CENTRE DE PHYSIQUE DES PARTICULES DE MARSEILLE
CPPM

Aix*Marseille université

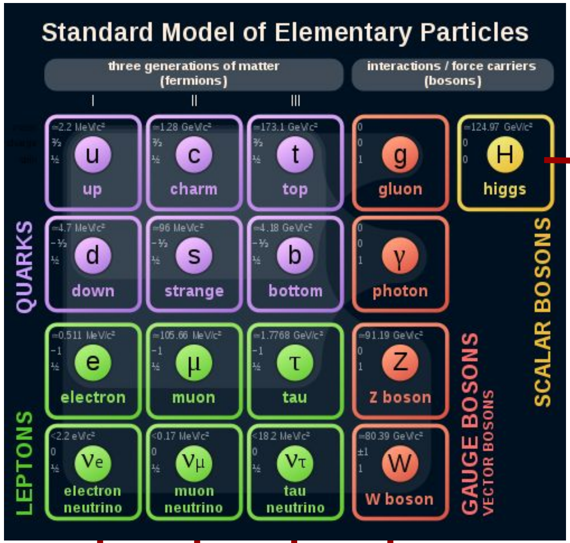
cnrs

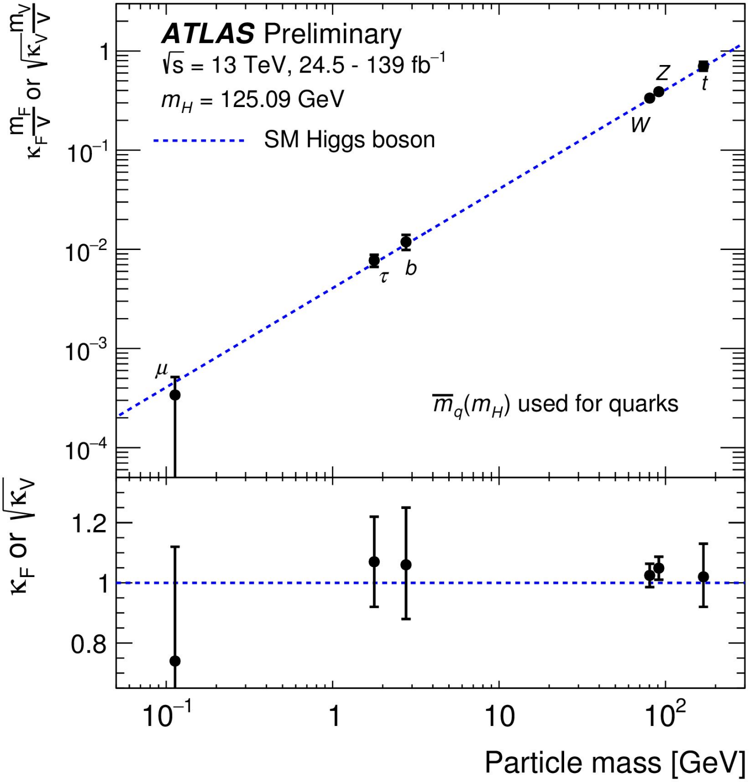
The Standard Model is complete, what’s next

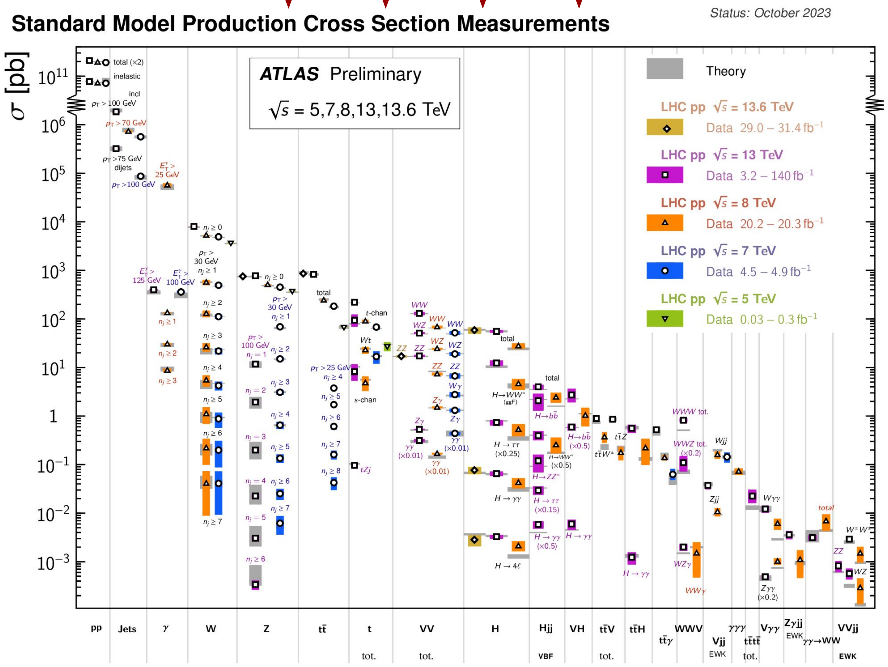
HDR defense

Georges Aad
15/03/2024

The Standard Model

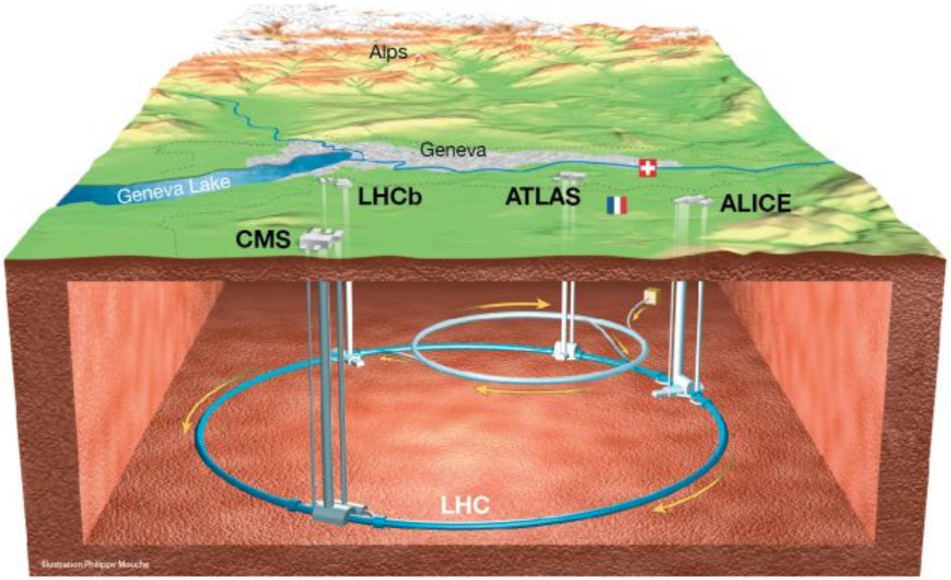


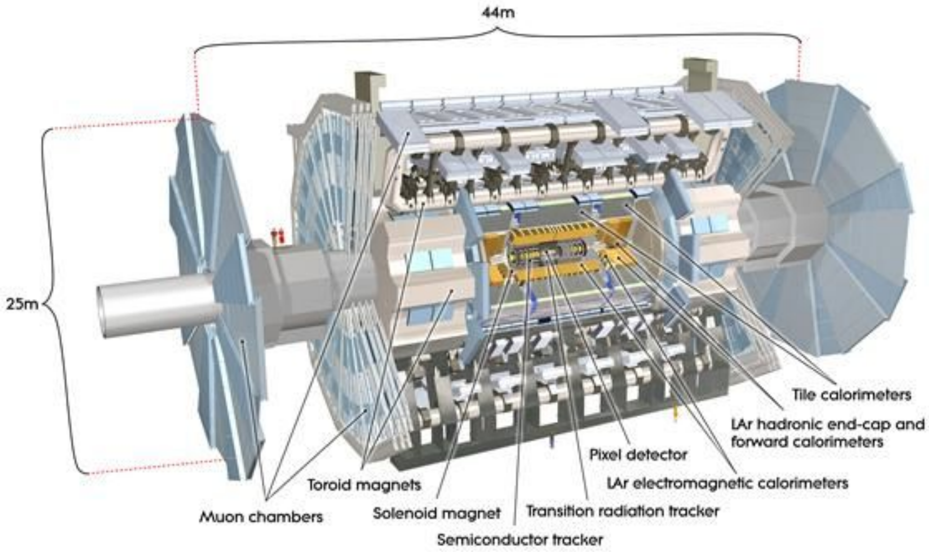




- The Standard Model is complete
    - Well, at least for its particle content
- Impressive agreement with measurements
- What's next?
    - Surprise, surprise, no one knows
    - Measure more precisely (especially Higgs)
    - Find new particles, forces, …

LHC and ATLAS





- LHC: proton-proton collider with a center-of-mass energy up to 14 TeV
  - 40 MHz collision rate
- ATLAS: general purpose detector
  - Optimized to discover the Higgs boson and find new particles in the TeV scale
- As of today the LHC has provided to ATLAS:
  - 12 millions Higgs bosons
  - 180 millions ttbar events
  - 40 billions W bosons and 13 billions Z bosons
  - And A LOT of annoying jets

Overview

LHC Run 1

Wc production cross section

Wbb production cross section

Inelastic scattering cross section (minimum bias)

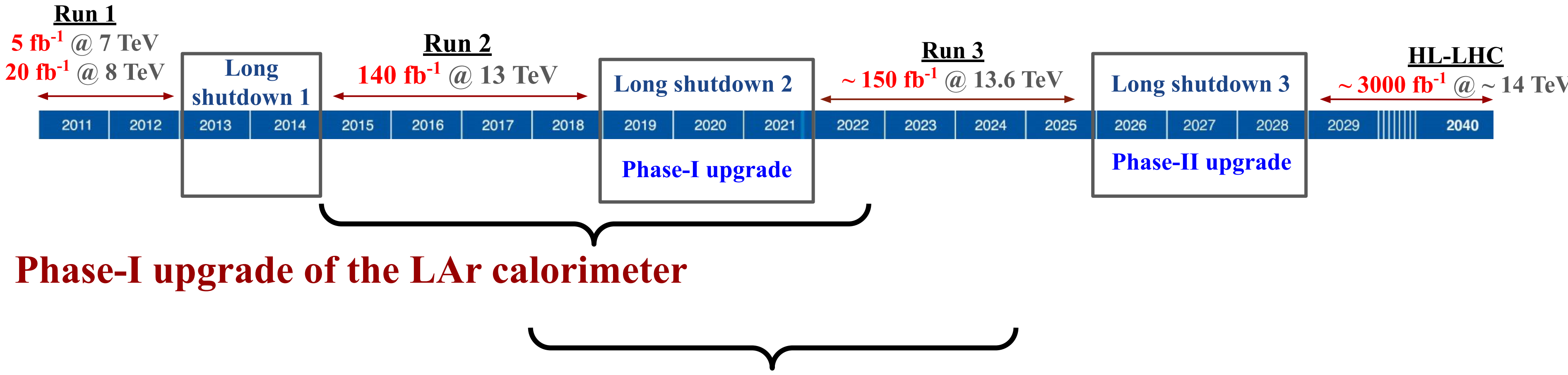
Search for the Higgs boson in WH(H→bb)

Tagging b-jets

LHC Run 2

Search for the Higgs boson in ttH(H→bb)

ttbb production cross section



Phase-I upgrade of the LAr calorimeter

Phase-II upgrade of the LAr calorimeter

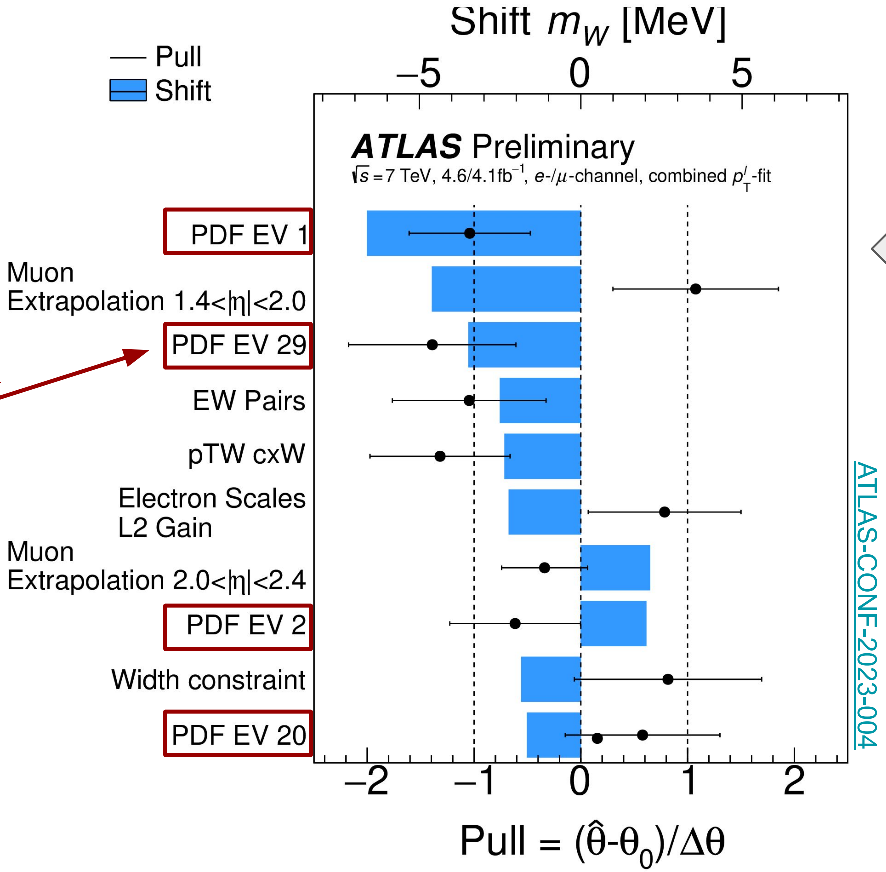
Wc cross section measurements

Probing the strange-quark PDF

JHEP 05 (2014) 068

Parton Density Functions (PDFs)

- All starts with PDFs in high-energy proton-proton collisions
  - Fundamental at LHC
- e.g. main uncertainty in W mass measurements
  - With a non-negligible contribution for the s-quark PDF



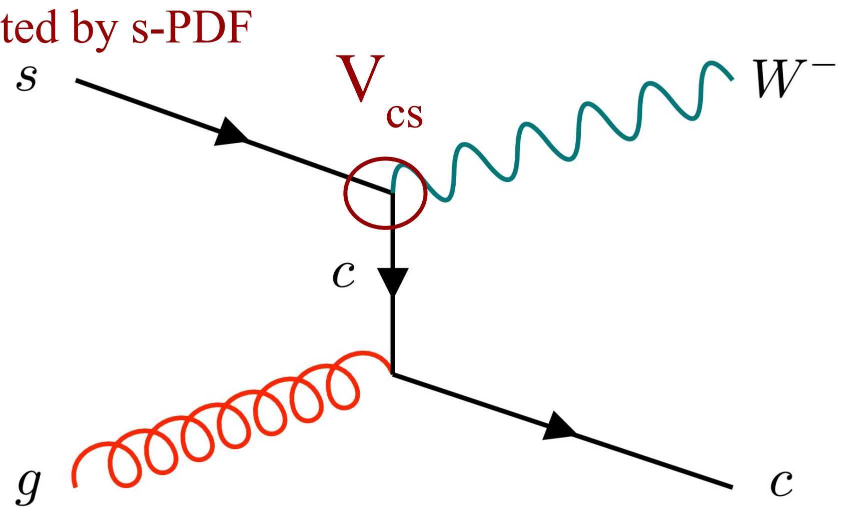
Significant contribution from s-PDF

W mass measurement Systematics ranking

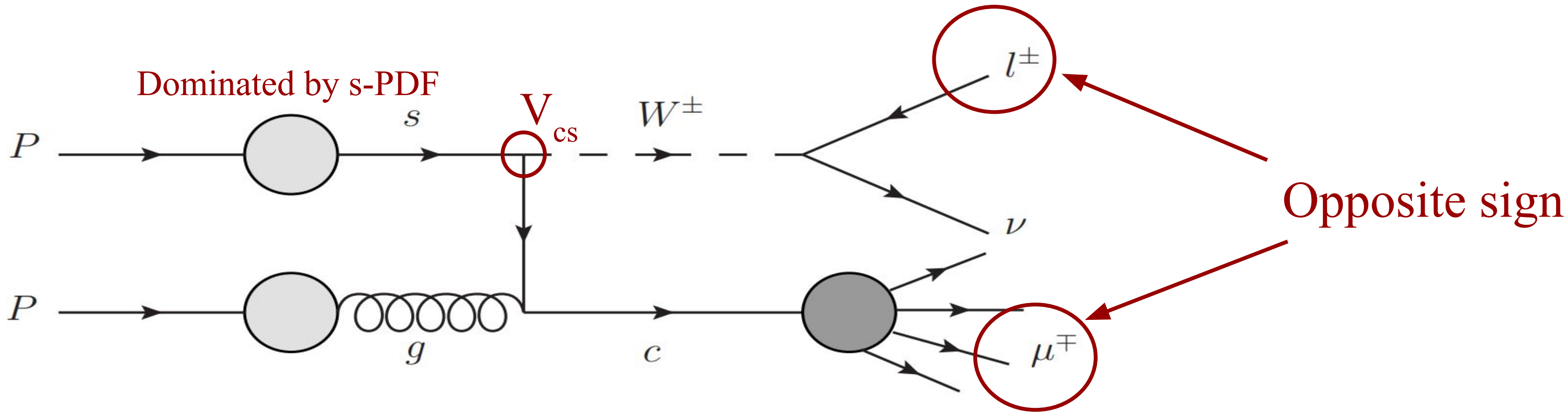
Wc: Introduction

- Situation of the s-PDF during LHC run 1
  - Constraints mainly from low energy fixed target experiments
  - s/$\overline{\mathrm{d}}$ ~50% in common PDFs
  - Hints of asymmetry between s and $\overline{\mathrm{s}}$ from the NuTeV experiment
- Tension between common PDFs and ATLAS W/Z data

