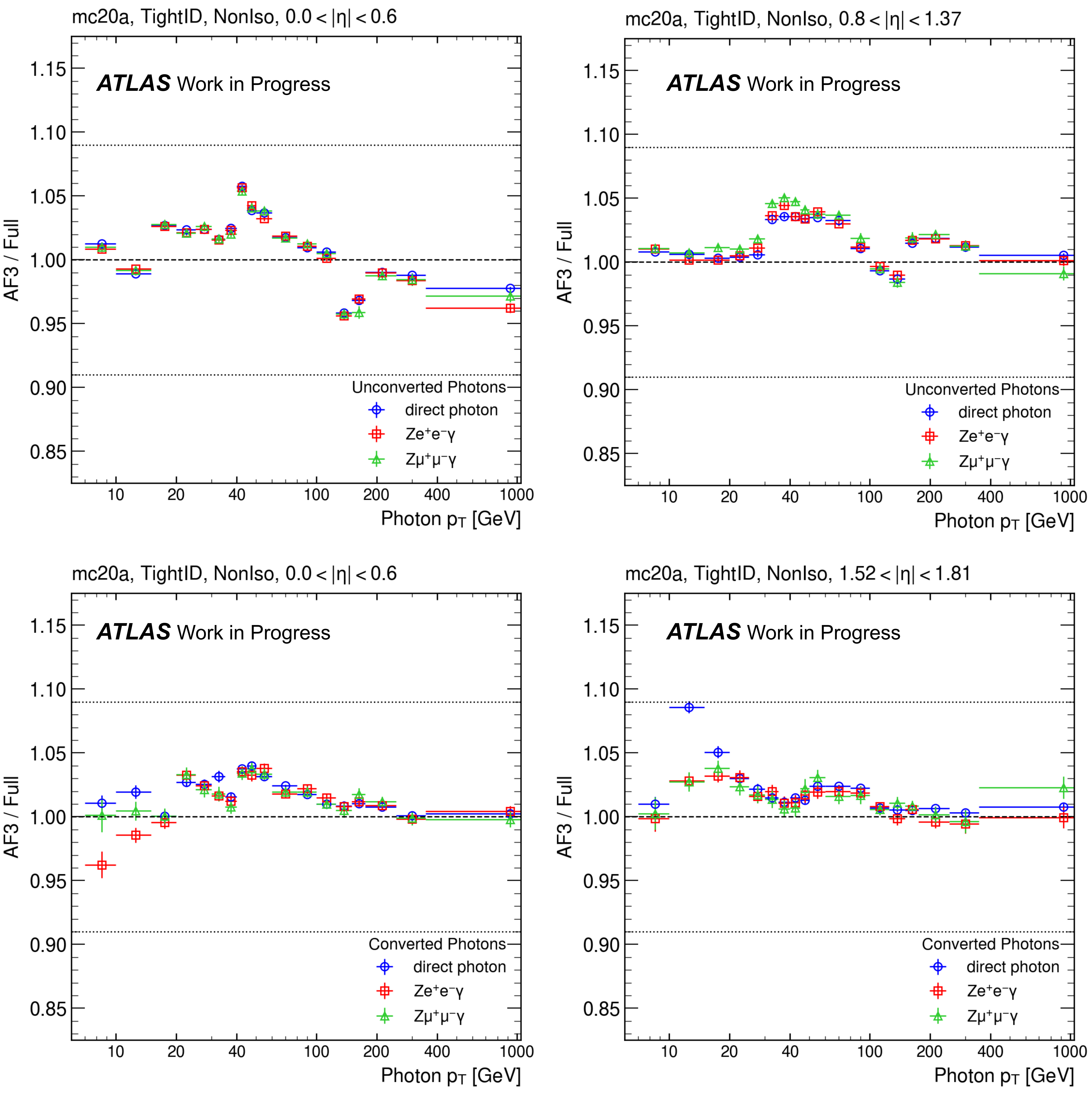
Direct photon production

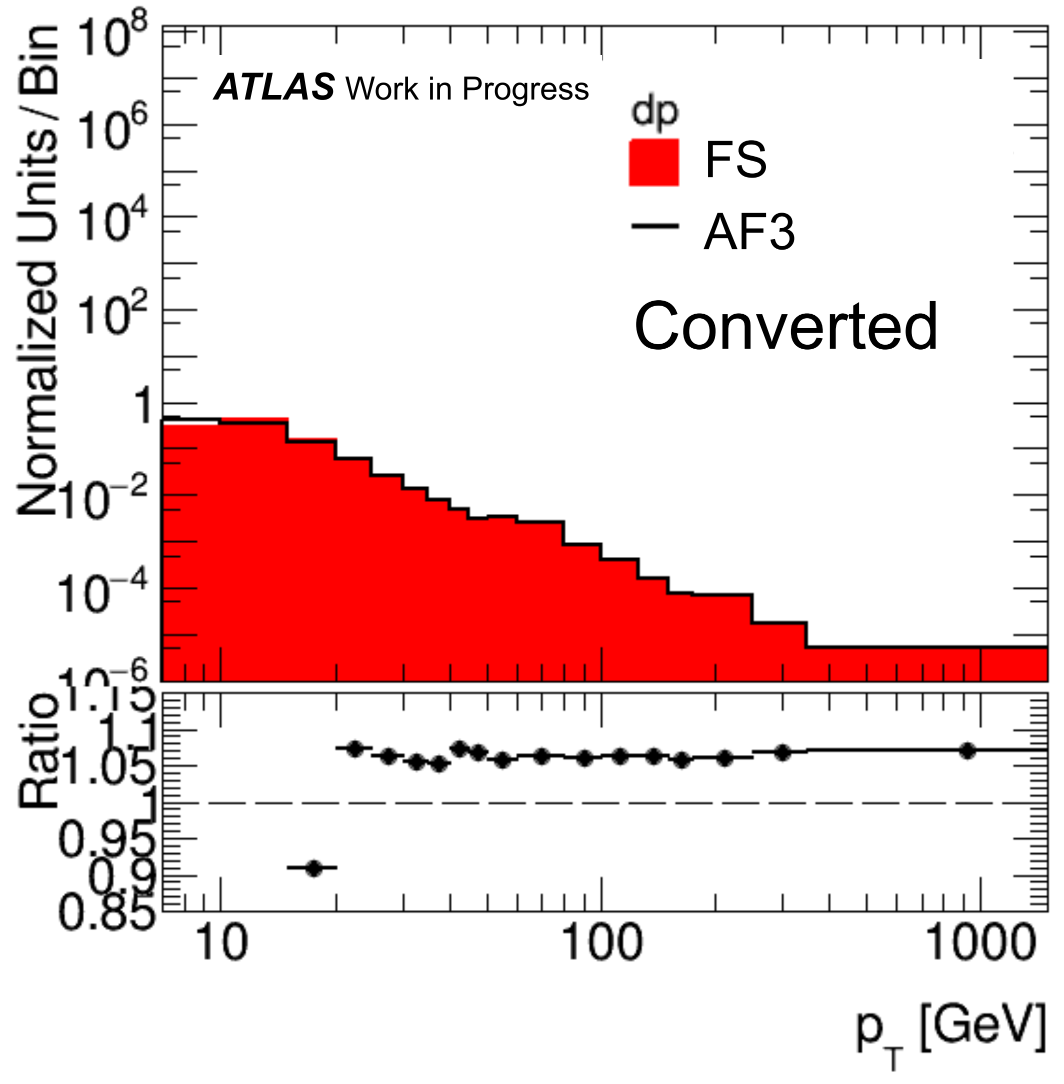


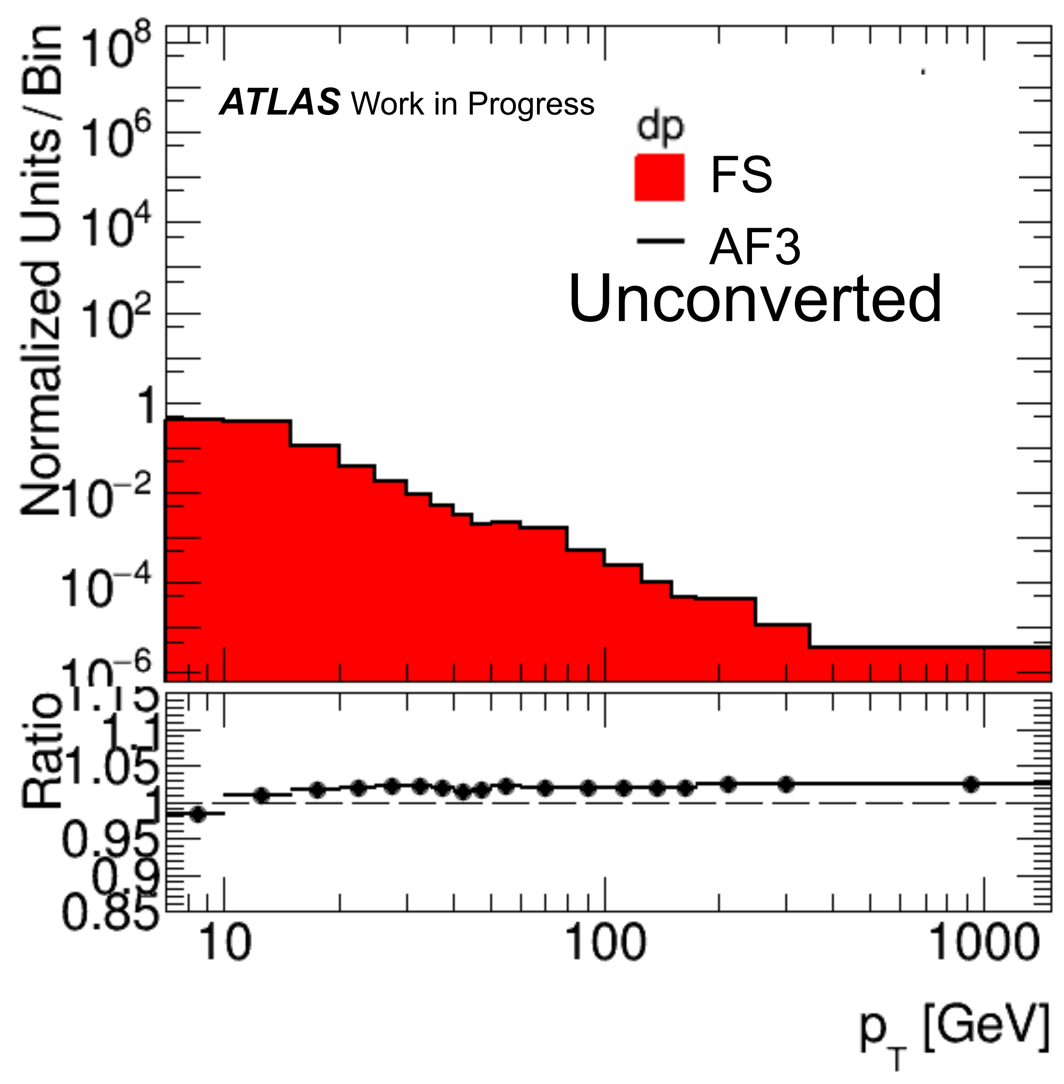


- ID Efficiency are different but Fast sim/Full sim ratio is supposed to be the same for all photons
- Samples: $Zee\gamma$, $Z\mu\mu\gamma$ and Direct photon production (DP)