The Production of a W boson with a single c-quark provides a unique direct access to the s-quark PDF



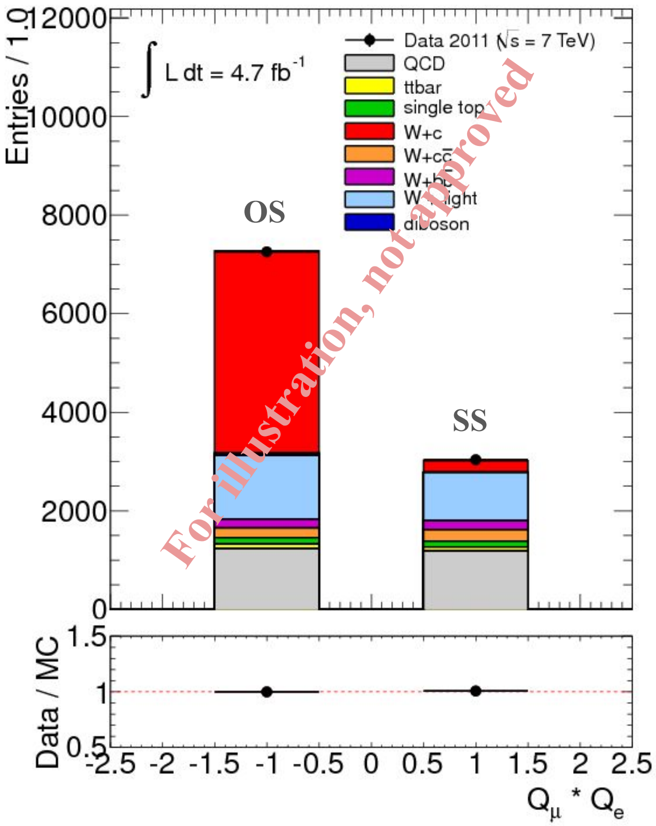
Wc: Analysis strategy



Dominated by s-PDF

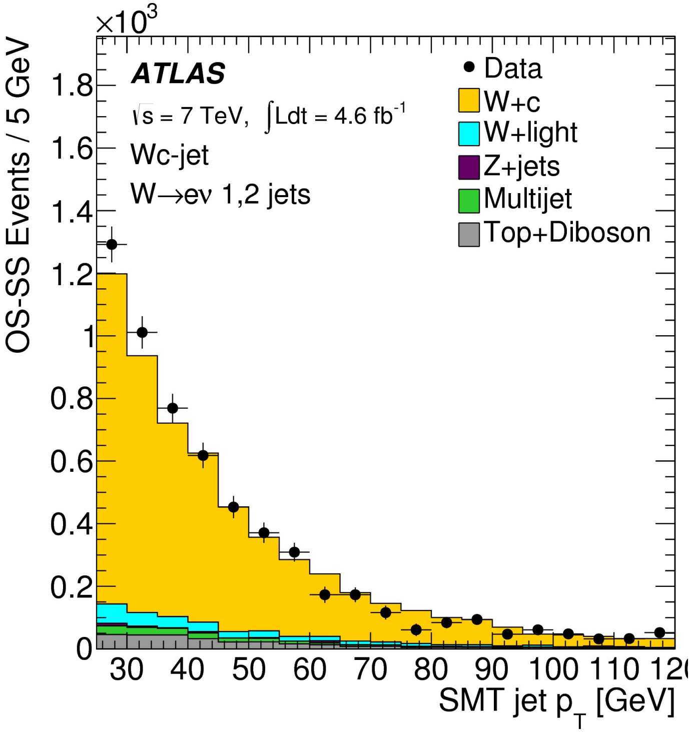
Opposite sign

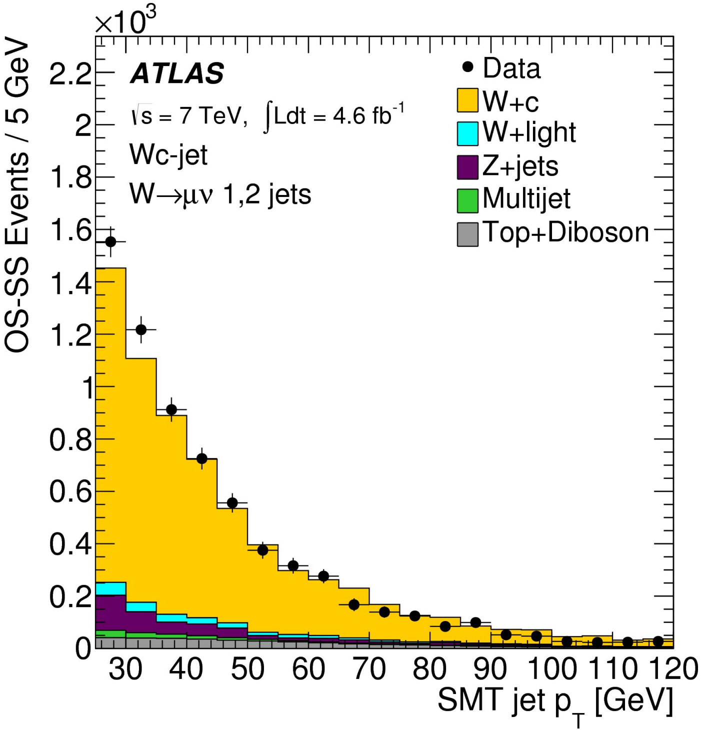
- Opposite sign of the W and the c-quark charge
- Most backgrounds are charge symmetric
  - Remove same-sign (SS) from opposite-sign (OS)
- Obtain very pure Wc contribution
- Access charm-jet charge through soft muon decay
  - Soft-muon tagger to identify the charm jet
- Combine with measurements using D-meson reconstruction to identify the c-quark



Wc: Selection and backgrounds

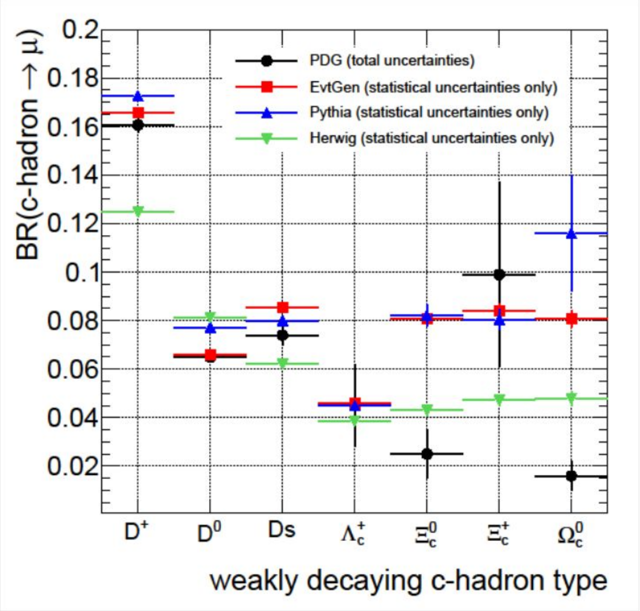
- Typical W+jets selection
    - One high pT lepton with missing energy
    - At least one jet with pT > 25 GeV
    - Exactly one c-tagged jet
- OS-SS strategy removes most backgrounds
    - Small remaining contribution from multijet, W+jets and Z+jets (muon channel)
    - Data driven estimation for all these 3 backgrounds





Wc: Systematic Uncertainties

- Dominant uncertainties
  - Jet energy scale
  - Background yields
  - c fragmentation and decay
- Large statistical contribution to background uncertainties
  - Due to OS-SS procedure



Relative systematic uncertainty in %	$W(e\nu)c$-jet	$W(\mu\nu)c$-jet
Lepton trigger and reconstruction*	0.7	0.8
Lepton momentum scale and resolution*	0.5	0.6
Lepton charge misidentification	0.2	-
Jet energy resolution*	0.1	0.1
Jet energy scale	2.4	2.1
E_T^{miss} reconstruction*	0.8	0.3
Background yields	4.0	1.9
Soft-muon tagging	1.4	1.4
c-quark fragmentation	2.0	1.6
c-hadron decays	2.8	3.0
Signal modelling	0.9	0.2
Statistical uncertainty on response	1.4	1.4
Integrated luminosity*	1.8	1.8
Total	6.5	5.3

Fragmentation function/fraction
and
c-hadron semi-muonic decay
reweighted to measurements

Wc: Measurement procedure

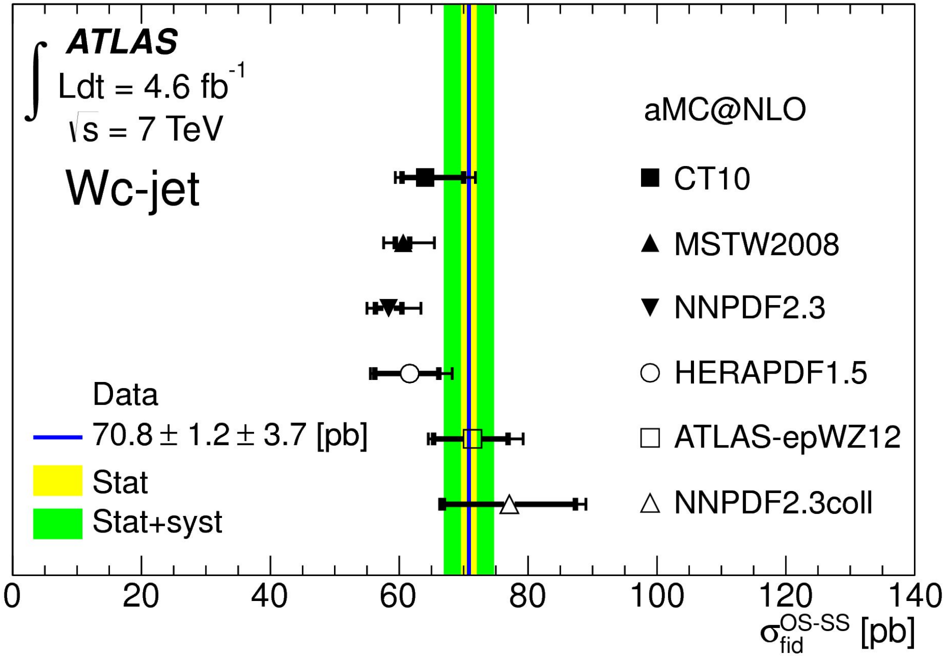
- Global chi2 fit with uncertainties as nuisance parameters
    - Includes measurements from the W+c-jet and W+D-meson analyses
    - Using the HERAFitter framework
- Individual measurements that contribute to a common cross section are combined
    - e.g. electron and muon channels

$$\chi^2 = \sum_{k,i} w_k^i \frac{\left[\overset{\text{Individual measurements}}{\boxed{\mu_k^i}} - \left(\overset{\text{Combined measurement}}{\boxed{m^i}} + \sum_j \gamma_{j,k}^i \overset{\text{Uncertainties as nuisance parameters}}{\boxed{m^i b_j}}\right)\right]^2}{(\delta^i_{\text{sta},k})^2 \mu_k^i \left(m^i - \sum_j \gamma^i_{j,k} m^i b_j\right) + (\delta^i_{\text{unc},k} m^i)^2} + \sum_j \overset{\text{Penalty term}}{\boxed{b_j^2}}.$$

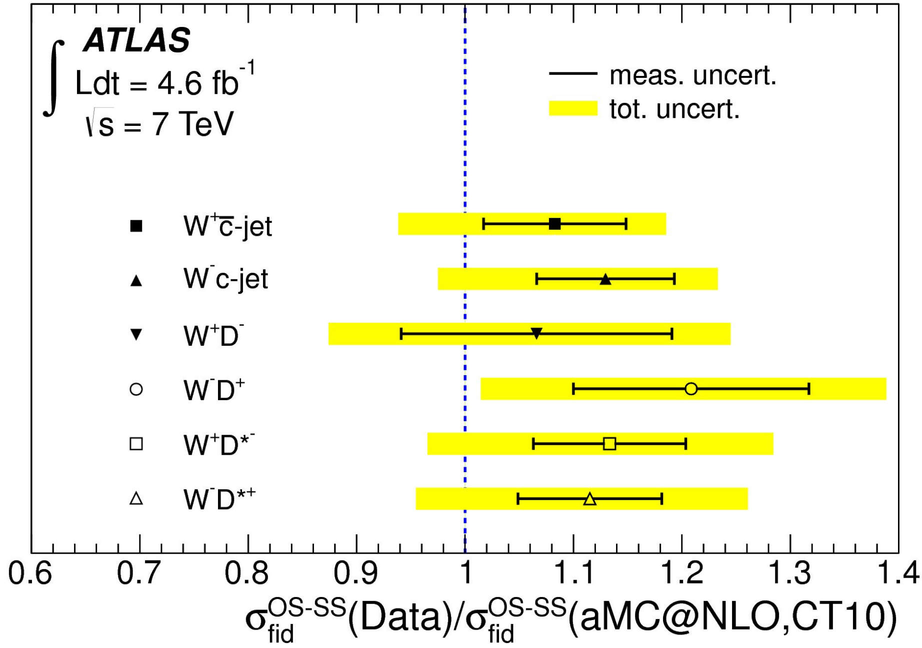
Wc: Cross section measurements

- Inclusive cross section compared with different PDFs
  - With aMC@NLO generator (Wc @NLO in QCD)
- PDFs with low strange-to-down sea ratio underestimate the cross section
- Compatible with ATLAS PDF extracted with W/Z data in 2012
- Consistent results between all Wc channels

Comparison of different PDF sets

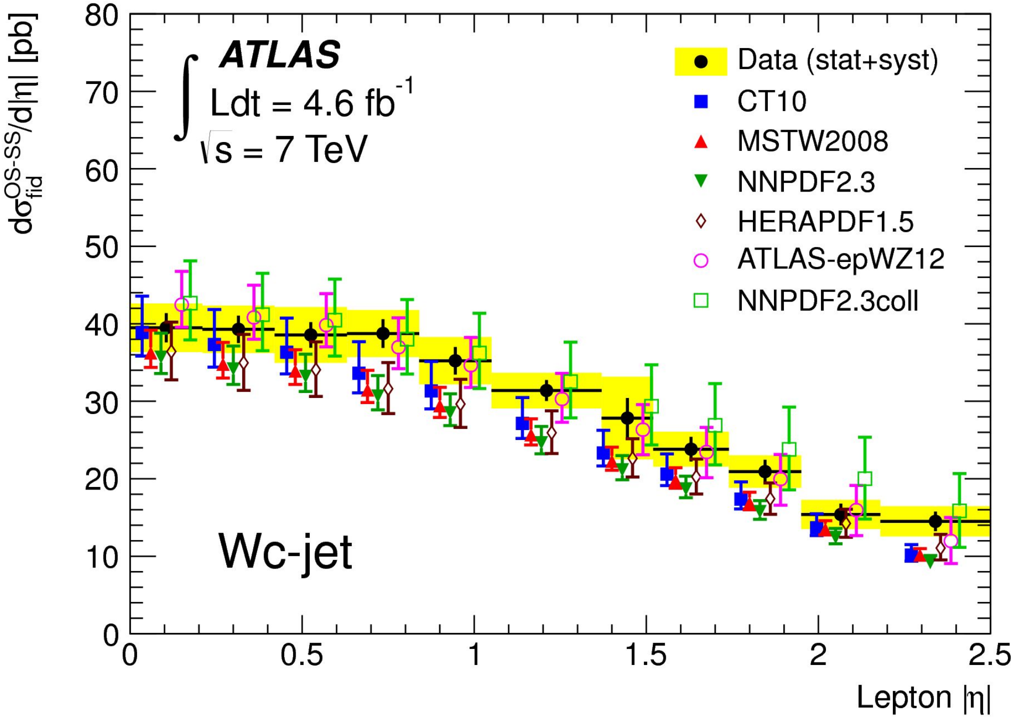


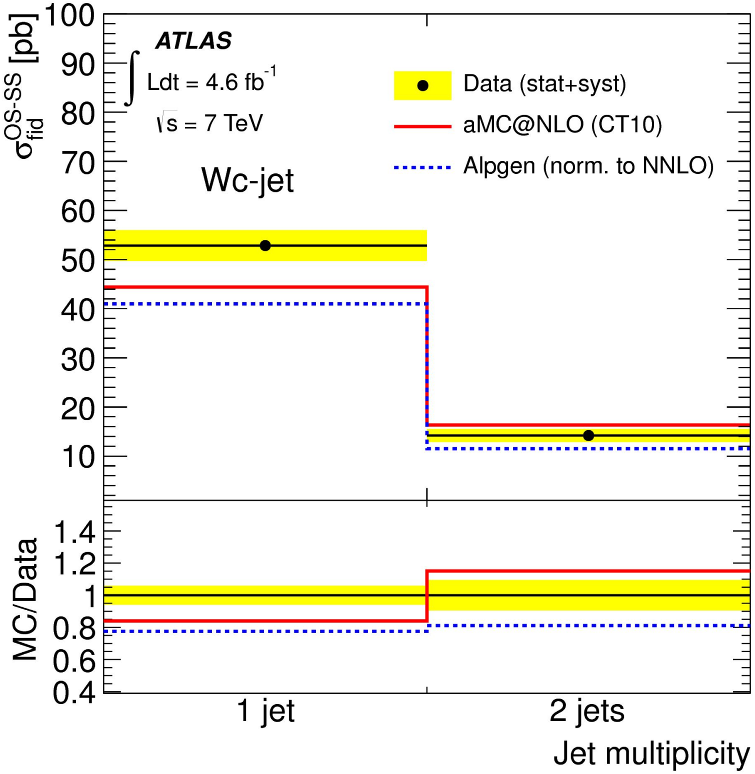
Comparison of different Wc channels



Wc: Differential cross section measurements

- Cross section as function of the pseudo-rapidity
  - Shape agreement with all PDFs within the uncertainties
  - Main mismodeling comes from normalisation
- Test NLO QCD generator modeling with N_{jets} distribution
  - LO multi-leg generators (Alpgen) describe better the N_{jets} distribution
  - But LO underestimates the inclusive fiducial cross section