Efficiency Ratio

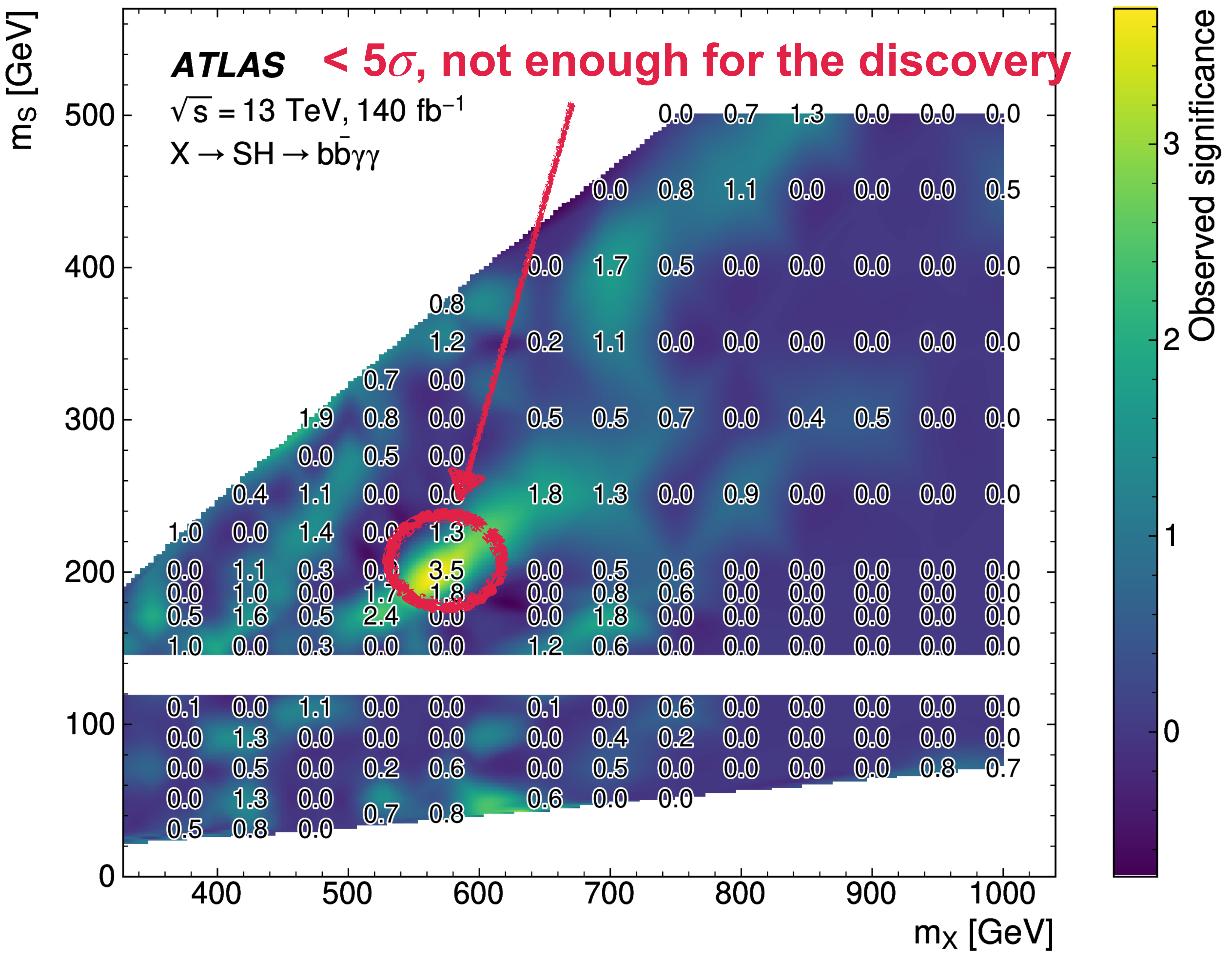




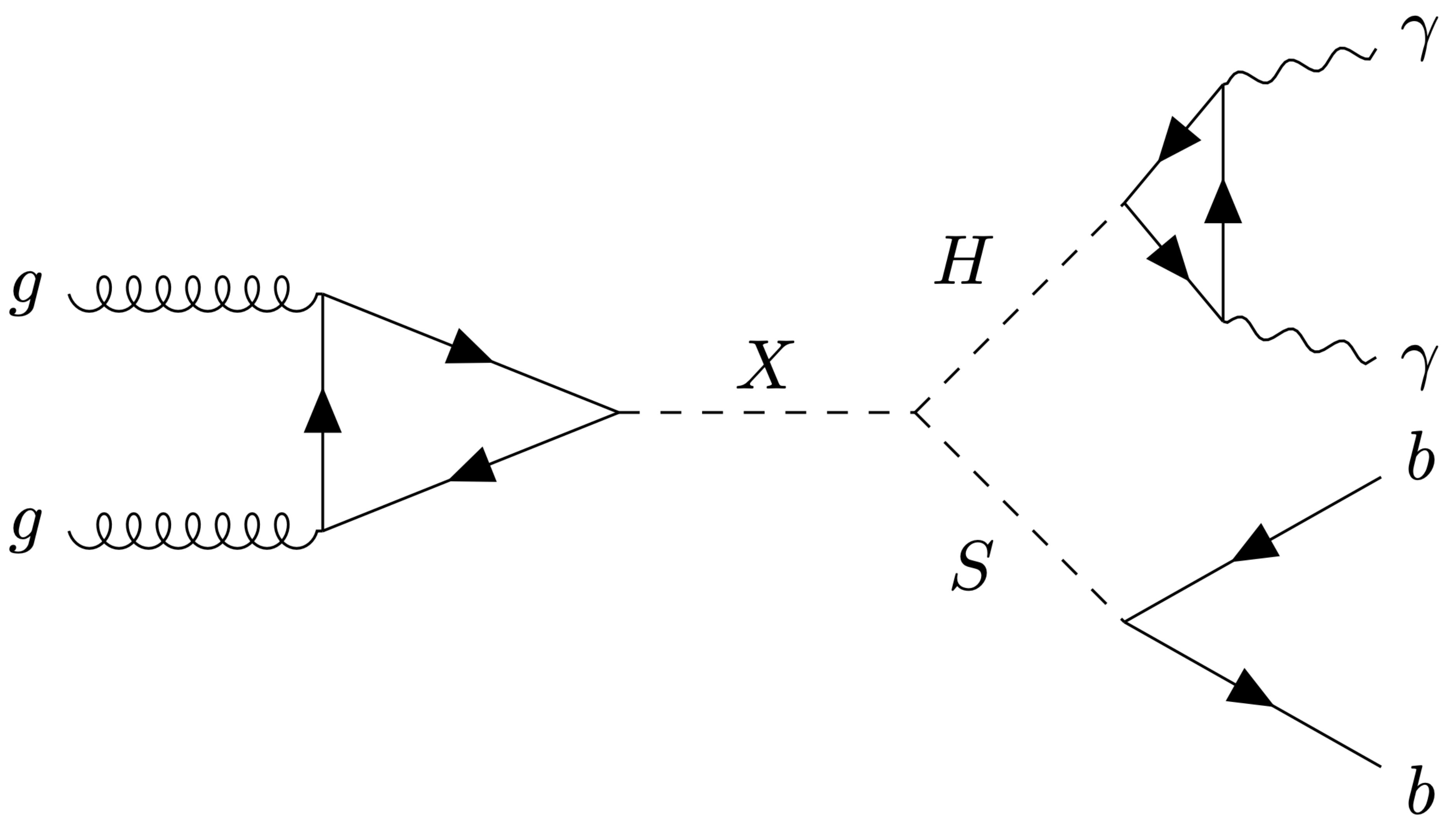


- The photon reconstruction algorithm allows to reconstruct photons even when they convert into e^-e^+ pairs
- Good agreement for unconverted photons from different processes: $Zee\gamma$, $Z\mu\mu\gamma$ and DP
- Discrepancy in low pt between $Zll\gamma$ and Direct photon production
- Discrepancy at reconstruction level for converted DP. Something is weird in AF simulation.
- Agreement above 20 GeV
- SFs were provided for everyone. Can use it physics analysis

Motivation



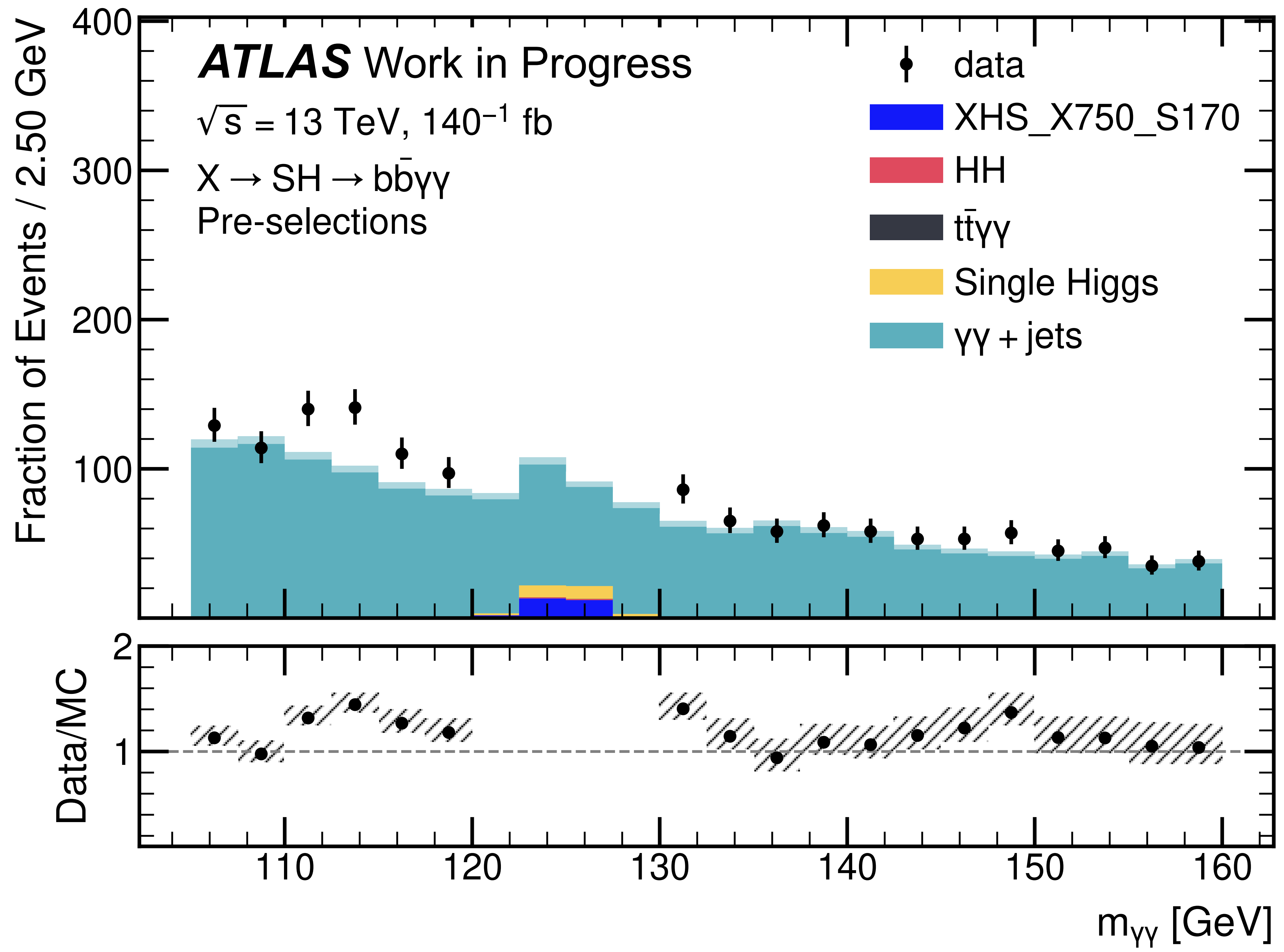
- Many theoretical BSM models predict the existence of new scalar particles in the Higgs sector, denoted as X and S, which could be produced in proton-proton collisions in association with a Higgs boson: $pp \to X \to SH$
- The same final state for $HH \to bb\gamma\gamma$ (Arthur's slides)
- Mass order $m_S < m_H < m_X$ or $m_H < m_S < m_X$



- Searches for such particles have been conducted by several analysis teams in the ATLAS and CMS experiments
- Analysis with Run 2 data reported a local (global) excess of 3.5 (2.0) for $m_X, m_S = (575, 200)$ GeV compared to the background only hypothesis in the decay channel
- Follow-up on the excess seen in the Run2 data using Run2 + partial Run3 dataset

Event Selection

Invariant photon mass distribution



Selection

	2 b-tagged	1 b-tagged
Number of 'tight' and isolated photons	≥ 2	
$m_{\gamma\gamma}$ [GeV]	$\in [105, 160]$	
Number of leptons	$== 0$	
Number of central jets	$\in [2, 5]$	
Number of b-tagged jets @ 77% WP	$== 2$	$== 1$

Yields for Run 2

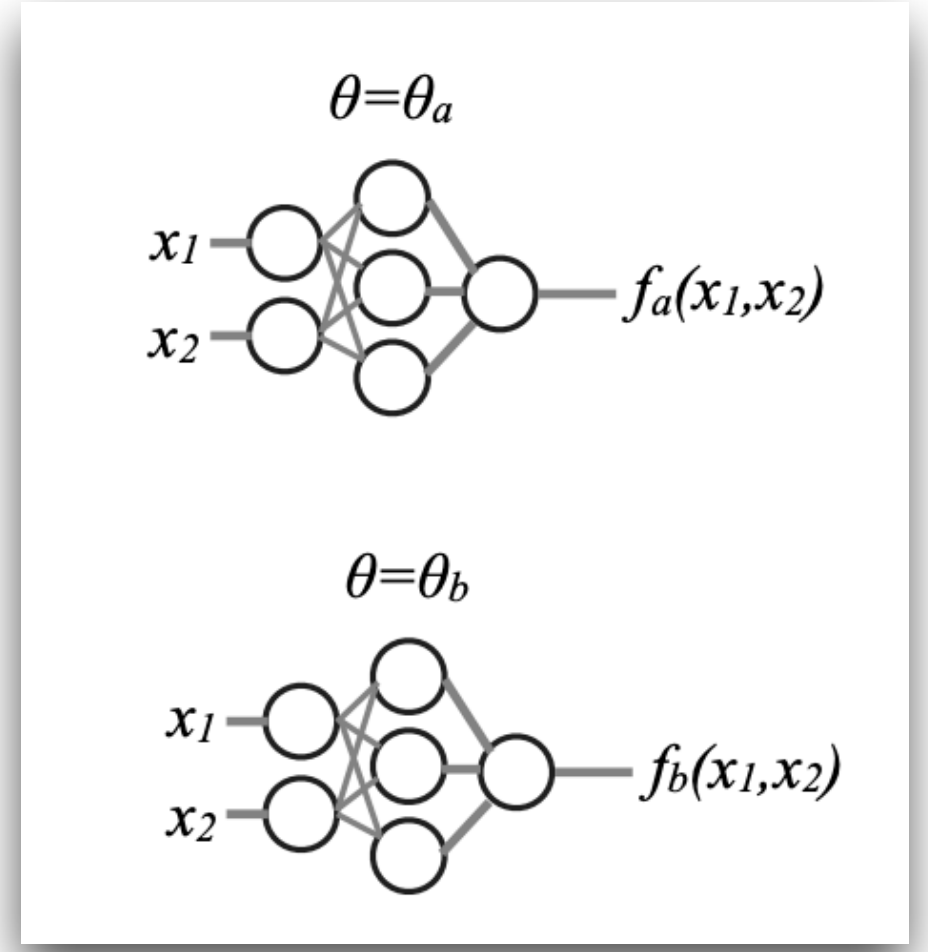
	CR (WP 77%)	SR (WP 77%)
ggH	0.101 ± 0	5.409 ± 0.019
$VBFH$	0.006 ± 0	0.332 ± 0.0004
$qq \rightarrow ZH$	0.029 ± 0	1.548 ± 0.001
$gg \rightarrow ZH$	0.012 ± 0	0.7127 ± 0.0009
W^-H	0.001 ± 0	0.036 ± 0.00016
W^+H	0.001 ± 0	0.053 ± 0.00015
bbH	0.015 ± 0	0.543 ± 0.0003
$tHjb$	0.016 ± 0	0.823 ± 0.0005
tWH	0.001 ± 0	$0.075 \pm 5.94e-05$
ttH	0.183 ± 0	7.438 ± 0.0053
$ggHH$	0.069 ± 0	1.666 ± 0.0004
VBFHH	0.003 ± 0	$0.060 \pm 9.6e-06$
$t\bar{t}\gamma\gamma$ (noallhad)	0.374 ± 0	$0.0925 \pm 6.8e-05$
$\gamma\gamma$+jets	622.493 ± 0	157.254 ± 0.402
Total SM	623.552 ± 0	181.586 ± 0.4953
Data	1123 ± 0	-
NF$_{\gamma\gamma\text{+jets}}$	1.80 ± 0.05	-
$m_{X,S} = (260, 30)$	0.246 ± 0	5.544 ± 0.128
$m_{X,S} = (275, 145)$	0.688 ± 0	12.340 ± 0.200
$m_{X,S} = (300, 70)$	0.356 ± 0	8.366 ± 0.156
$m_{X,S} = (575, 200)$	0.889 ± 0	21.984 ± 0.193
$m_{X,S} = (650, 90)$	0.833 ± 0	23.052 ± 0.225
$m_{X,S} = (700, 570)$	1.094 ± 0	21.068 ± 0.186

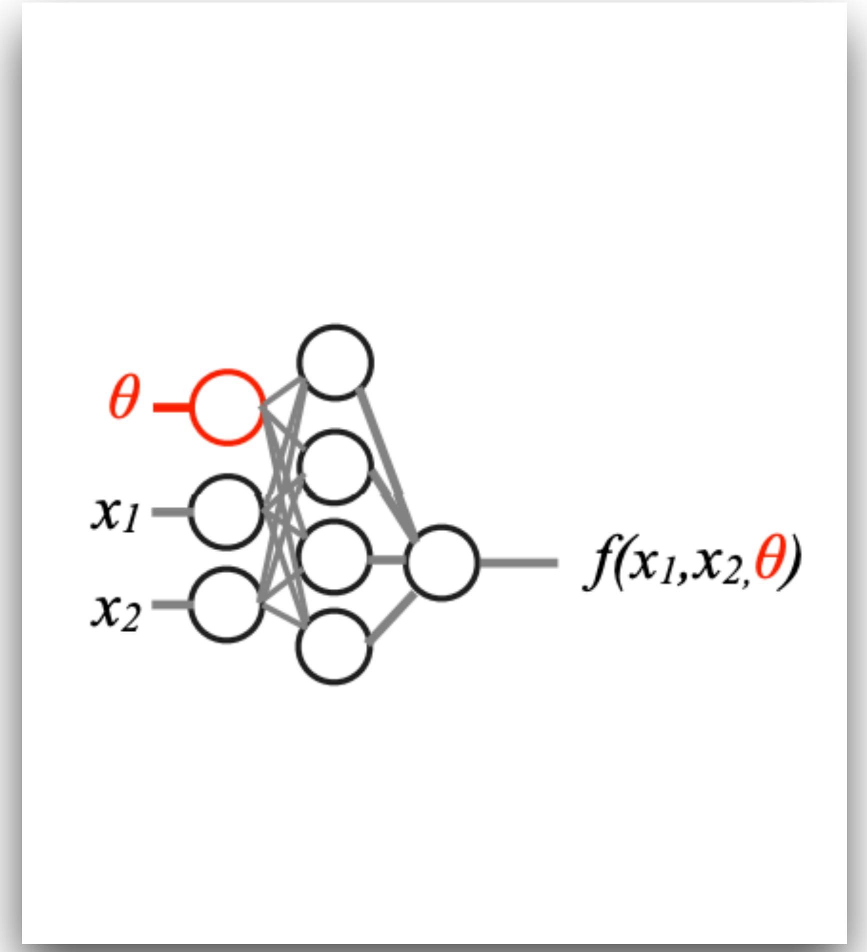
- Main background $\gamma\gamma$ + jets (non resonant)