Wc: W^+/W^- ratio

- W^+/W^- ratio probes the $s/\bar{s}$ asymmetry
- Slightly higher W^- cross section due to contribution from the d-PDF
- Results comparable with symmetric contribution of s and $\bar{s}$

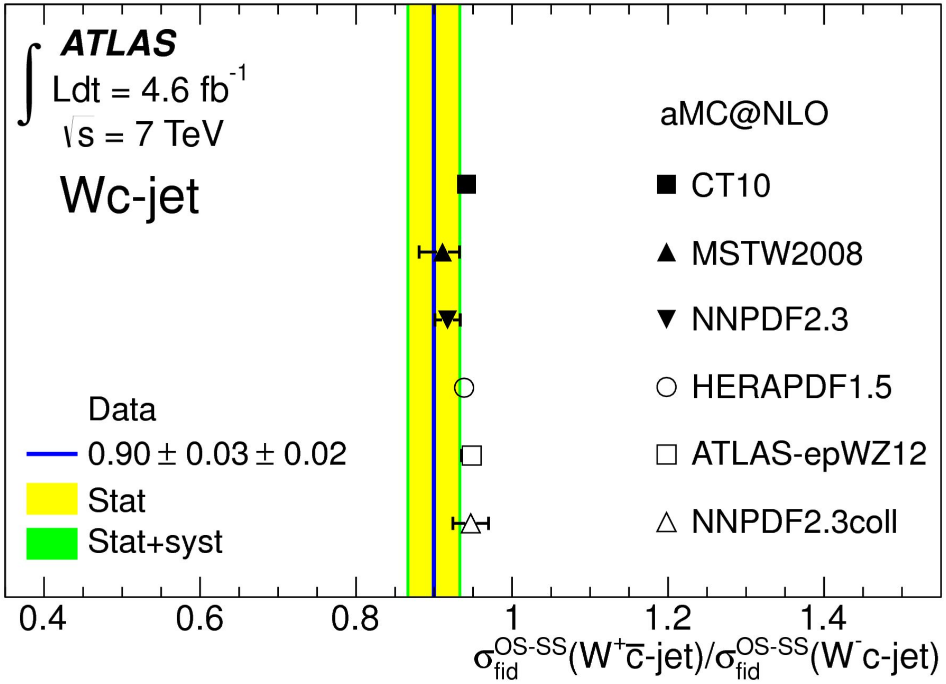
$$A_{s\bar{s}} = \frac{\langle s(x,Q^2)\rangle - \langle \bar{s}(x,Q^2)\rangle}{\langle s(x,Q^2)\rangle}$$

CT10: s-PDF = $\bar{\text{s}}$-PDF

⇩

$$A_{s\bar{s}} \approx R_c^{\pm}(\text{CT10}) - R_c^{\pm}(\text{Data})$$

$$A_{s\bar{s}} = (2 \pm 3)\%$$



Wc: PDF compatibility

Measurements

Predictions

PDF uncertainty as nuisance parameter

PDF uncertainty penalty term

$$\chi^2 = \sum_{k,i} w_k^i \frac{\left[\mu_k^i - m^i\left(1+\sum_j \gamma_{j,k}^i b_j + \sum_j (\gamma^{\text{theo}})_{j,k}^i b_j^{\text{theo}}\right)\right]^2}{(\delta_{\text{sta},k}^i)^2 \Delta_i^k + (\delta_{\text{unc},k}^i m^i)^2} + \sum_j b_j^2 + \sum_j \left(b_j^{\text{theo}}\right)^2$$

Chi2 fit to estimate the compatibility with different PDFs
All compatible within 2 sigma

Two sigma tension with NNPDF2.3

HERAPDF with $s/\bar{d} \sim 50\%$ at $Q^2 = 1.9$ GeV^2

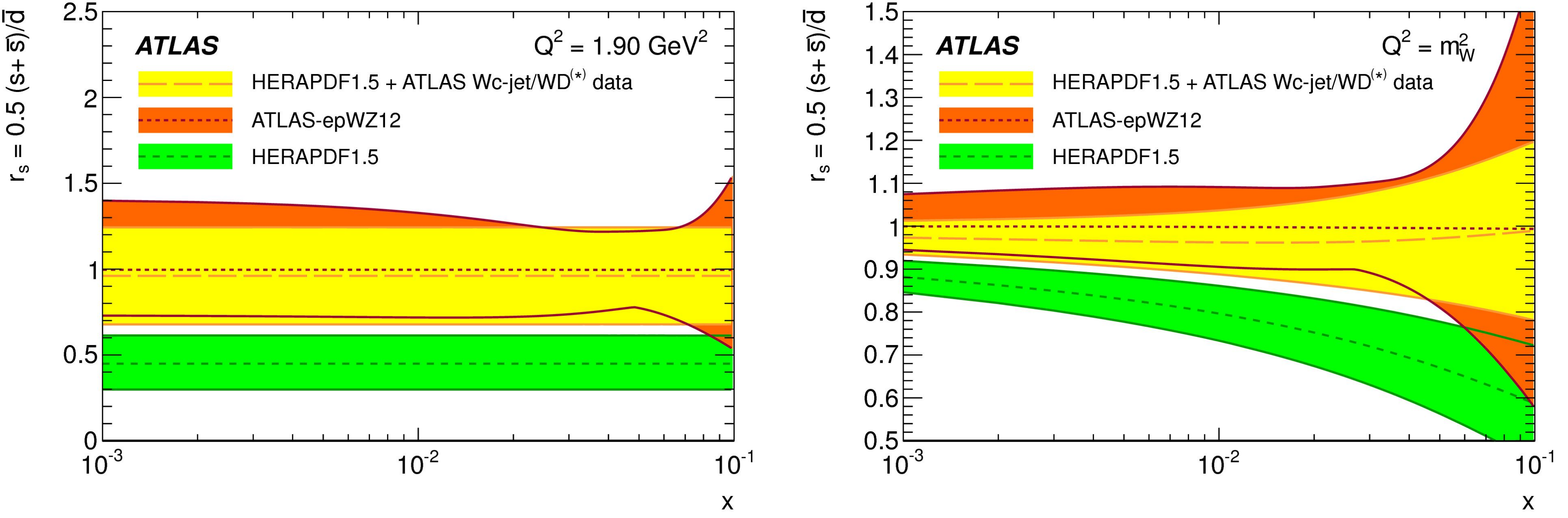
Best compatibility with NNPDF (with only collider data) and ATLAS PDF

	NNPDF2.3	MSTW2008	CT10	HERAPDF1.5	ATLAS-epWZ12	NNPDF2.3COLL
Chi2/ndof	52.1/38	41.3/38	33.6/38	28.0/38	19.2/38	18.2/38
Probability	6.3%	32.8%	67.3%	88.3%	99.5%	99.7%

Wc: Constraints on the strange-PDF

- Strange-to-down PDF ratio modeled as a single parameter in HERA PDF
    - Fit this parameter by removing the corresponding constraints from HERA PDF
- Strange-to-down PDF ratio compatible with one
    - Confirms SU(3) symmetry in the PDF (seen in ATLAS W/Z PDF fits)

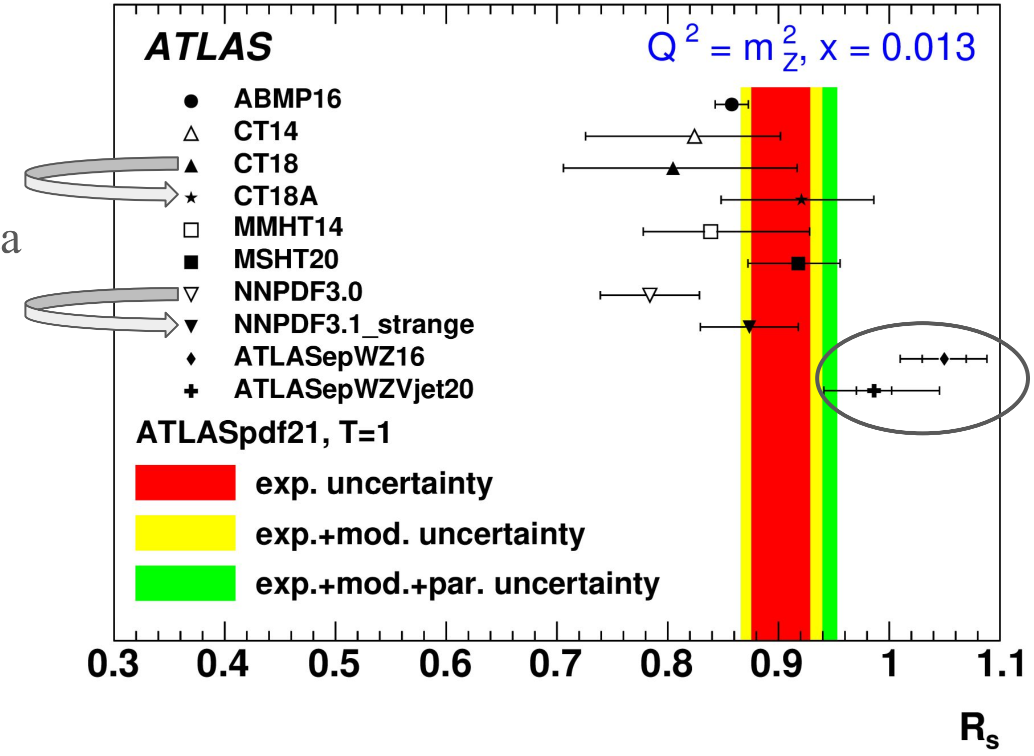
$$r_s \equiv 0.5(s+\bar{s})/\bar{d} = 0.96 \begin{smallmatrix}+0.16 \\ -0.18\end{smallmatrix} \begin{smallmatrix}+0.21 \\ -0.24\end{smallmatrix}$$



Wc: Compatibility with recent measurements

- Recent ATLAS PDF fit confirms enhancement of strange-to-down density
  - However the enhancement is slightly lower than previous ATLAS PDFs
  - Still compatible with the Wc measurement presented above
- Most recent PDFs implement higher strange density fraction
  - With respect to the PDFs we had in LHC run 1

Eur. Phys. J. C **82** (2022) 438



Include LHC data

Previous ATLAS PDFs

Search for ttH(bb)

Probing the top and bottom Yukawa couplings

Phys. Rev. D 97 (2018) 072016

Higgs coupling at the beginning of Run 2

- Higgs coupling to bosons established in Run 1
- Higgs decay to **τ** leptons observed
- Run 2 focuses on Higgs coupling to third generation quarks
    - Top Yukawa largest in the Standard Model
    - H→bb decay largest in the Standard Model

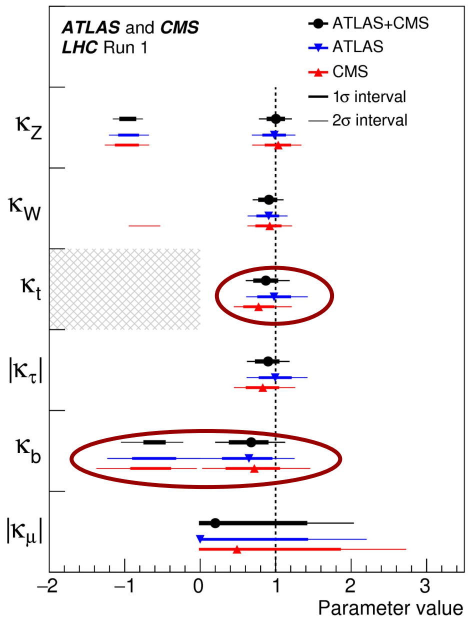
ATLAS+CMS Run 1

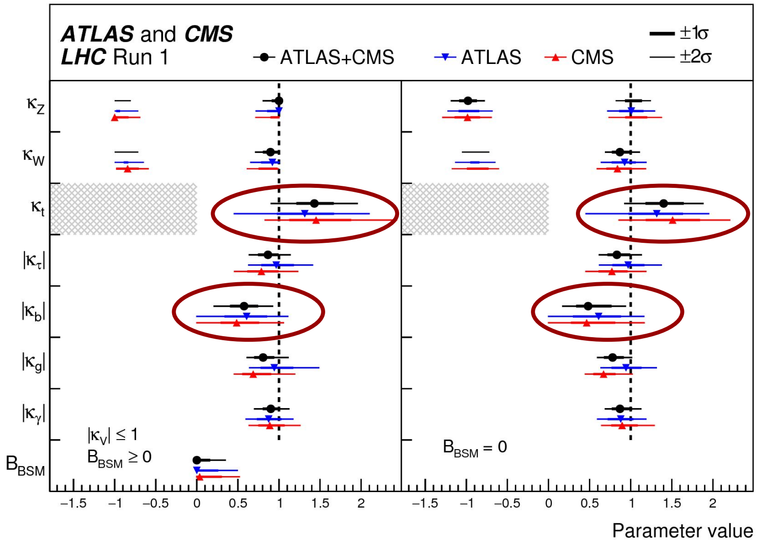
Significance	Expected	Observed
H→ττ	5.0σ	5.5σ
H→bb	3.7σ	2.6σ
ttH	2.0σ	4.4σ

	No BSM	BSM in loops
$\boldsymbol{\kappa_\tau}$	$\mathcal{O}(15\%)$	$\mathcal{O}(15\%)$
$\boldsymbol{\kappa_b}$	$\mathcal{O}(25\%)$	$\mathcal{O}(20-30\%)$
$\boldsymbol{\kappa_t}$	$\mathcal{O}(15\%)$	$\mathcal{O}(30\%)$

$$\kappa_i^2 = \frac{\Gamma^i}{\Gamma^i_{SM}}$$

JHEP 08 (2016) 045



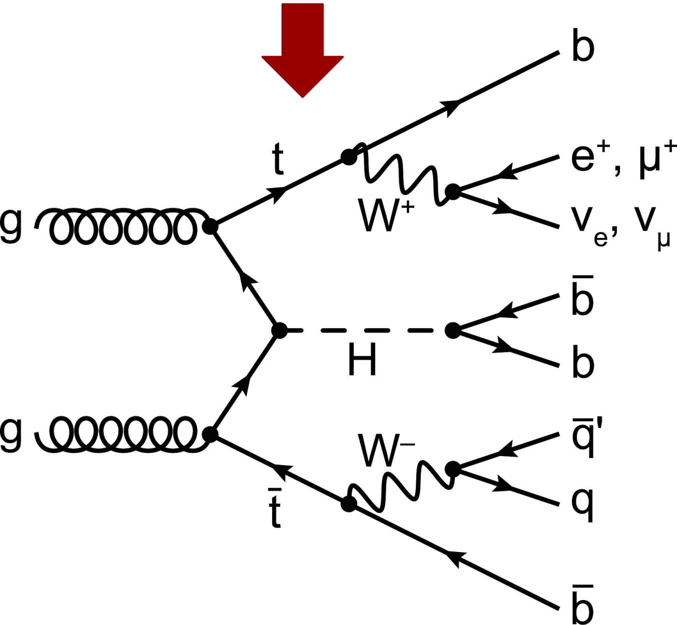


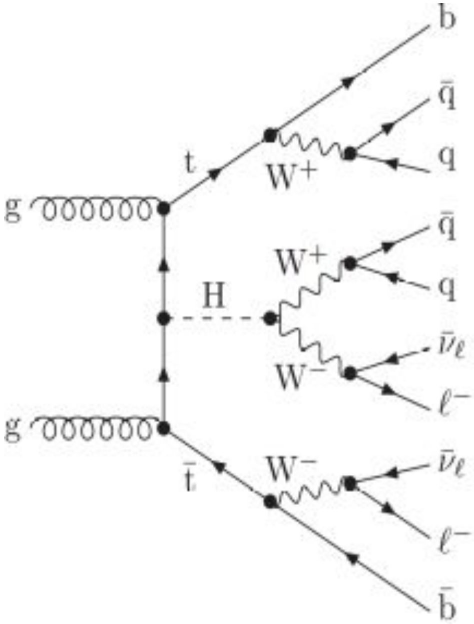
ttH(bb): Introduction

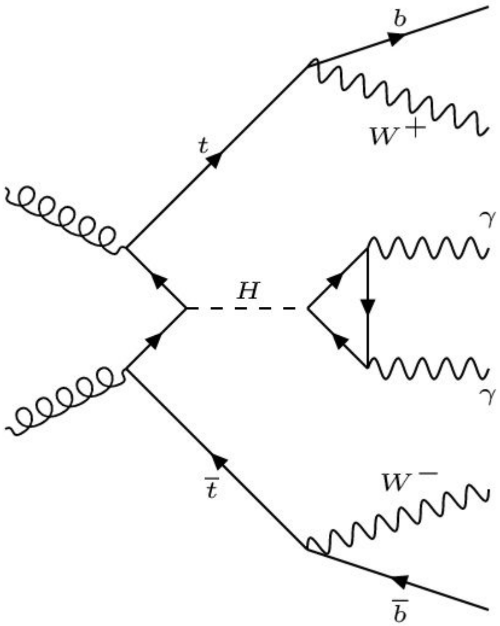
- ttH provides a direct way to probe the top yukawa coupling
    - Complementary to loop-induced sensitivity in gg→H and H→γγ
- Small cross section and complex final state
    - Explore all available channels
- ttH(H→bb) channel described here
    - Large branching ratio of H→bb
    - However large ttbb background which is hard to model