Parametrised neural network

Individual networks with parameters θ_a

Single network trained with input features and parameter



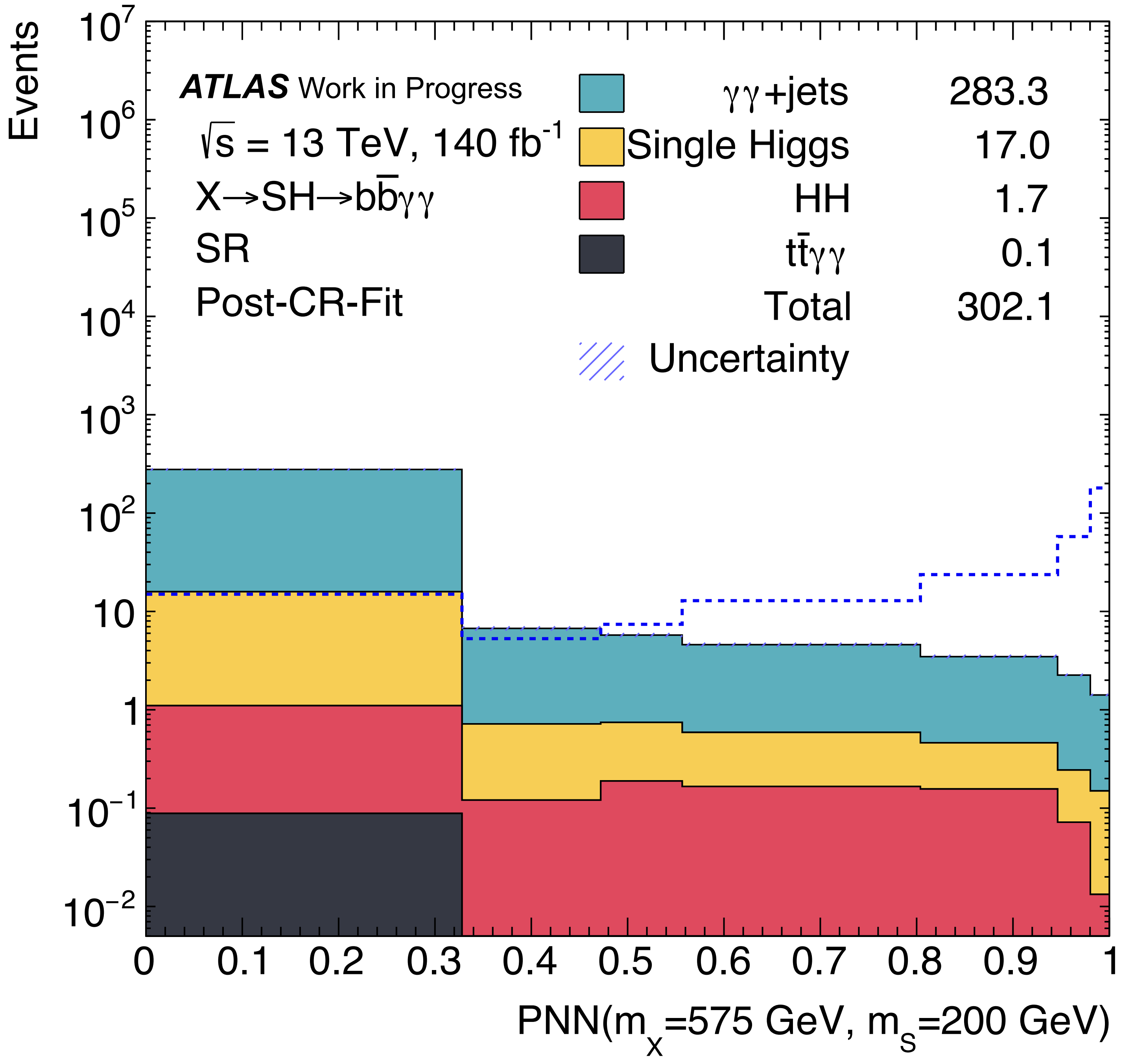


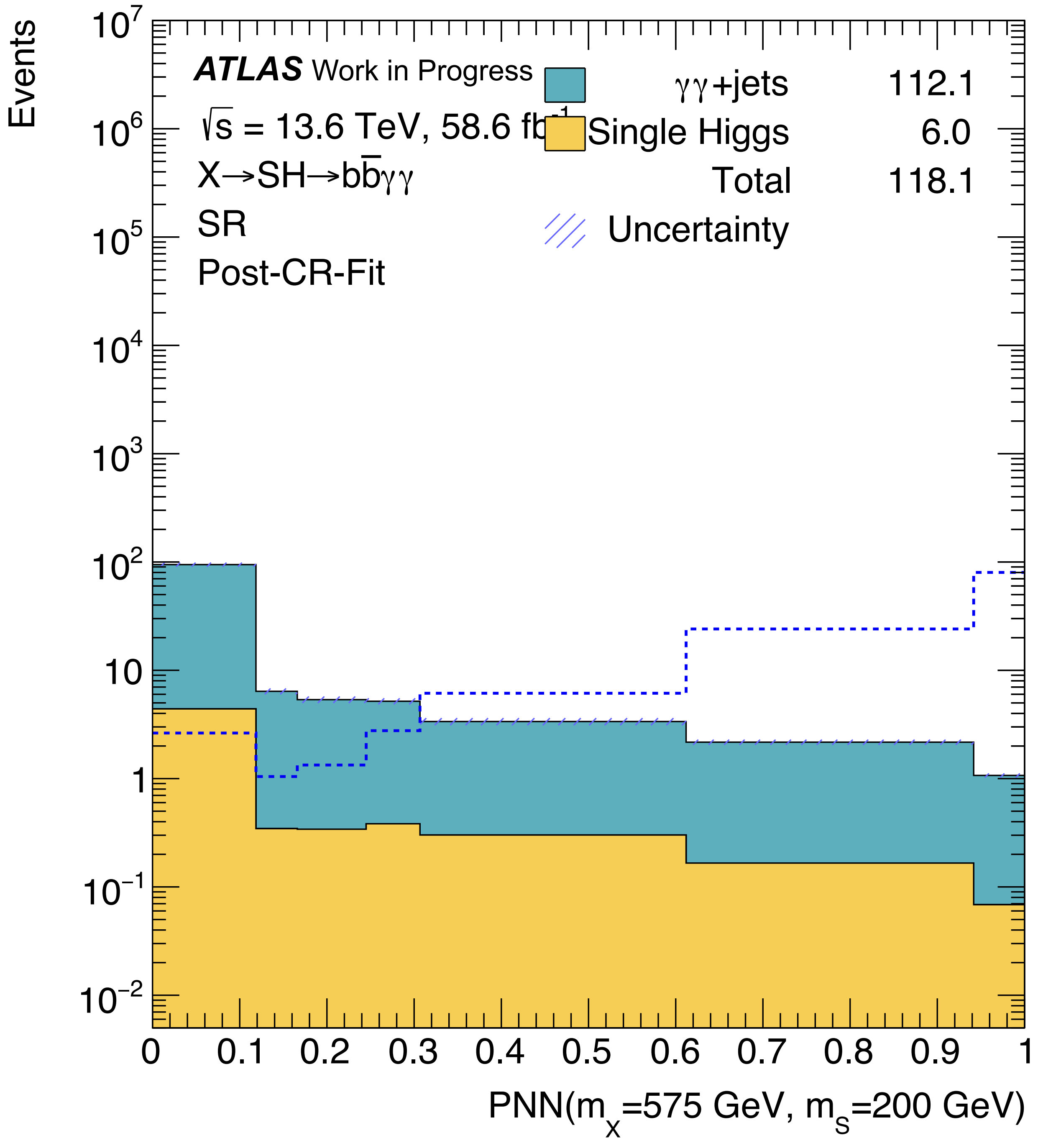
x_1 and x_2 - input variables

Training Setup:

- Train two different PNN for Run 2 and Run 3
- Input variables: m_{bb}, $m^*_{bb\gamma\gamma} = m_{bb\gamma\gamma} - (m_{bb} - 125$ GeV$)$, $\theta = (m_X, m_S)$
- Final PNN score is used in binned likelihood to estimate limits and significance