Analysis with partial Run 2 data (2016)
36.1 fb^{-1}

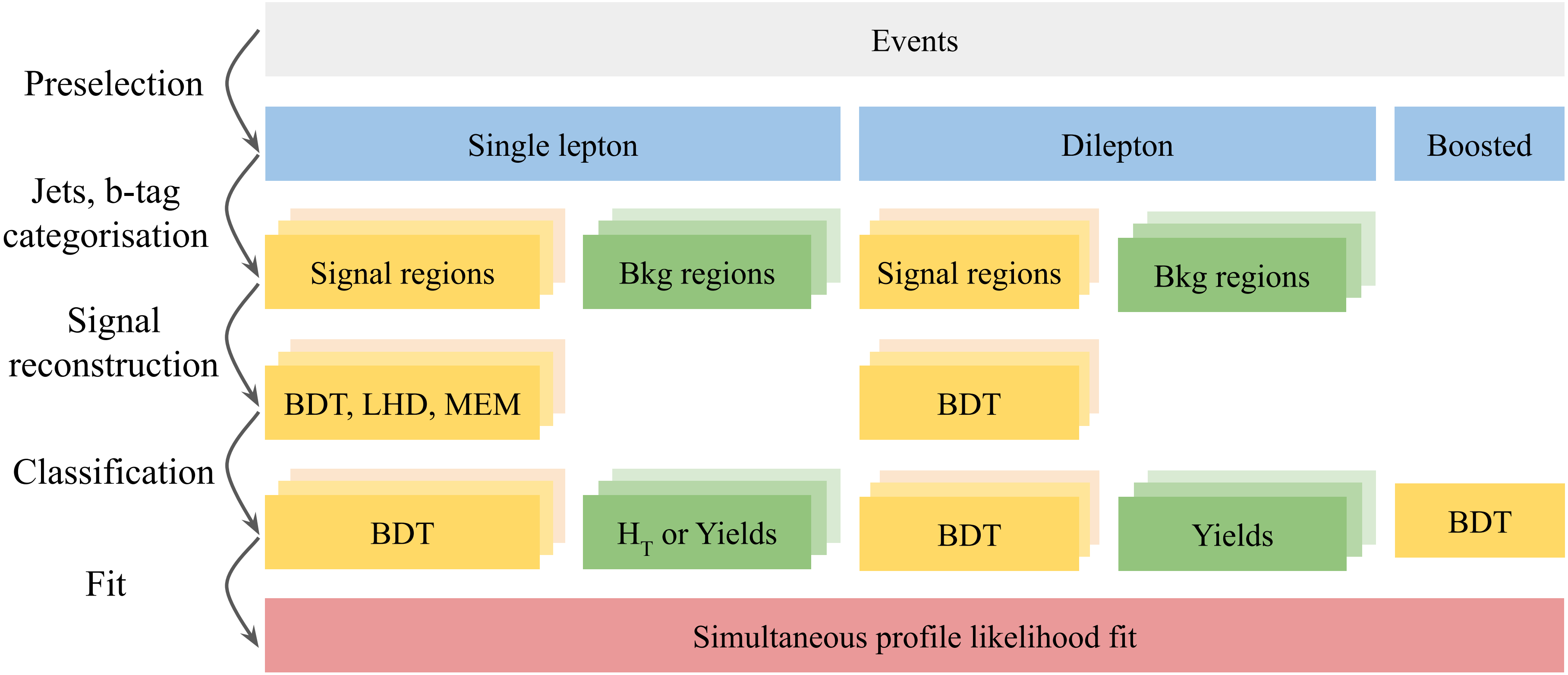
Increased statistics ← → Increased purity





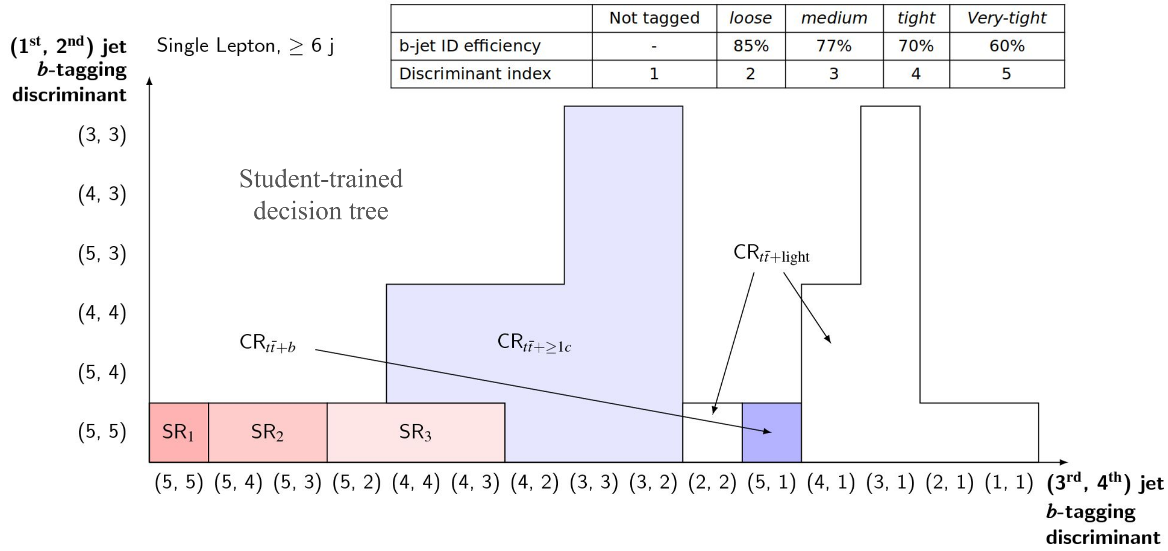


ttH(bb): Analysis strategy

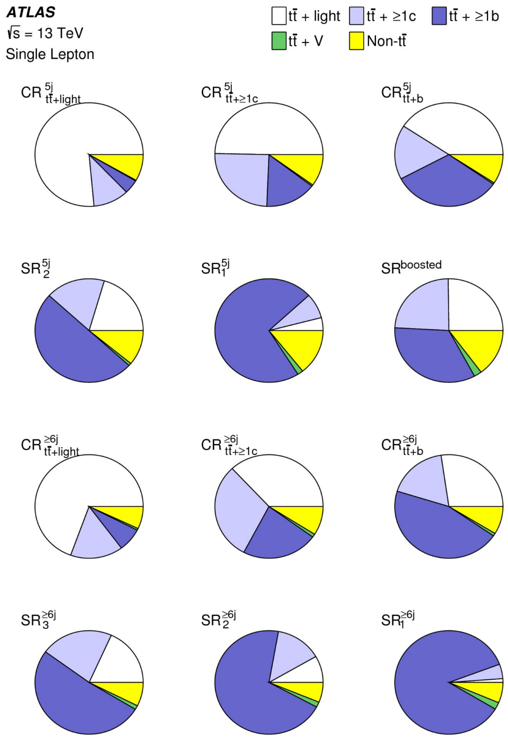


- Explore different W-boson (from top) decays and topologies
  - Single lepton, dilepton and boosted
- Categorisation based on the number of jets and b-jets
  - Separate ttH, ttbb, ttcc, and tt+light contributions in different regions
- Extensive use of MVA for signal reconstruction and background separation

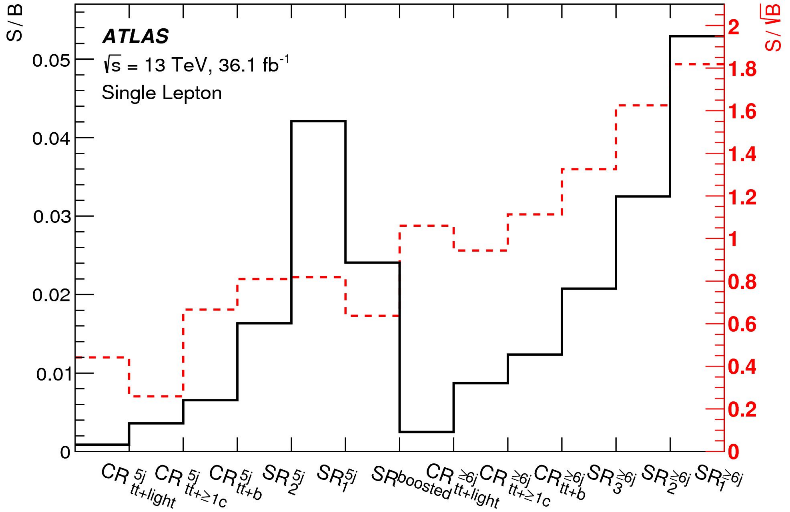
ttH(bb): Categorisation



	Not tagged	*loose*	*medium*	*tight*	*Very-tight*
b-jet ID efficiency	-	85%	77%	70%	60%
Discriminant index	1	2	3	4	5

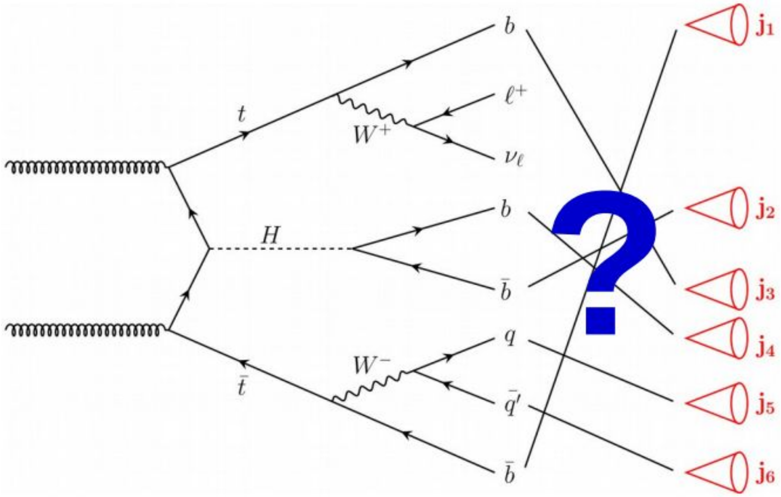


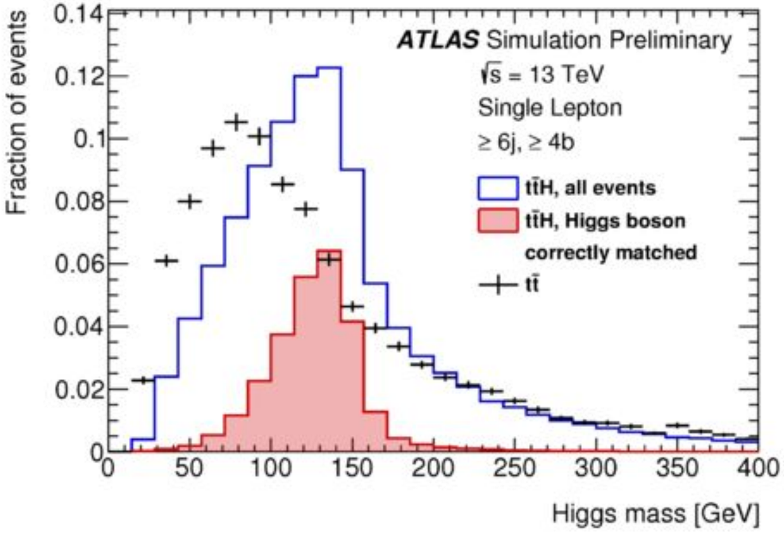
- Categorize jets based on b-tagging probability
    - b-tagging discriminant index for each jet
- Categorise events based on b-jets index
    - Regions with different tt+jets composition
    - Constrain tt+l-jets, tt+c-jets, and tt+b-jets
- Signal regions dominated by ttbb background
    - Low S/B: ~ 5%
    - Extensive usage of MVAs needed

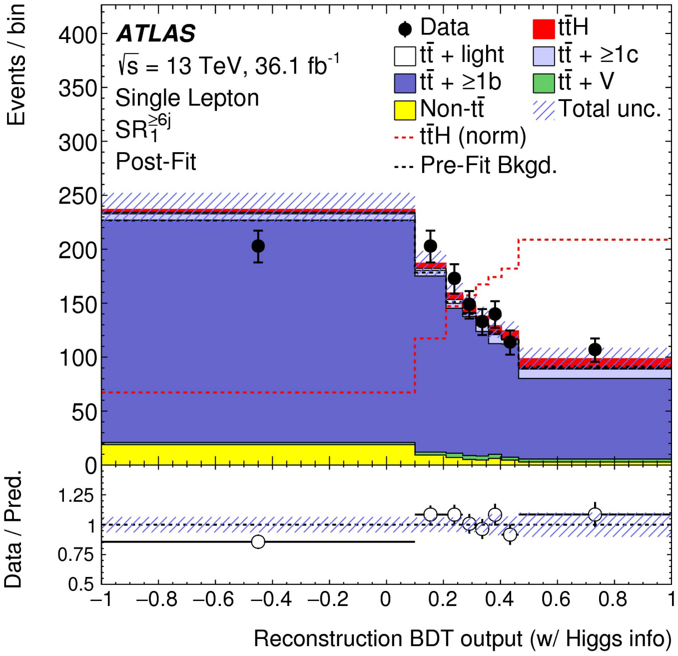


ttH(bb): Signal reconstruction

- 4 b-jets in the event
  - Solve combinatorics to reconstruct the Higgs boson
- BDT used to find the correct jet matching
- Correct Higgs combination:
  - 30% if Higgs decay product not used in the training
  - 50% if Higgs information used
- Reconstruction BDT output has large separation power with ttbb background
  - Likelihood and MEM discriminants in some regions to further increase separation

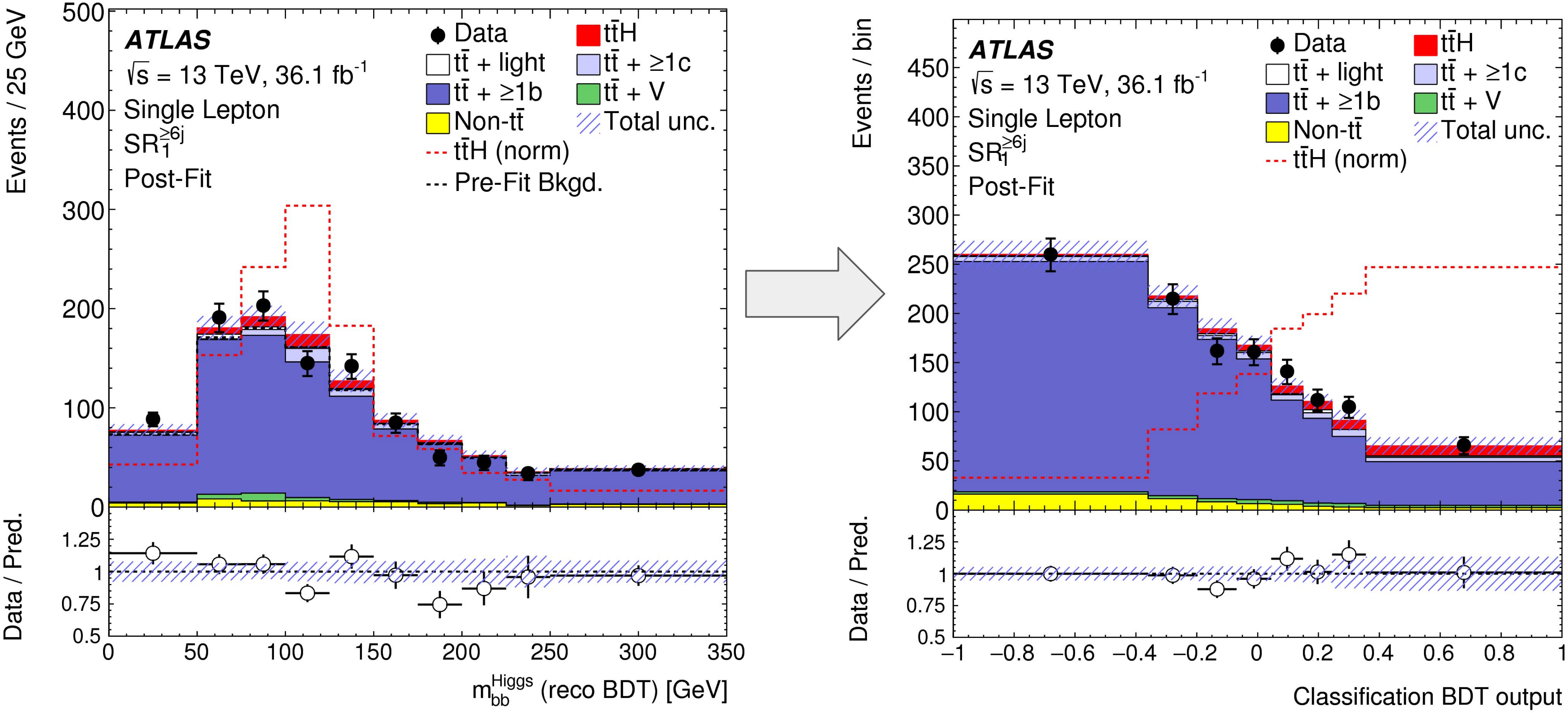






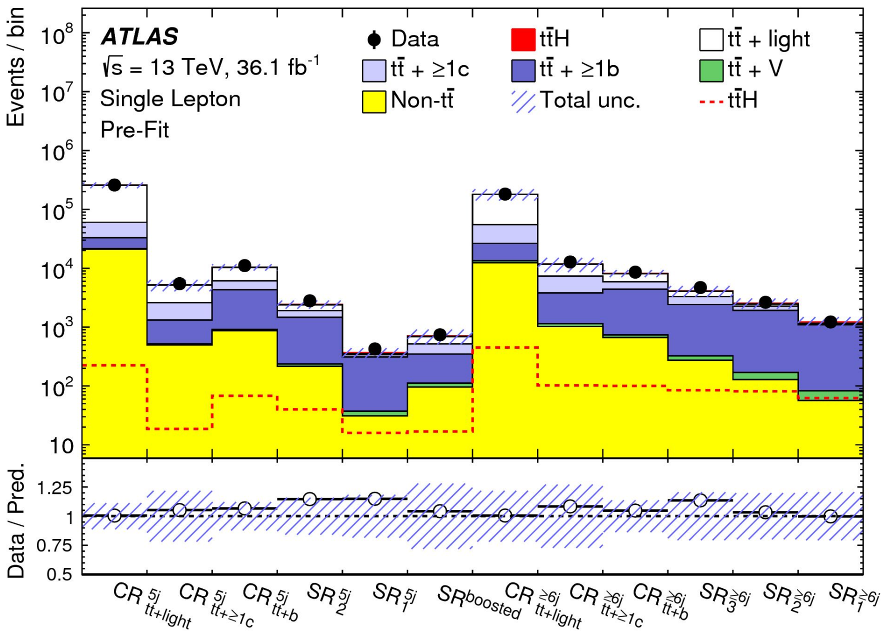
ttH(bb): Signal/background separation

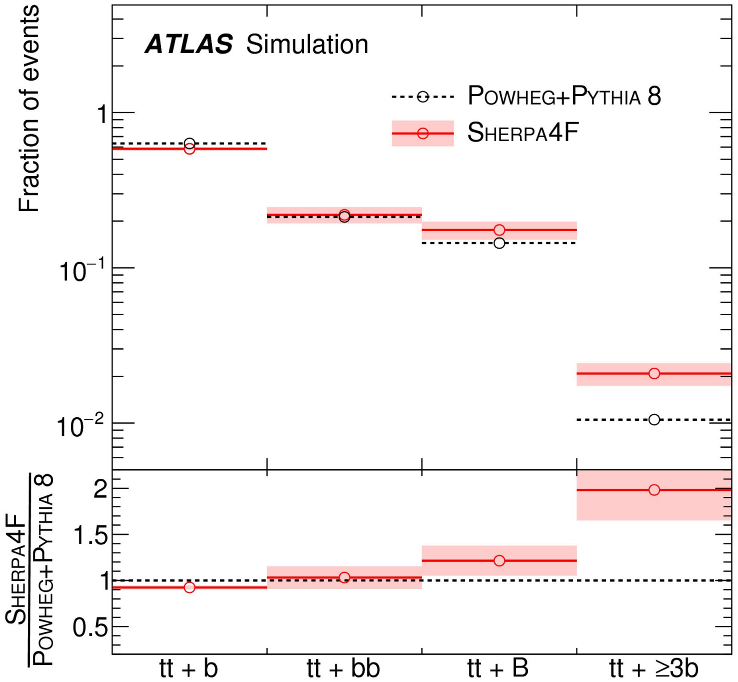
- Final classification BDT used for discrimination
  - Including reconstruction variables + global variables



ttbb: Modeling

- Important ttbb mismodeling observed in previous searches and ATLAS measurements
  - Consistent underestimation of ttbb for different MC generators
- Complex ttbb model to combine the advantages of different ttbb computations
  - 4 flavor scheme (4F) allows “better” description of ttbb process (massive b quarks)
  - 5 flavour scheme (5F) allows a complete description of tt+jets
  - Split ttbb in several sub-components and reweight 5F to 4F
  - Include a list of systematics uncertainties on each components by comparing several generators

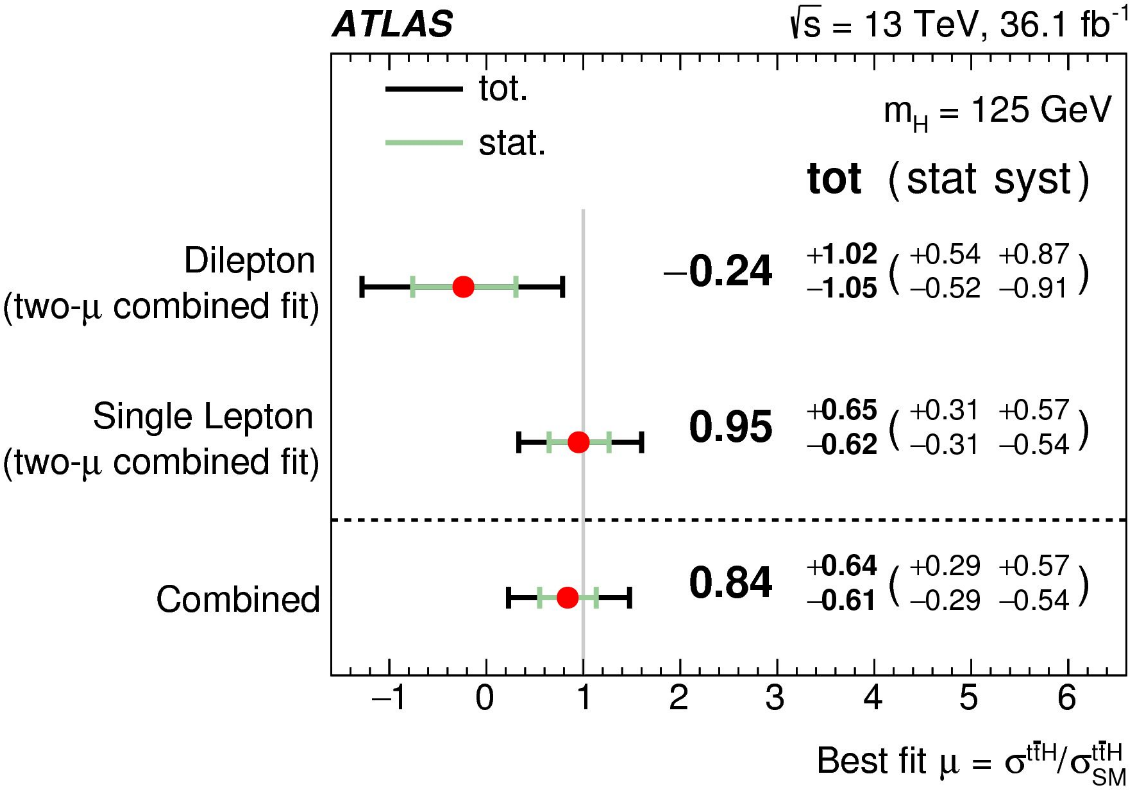




ttH(bb): Results

- Profile likelihood fit in 9 signal regions and 10 control regions
- No significant excess found but results compatible with the SM
- Tremendous effort to understand the modeling of tt+jets in the fit
    - Results completely dominated by systematic uncertainties

Expected significance: 1.6
Observed significance: 1.4



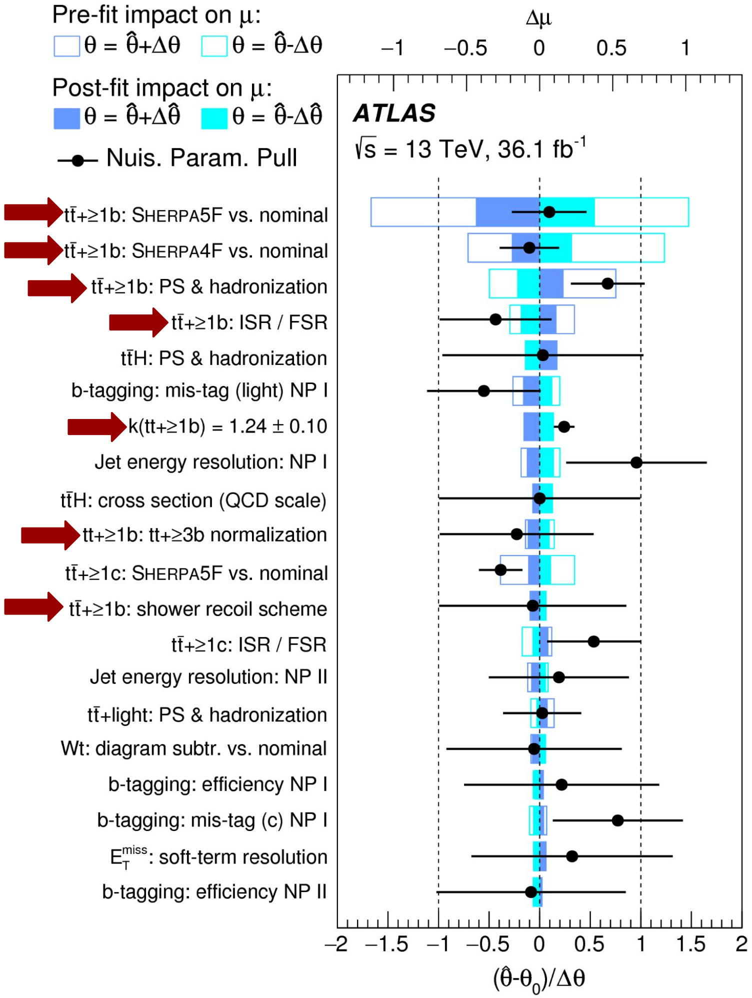
ttH(bb): Systematic uncertainties

- Paranoia-based ttbb uncertainties
  - Comparing every possible computation on the market
  - Large statistical component in these systematics
- ttbb systematics completely dominates the sensitivity
  - Everything else negligible

Dominated by 4.5 flavour systematics
But we avoided pythia7 systematics
Pythia7 systematics: compare Pythia6 and Pythia8*

*Timothée T.P.: Private communication

Uncertainty source	$\Delta\mu$	
$t\bar{t}+\geq 1b$ modeling	+0.46	−0.46
Background-model stat. unc.	+0.29	−0.31
b-tagging efficiency and mis-tag rates	+0.16	−0.16
Jet energy scale and resolution	+0.14	−0.14
$t\bar{t}H$ modeling	+0.22	−0.05
$t\bar{t}+\geq 1c$ modeling	+0.09	−0.11
JVT, pileup modeling	+0.03	−0.05
Other background modeling	+0.08	−0.08
$t\bar{t}$ + light modeling	+0.06	−0.03
Luminosity	+0.03	−0.02
Light lepton (e, μ) id., isolation, trigger	+0.03	−0.04
Total systematic uncertainty	+0.57	−0.54
$t\bar{t}+\geq 1b$ normalization	+0.09	−0.10
$t\bar{t}+\geq 1c$ normalization	+0.02	−0.03
Intrinsic statistical uncertainty	+0.21	−0.20
Total statistical uncertainty	+0.29	−0.29
Total uncertainty	+0.64	−0.61



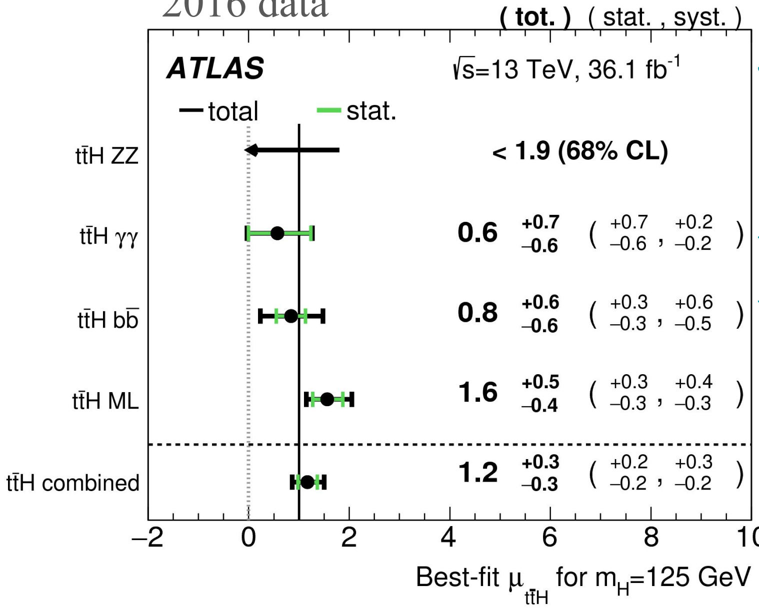
ttH: Combination and discovery

- Combination of all ttH channels
  - Discovery of the ttH process in 2017
  - Lead by the purity of the ttH(γγ) channel
- All channels compatible with SM predictions

ttH(bb) competitive @ 36fb^{-1}
ttH(γγ) takes over at higher lumi
ttH(bb) still interesting for differential measurements

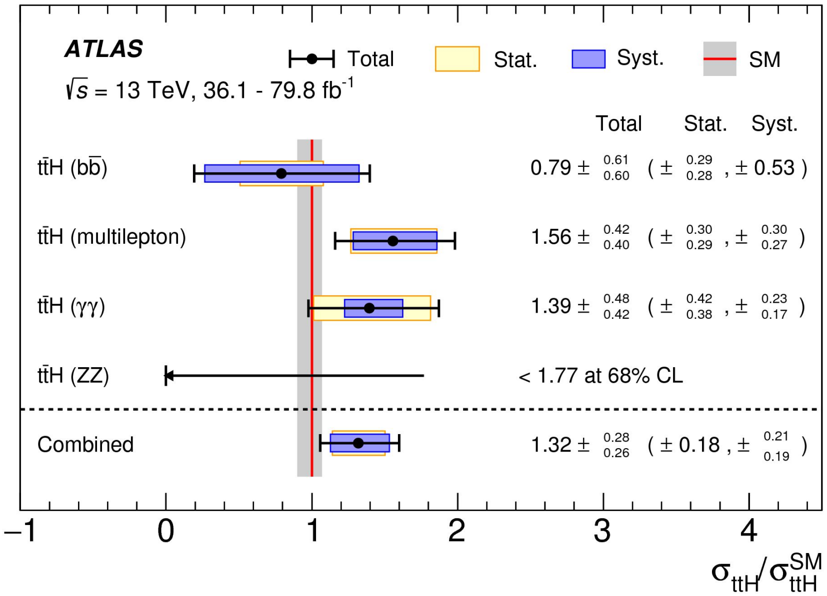
Analysis	Integrated luminosity [fb^{-1}]	Obs. sign.	Exp. sign.	**Exp. @36.1fb^{-1}**
$H \to \gamma\gamma$	79.8	4.1σ	3.7σ	**1.7σ**
$H \to$ multilepton	36.1	4.1σ	2.8σ	
$H \to b\bar{b}$	36.1	1.4σ	1.6σ	
$H \to ZZ^* \to 4\ell$	79.8	0σ	1.2σ	**0.6σ**
Combined (13 TeV)	36.1 − 79.8	5.8σ	4.9σ	**3.8σ**

2016 data



Phys. Rev. D 97 (2018) 072003

2016+2017 data

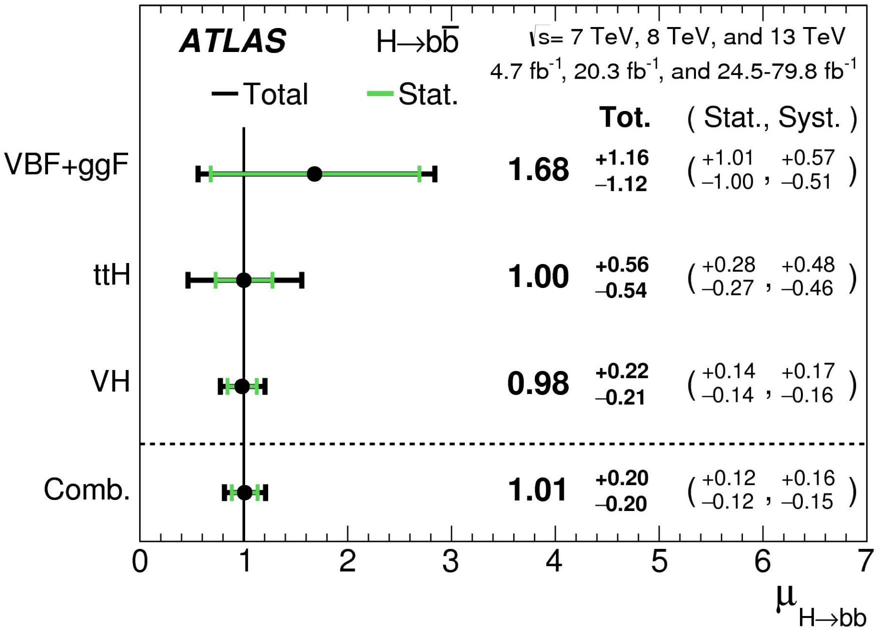


Phys. Lett. B 784 (2018) 173

H→bb: Combination and discovery

- Combination of all H→bb channels with partial Run 2 data
    - Discovery of H→bb in 2017
    - Small contribution from the ttH(bb) channel