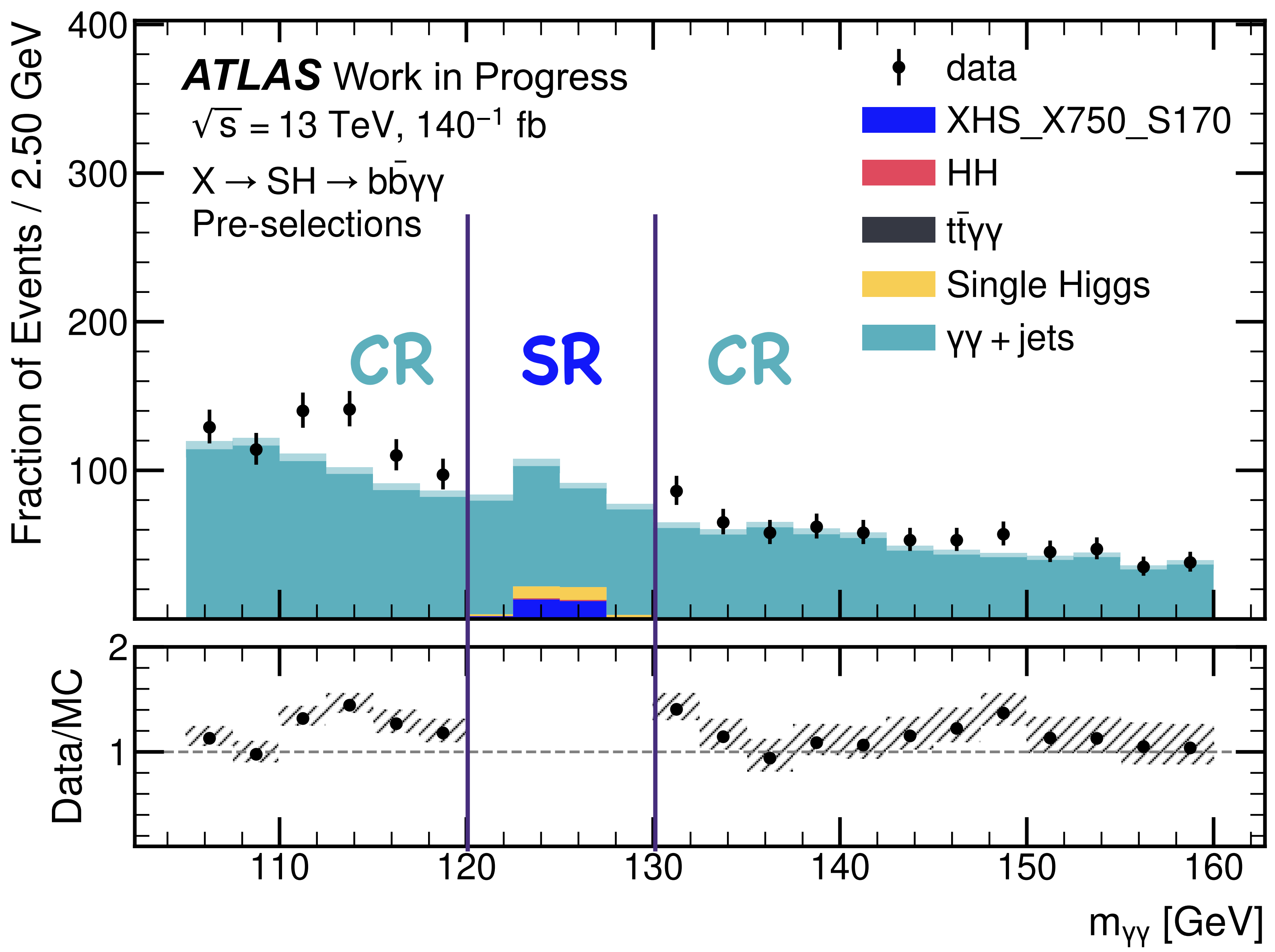
- Parametrised Neural Network (PNN) is an appropriate approach for this search which targets a large phase space of signals that differ from each other by 2 parameter m_X and m_S
- PNNs are trained with a set of training variables (**x**) that will be used to discriminate signal from background and a set of parameters (θ). Alternatively, a standard NN is only trained on **x**.