Phys. Lett. B 786 (2018) 59



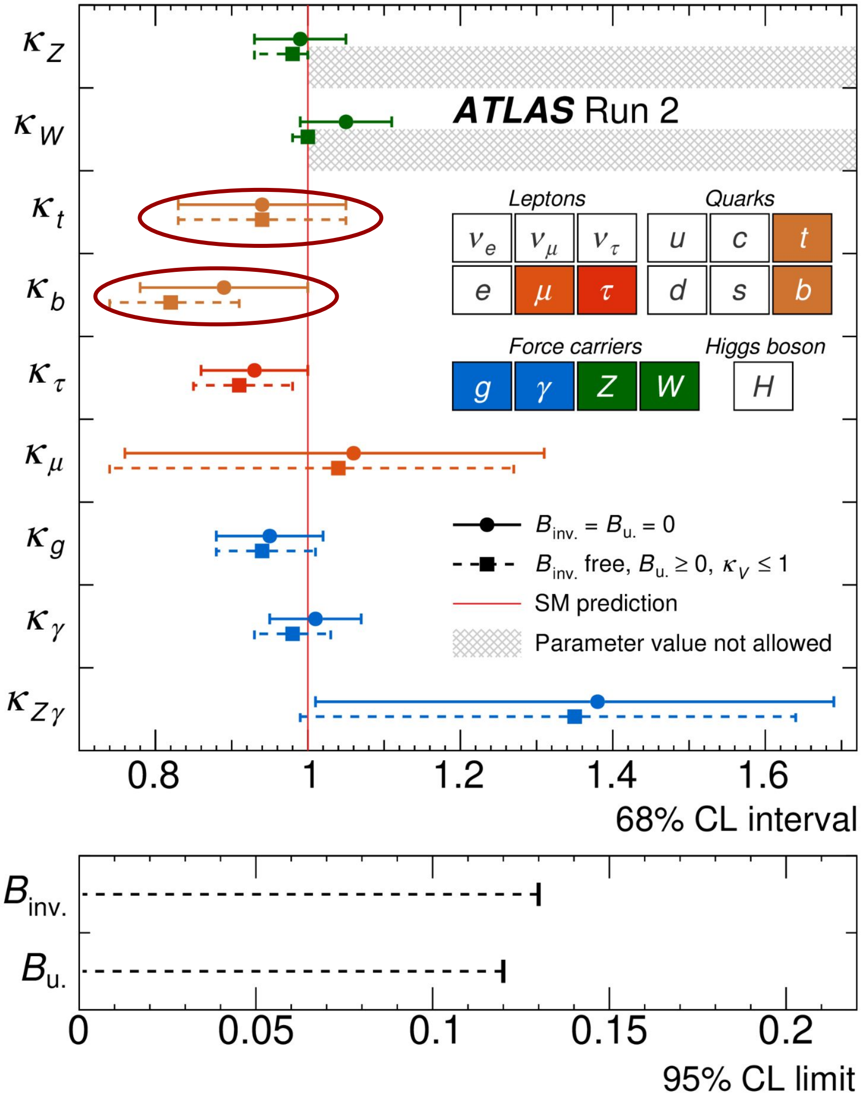
Channel	Significance	
	Exp.	Obs.
VBF+ggF	0.9	1.5
$t\bar{t}H$	1.9	1.9
VH	5.1	4.9
$H \rightarrow b\bar{b}$ combination	5.5	5.4

Higgs couplings after Run 2

- ATLAS run 2 results
  - Legacy measurements and combination with CMS yet to come
- Couplings to bosons and **τ**-leptons:
  - < 10%
- Coupling to third generation quarks:
  - O(10%)
- Couplings to muons and Zγ
  - O(30%)
- Invisible and undetected modes excluded to O(10%) level
- 61% compatibility with SM
  - If B_{inv} and B_u set to zero

It looks like a SM Higgs, it couples like a SM Higgs, …
Well there is still self coupling

Nature 607, 52 (2022)

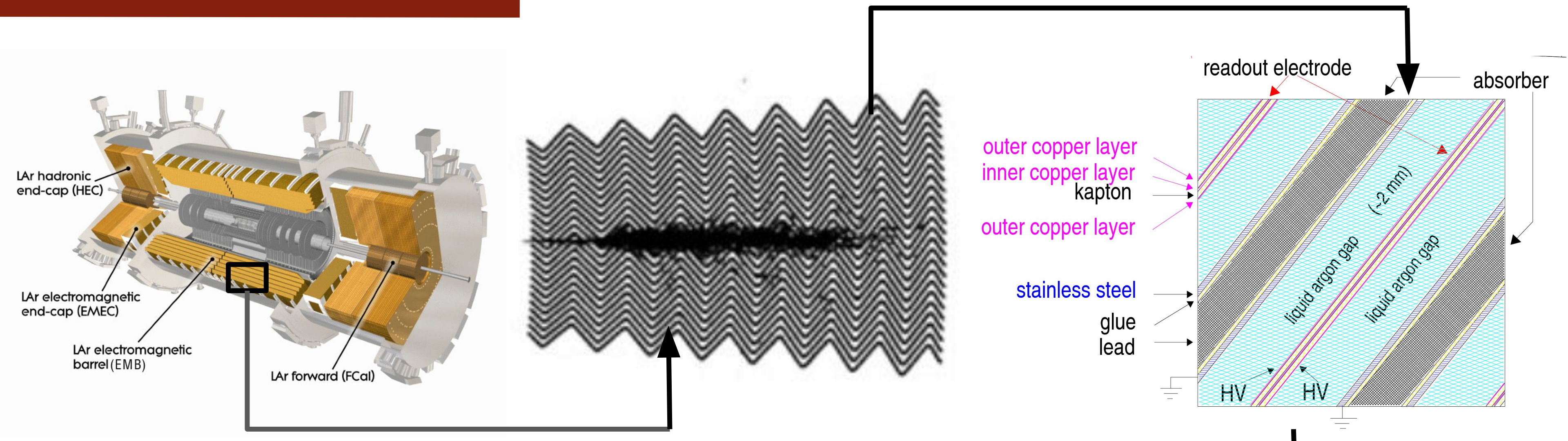


Upgrade of the Liquid Argon Calorimeter

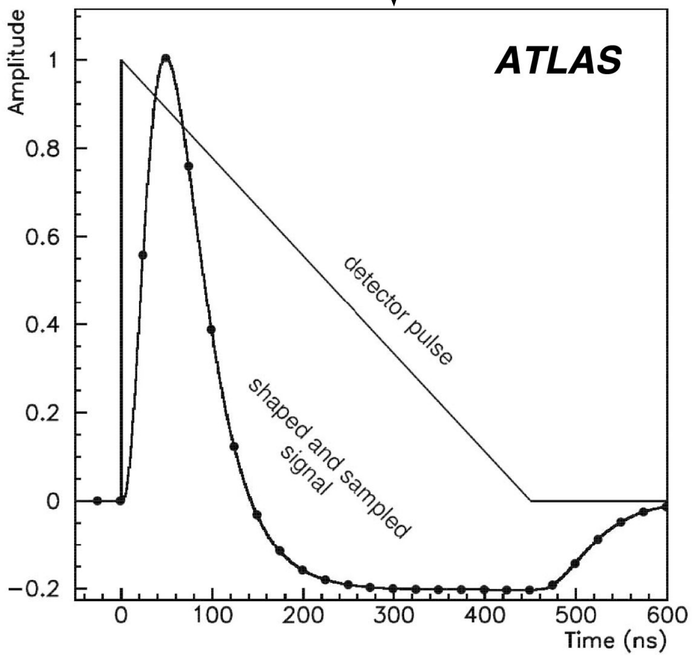
Probing the future

JINST 17 (2022) P05024

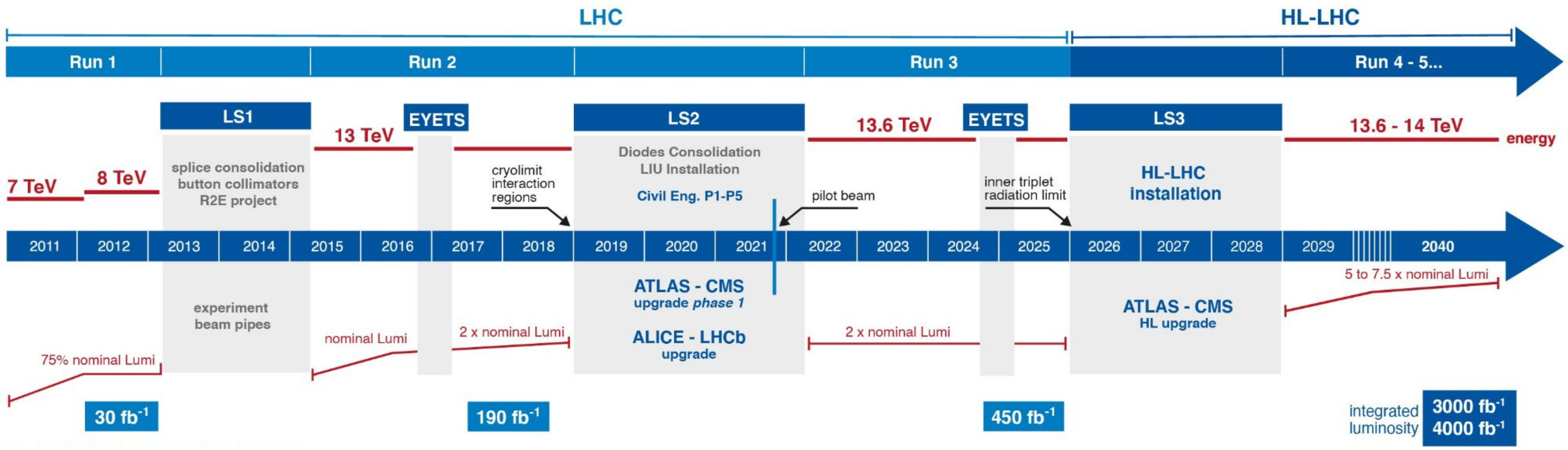
The Liquid Argon calorimeter (LAr)



- Sampling calorimeter with accordion shape
    - Liquid argon as active medium
    - Lead/copper/tungsten as absorber
    - 180k channels
- Measure the energy of (mainly) electromagnetic interacting particles
- Trigger capabilities



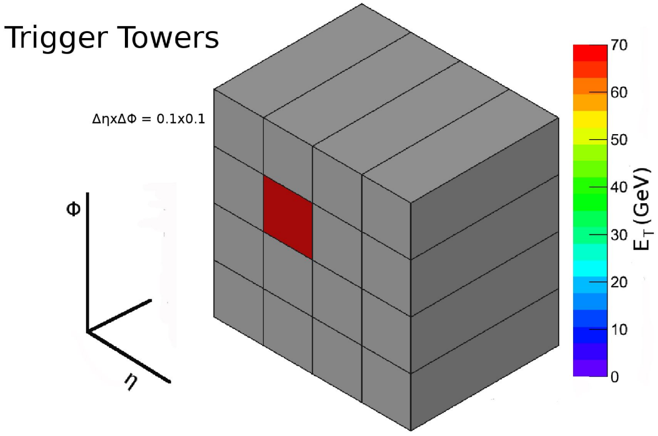
Towards the HL-LHC



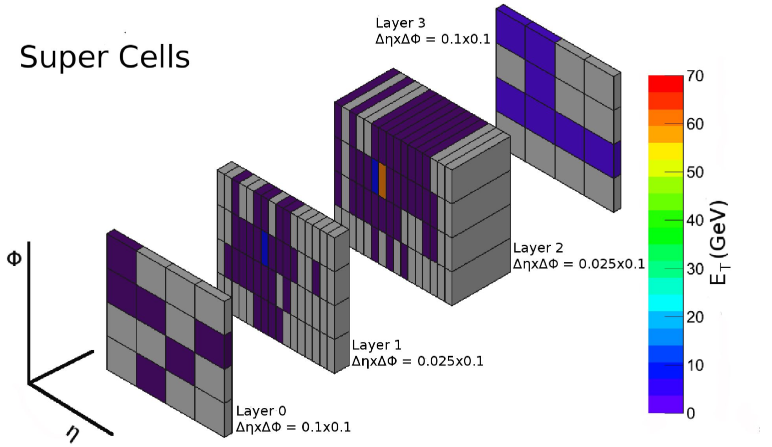
- It seems that new physics is hiding and we do not have enough Higgs bosons
  - So everyone (almost) wants more data
  - Lets increase the luminosity → increased pileup
- Phase-I upgrade of LAr:
  - Maintain trigger performance with higher pileup
  - Replace the trigger path electronics
- Phase-II upgrade of LAr:
  - Move to 1 MHz readout (100 kHz before HL-LHC)
  - Replace the full readout electronics

Phase-I upgrade of the LAr calorimeter

70 GeV electron as seen by the legacy (analog) trigger

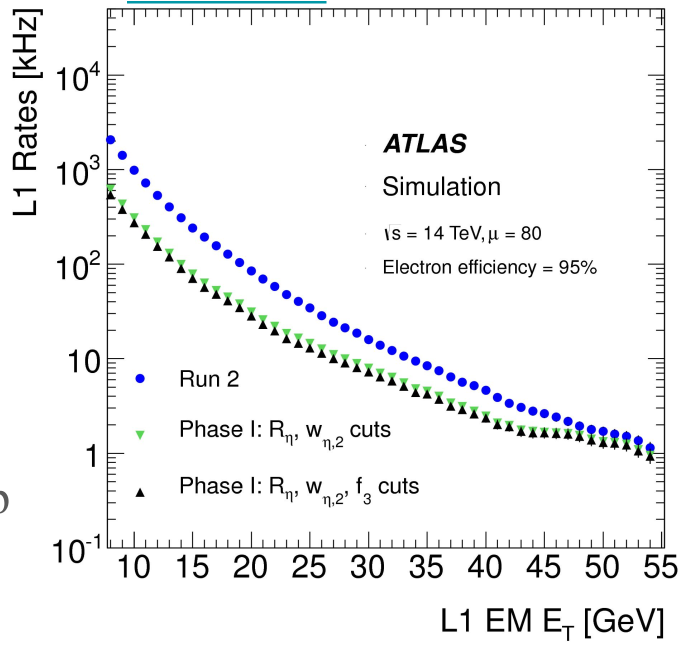


70 GeV electron as seen by the new (digital) trigger

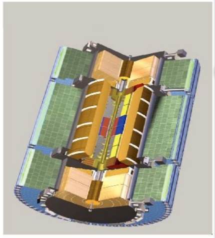


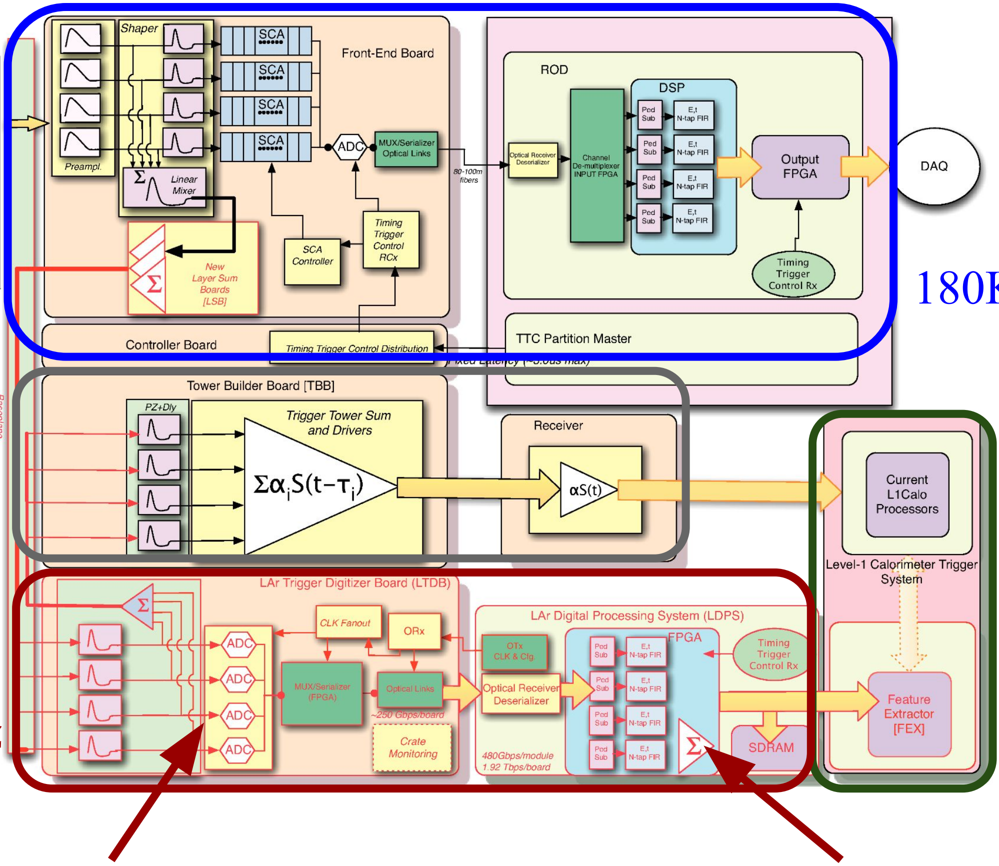
- Coarser granularity at trigger level
  - To handle large bandwidth at 40 MHz
- Digital trigger increases the granularity by factor 10
  - Access to longitudinal shower shape
- Shower shape variables increase background rejection
  - Maintain trigger thresholds and rates with higher pileup

ATLAS-TDR-022



Digital Trigger electronics





Readout system
180K channels @100 kHz

Legacy trigger chain
4K channels @40 MHz

L1 Trigger system

Digital trigger chain
34K channels @40 MHz

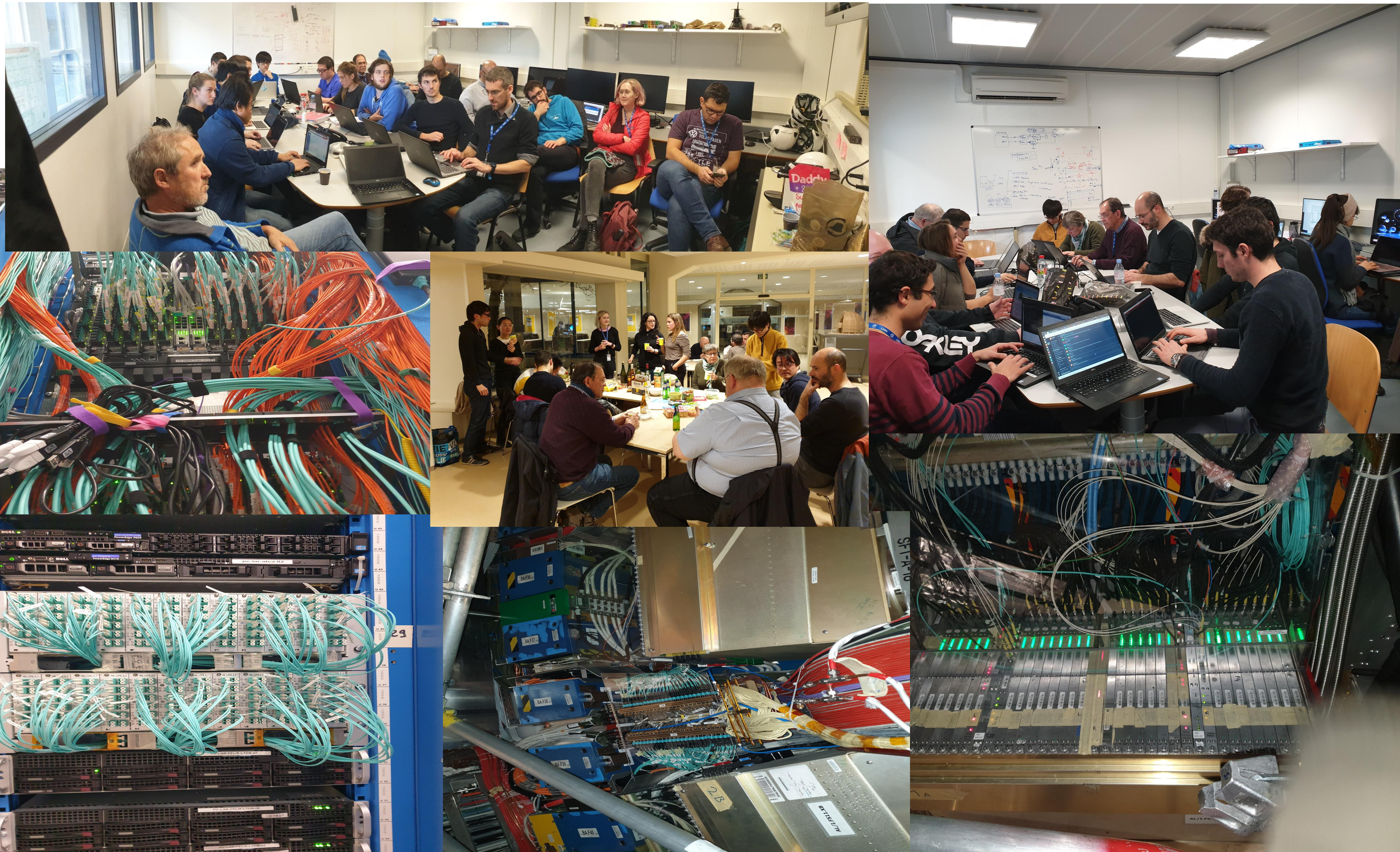
New frontend board
Digitize LAr pulse and send it to backend

New backend board
Compute the energy and send it to trigger

New firmware, online software, detector control system, offline software, and data analysis

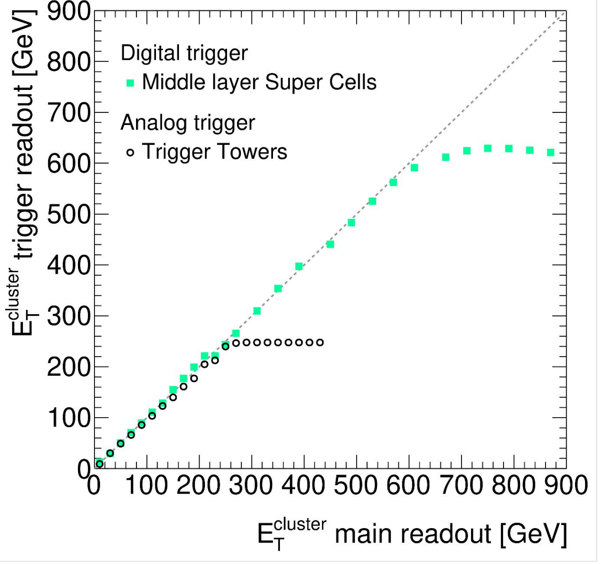
Commissioning of the Digital Trigger system

4 years of work and fun (and COVID)

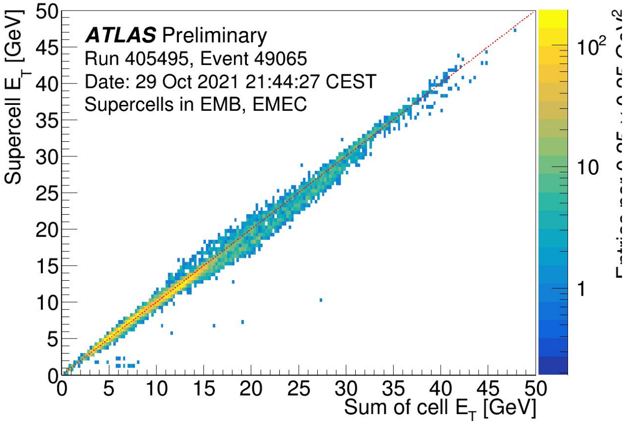


And after all this hard work

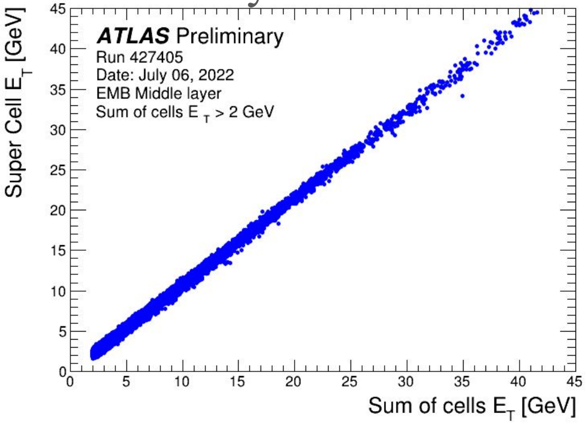
Calibration data



Pilot run data

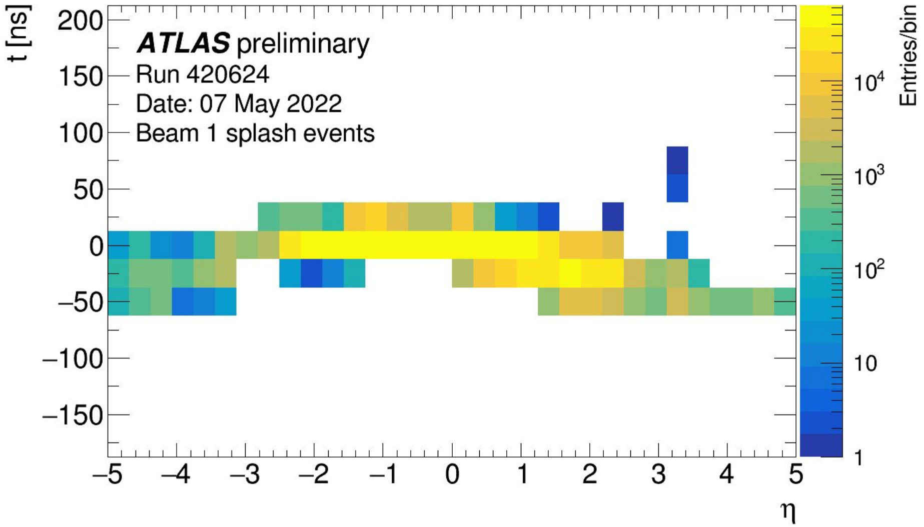


Early run 3 data



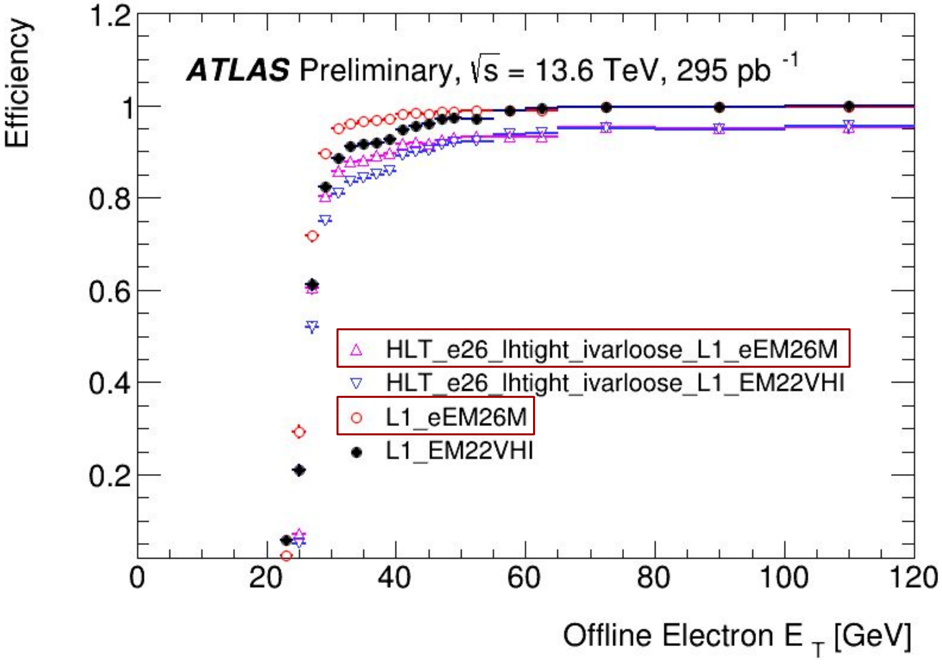
- Digital trigger system ready for the first LHC run 3 data
- Detector timing adjusted starting the first beam splash events
- Energy computation compared to well calibrated main readout system
  - Prepared with calibration data
  - Refined with collisions

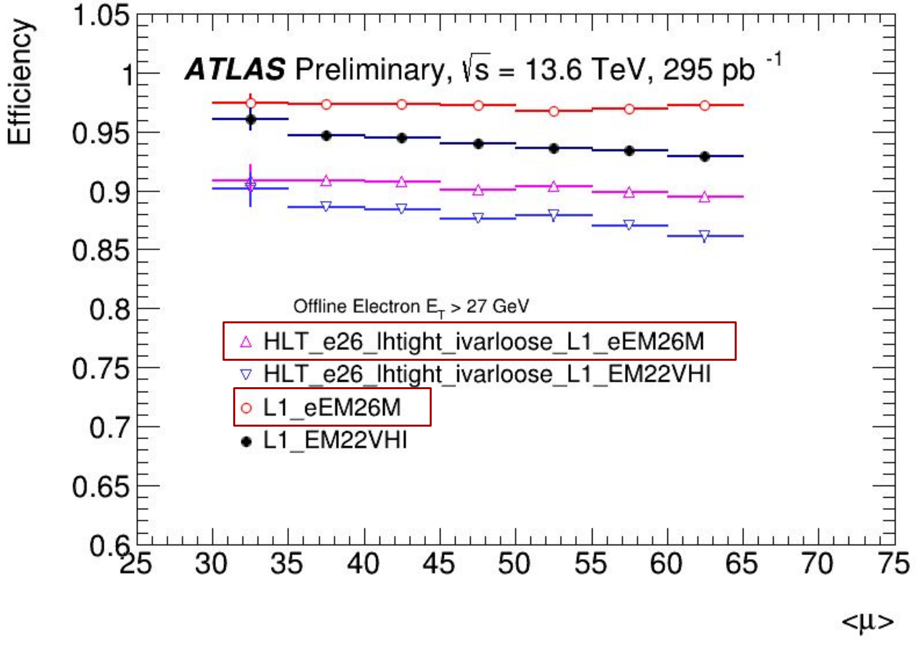
Beam splash events



Performance of the Digital Trigger system

- Digital trigger outperform the legacy trigger for egamma objects
    - Higher efficiency at lower energy
    - Stable efficiency with increased pileup
- Switched to new egamma trigger objects in 2023
    - The rest of the trigger objects will follow in 2024





And now a bit of the Phase-II upgrade

Probing new ideas

Comput Softw Big Sci **5**, 19 (2021)

JINST ***18*** (2023) P05017

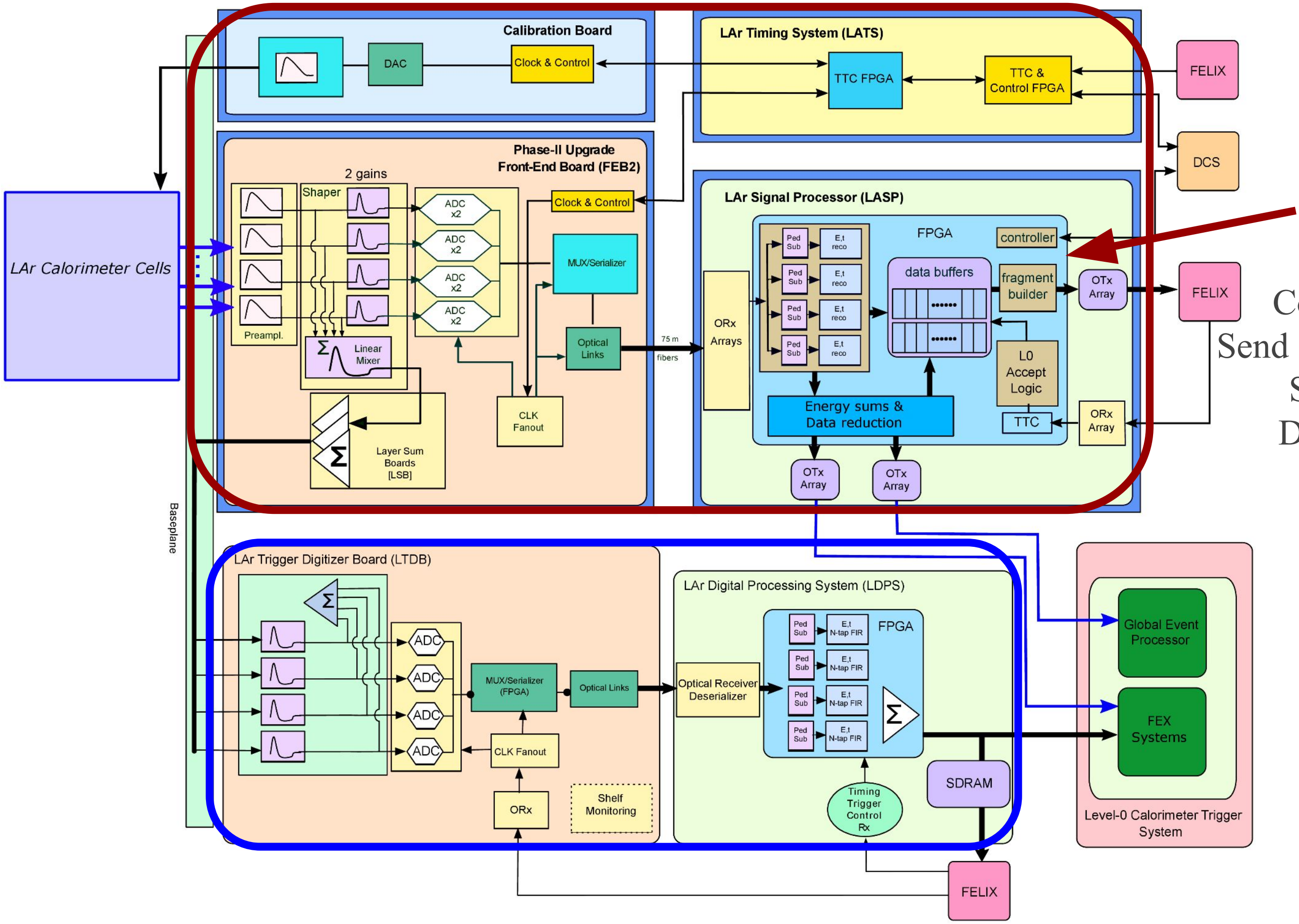
Etienne Fortin, PhD thesis (2022)

Lauri Laatu, PhD thesis (2023)