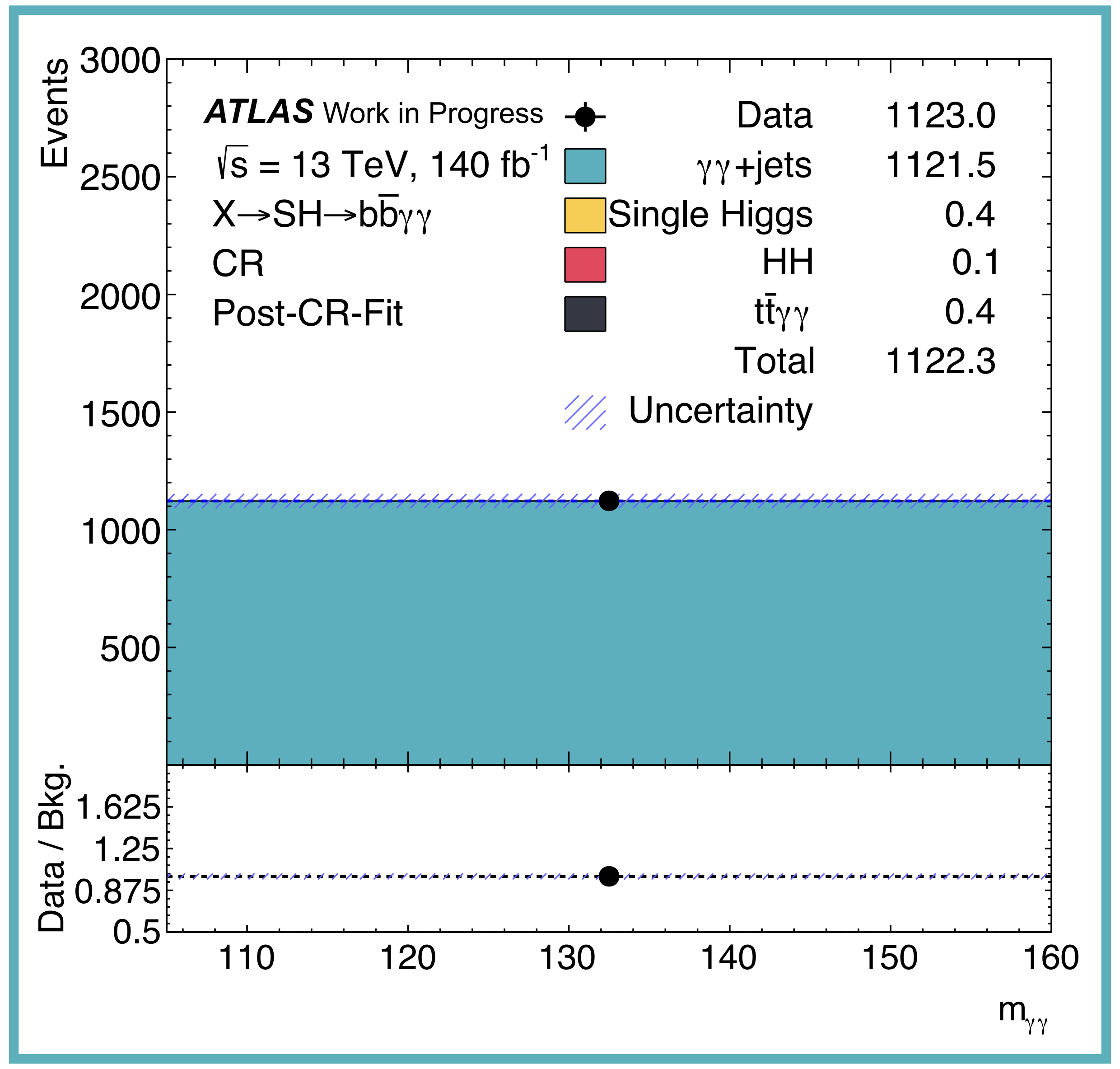




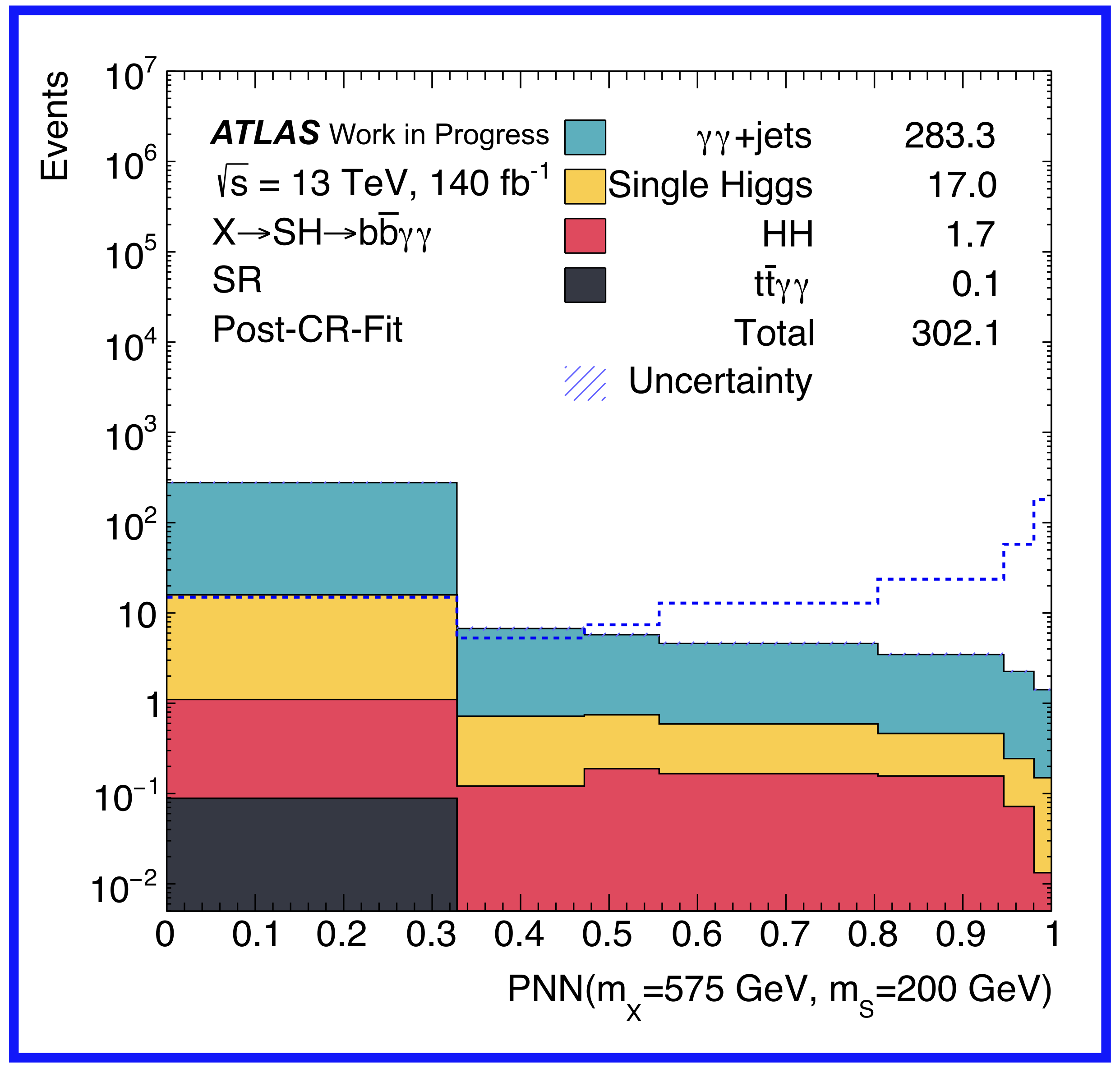
Analysis strategy

Invariant photon mass distribution





PNN distribution

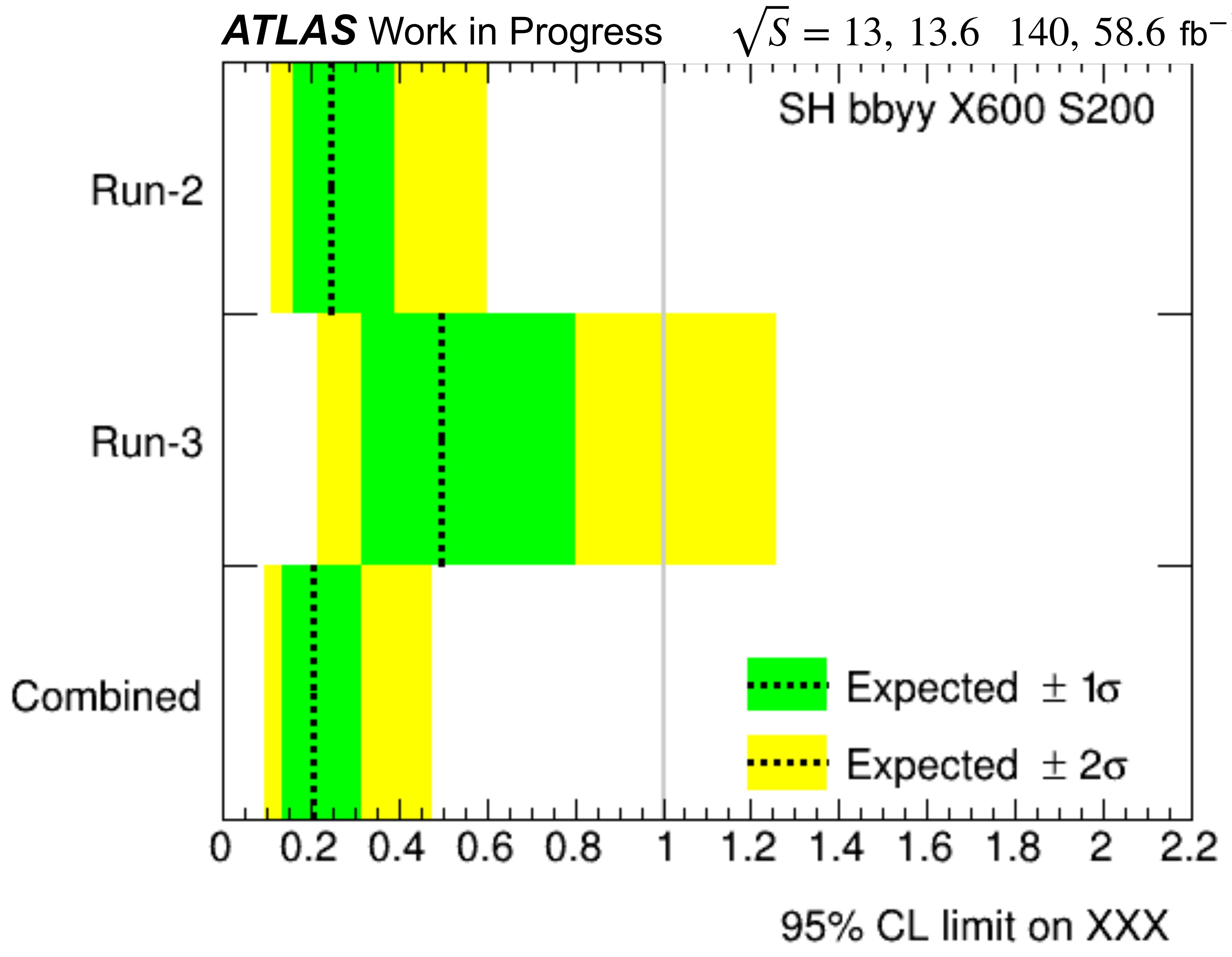


- Main background $\gamma\gamma$ + jets (non resonant)
- Signal Region (SR) with mass window $m_{\gamma\gamma} \in [120,130]$
- Control Region (CR) with $m_{\gamma\gamma} < 120$ and $m_{\gamma\gamma} > 130$ is used to correct the normalisation of $\gamma\gamma$ + jets
- Fit is blinded, we use simulation instead of data in SR
- Fitting the PNN distribution in the SR
- Fits are made without systematic uncertainties at the moment

Limit Run 2 + Run 3 Combinations for $m_X = 600, m_S = 200$

s

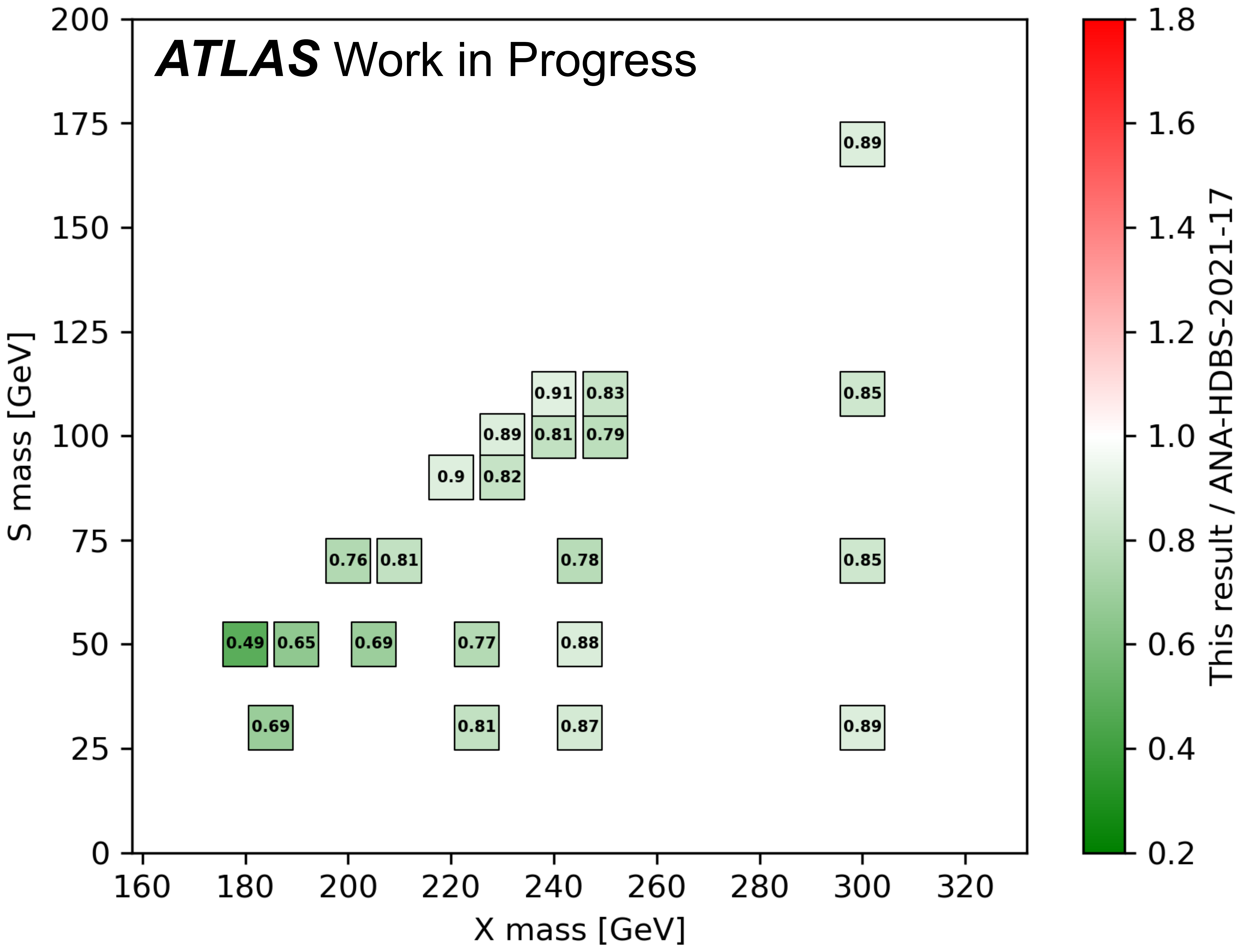
	-2sigma	-1sigma	Median	1sigma	2sigma
Run 2	0.111 fb	0.160 fb	0.247 fb	0.391 fb	0.602 fb
Run 3	0.244 fb	0.357 fb	0.560 fb	0.905 fb	1.419 fb
Run 2 + 3	0.094 fb	0.134 fb	0.202 fb	0.314 fb	0.476 fb
published Run 2	-	-	0.240 fb	-	-



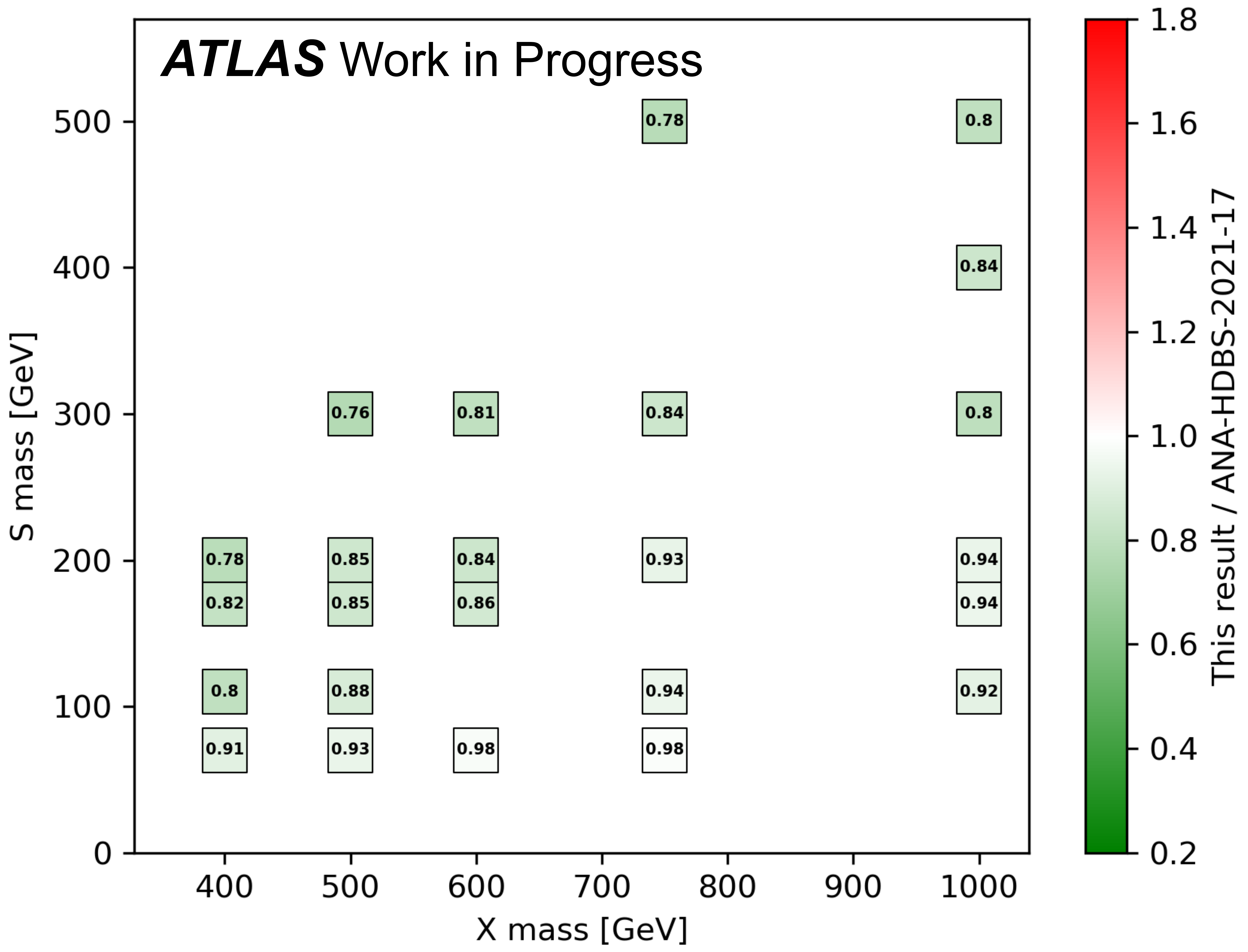
- Difference of around 3% in median limit value for run 2 vs published run 2
  - Consistent with difference in yields b/w previous analysis and new analysis
- Adding partial Run 3 to Run 2 resulted in an improvement of around 20%

Ratio of Run 2 + Run 3 expected to published limits

Low mass regime



High mass regime



- Ratio of the 95% CL upper limits on the signal strength from the current Run 2 plus partial Run 3 analysis to those from the previous Run 2 result
- Adding partitial Run 3 to Run 2 resulted in an improvement 2 - 50 %

Summary

- **Qualification Task in LAr Data Quality:**
  - Doing LADIeS(LAr Data Investigation and Signoff) shifts
  - Developing the tools used for DQ assessment. Contribute to the study of the DT performance
- **Photon identification:**
  - Retrieving photon MC-to-MC SFs between Full and Fast Simulation
- **Physics Analysis:**
  - Moving PNN from legacy software to new one. PNN Validation.
  - Data/MC comparison. Combined fit results

Backup