Nemer Chiedde, PhD thesis (2023)

Phase-II Upgrade electronics

Complete replacement of the readout system



New backend readout board (LASP)

Compute energy deposited in LAr

Send a selection of channels to the trigger

Send all channels to the readout

Designed and produced at CPPM

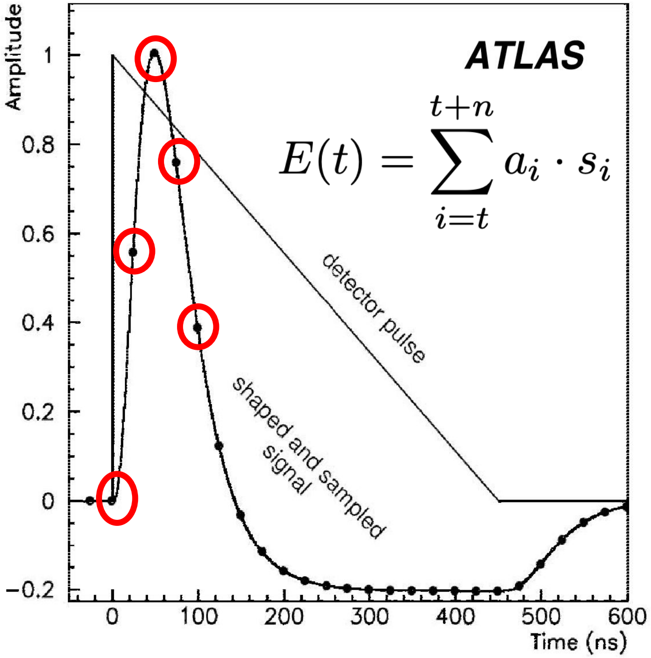
Based on INTEL FPGAs

Large FPGAs that can handle neural network computations

Phase-I digital trigger system remains operational for Phase-II

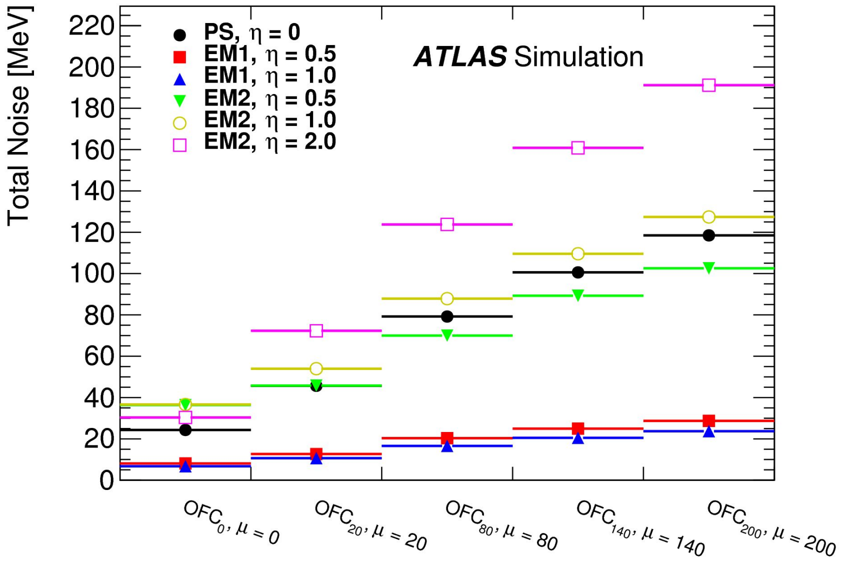
Energy reconstruction at the HL-LHC

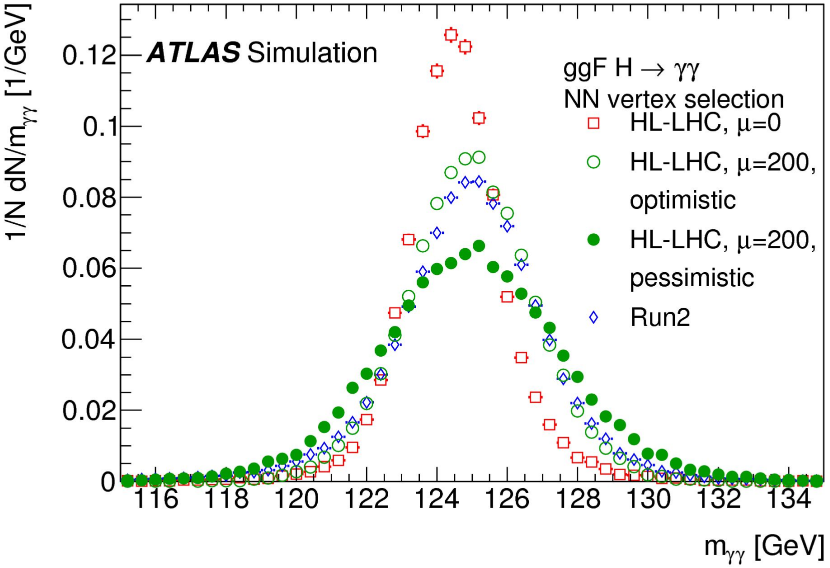
- Energy reconstruction using optimal filtering
    - Weighted sum of sampled pulse amplitudes
- Increased noise due to increased pileup
    - Up to a factor of 2 with respect to Run 3
- About 30% degradation in $m_{\gamma\gamma}$ resolution
    - Better energy reconstruction algorithms needed
    - Neural networks are obvious candidates



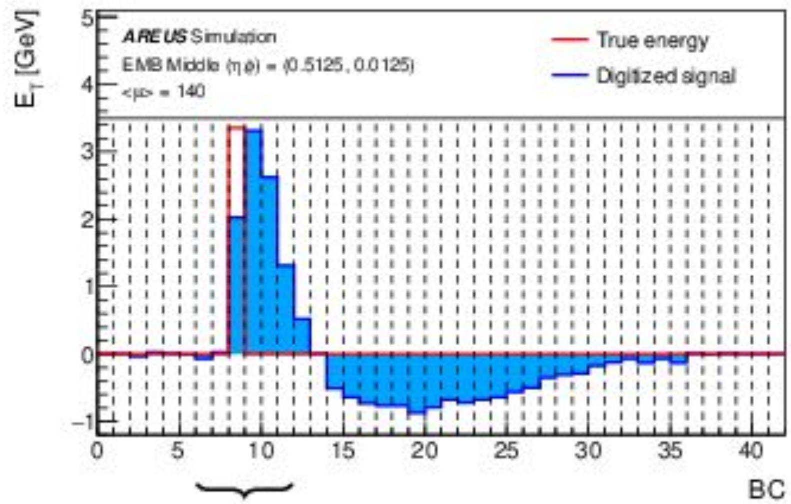
$$E(t) = \sum_{i=t}^{t+n} a_i \cdot s_i$$

ATL-COM-LARG-2017-030

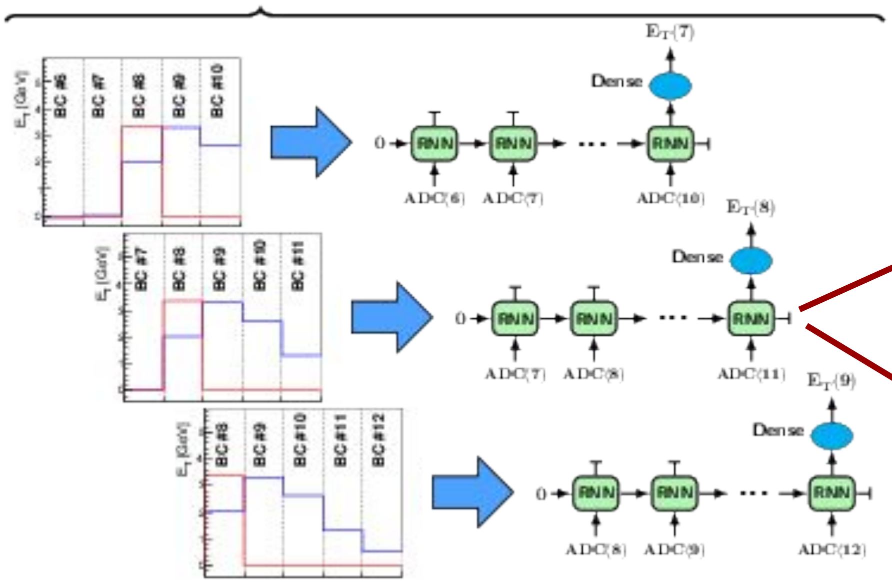




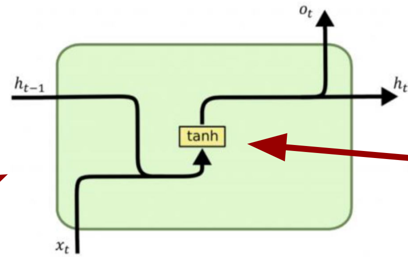
Neural network architecture



- Recursive neural network (RNN) to handle the series of samples
- RNN cell to processes each time step (sample)
- Sliding window architecture
    - ○ 4 samples on the pulse to probe the amplitude
    - ○ N samples in the past to probe pileup

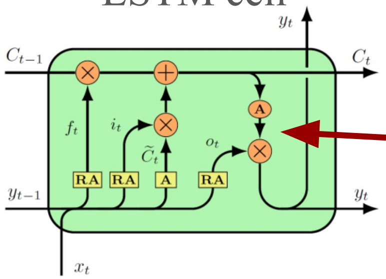


Vanilla RNN cell



One internal neural network

LSTM cell



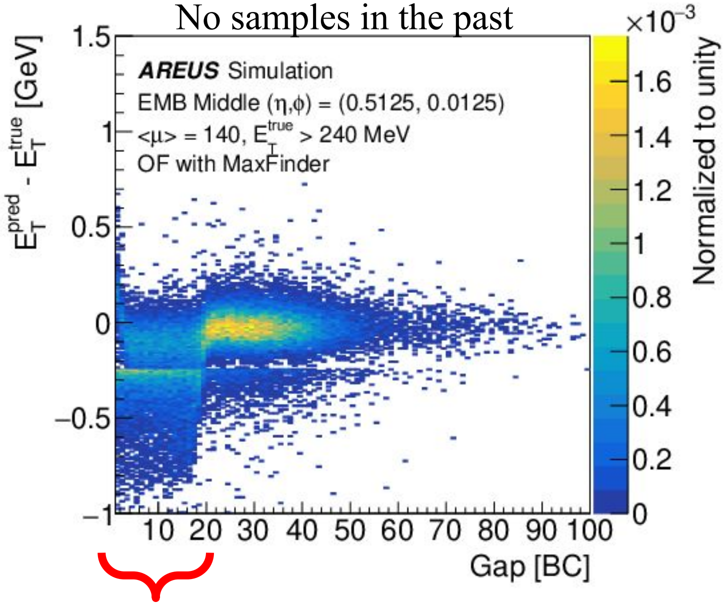
Several internal neural networks with dedicated gates

Performance as function to time gap

- Dramatic degradation of performance when pulses overlap (out-of-time pileup)
    - Time gap of less than ~ 20 BC
- Neural networks recover the performance
    - Better performance with increased sequence length to cover past events

Optimal Filtering

No samples in the past

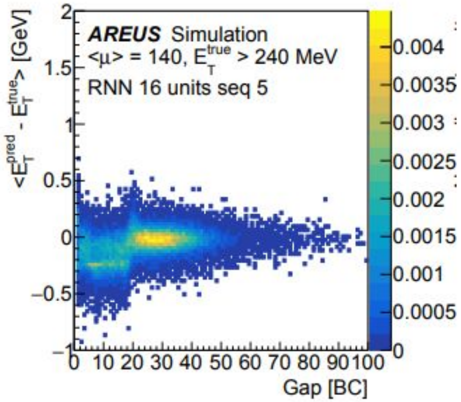


Overlapping signals region

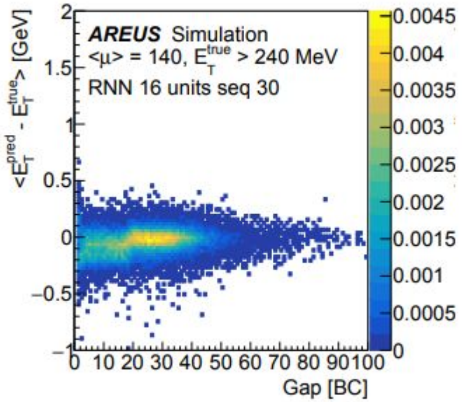
Vanilla RNN

LSTM

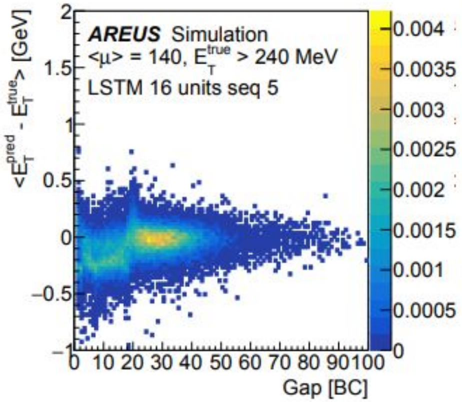
One sample in the past



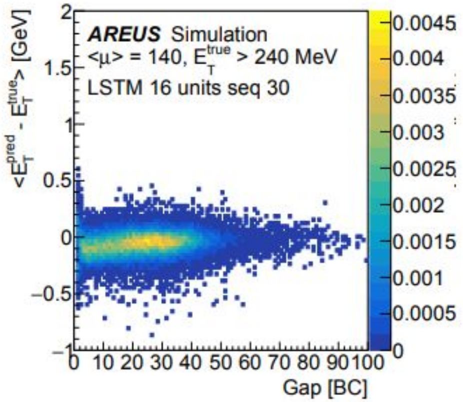
26 samples in the past



One sample in the past

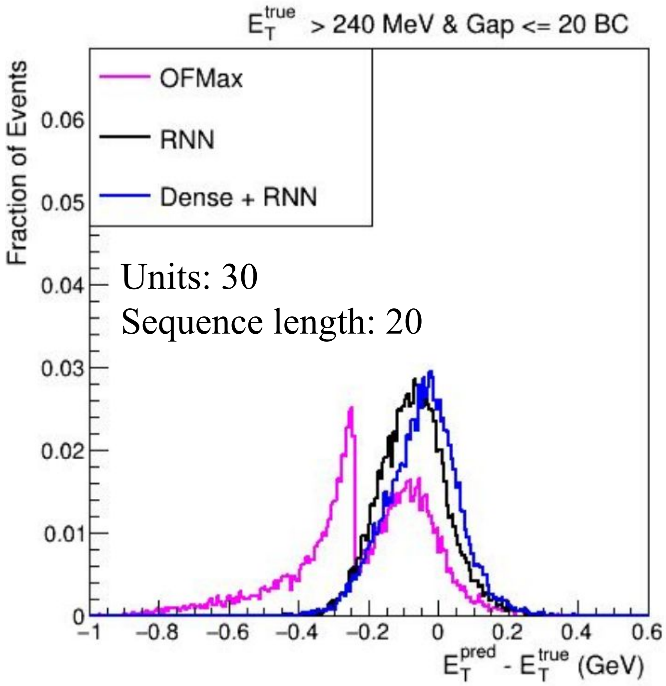


26 samples in the past



RNN optimisations

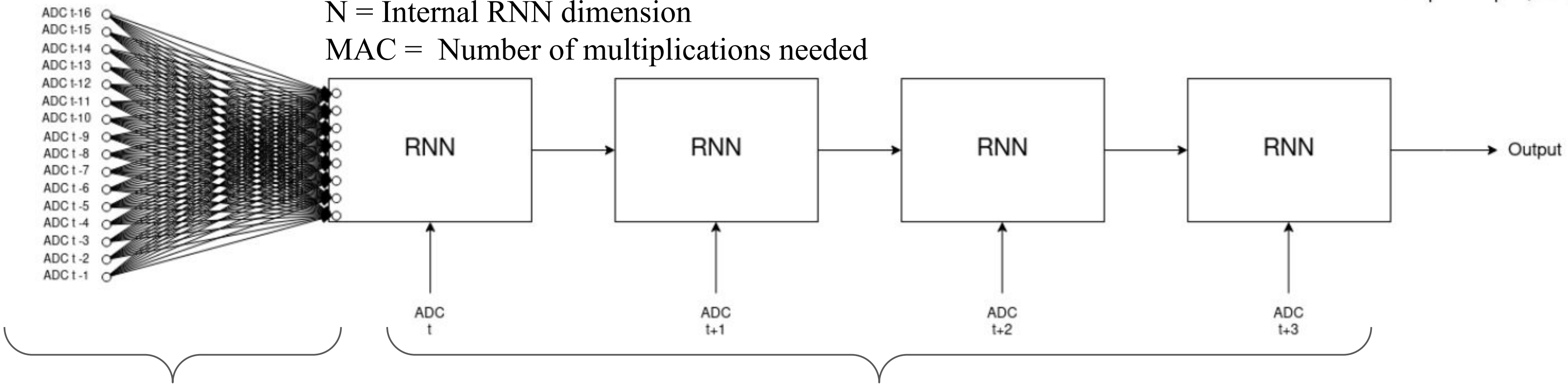
- Long sequences needed to efficiently correct for pileup
- Number of RNN cells scales with the sequence length
- RNN cells need significant processing resources
- Use dense layer to acquire pileup correction
- Reduce the number of multiplication by a factor 4
    - For a network internal dimension of 30
- No effect on performance



S = Number of samples

N = Internal RNN dimension

MAC = Number of multiplications needed

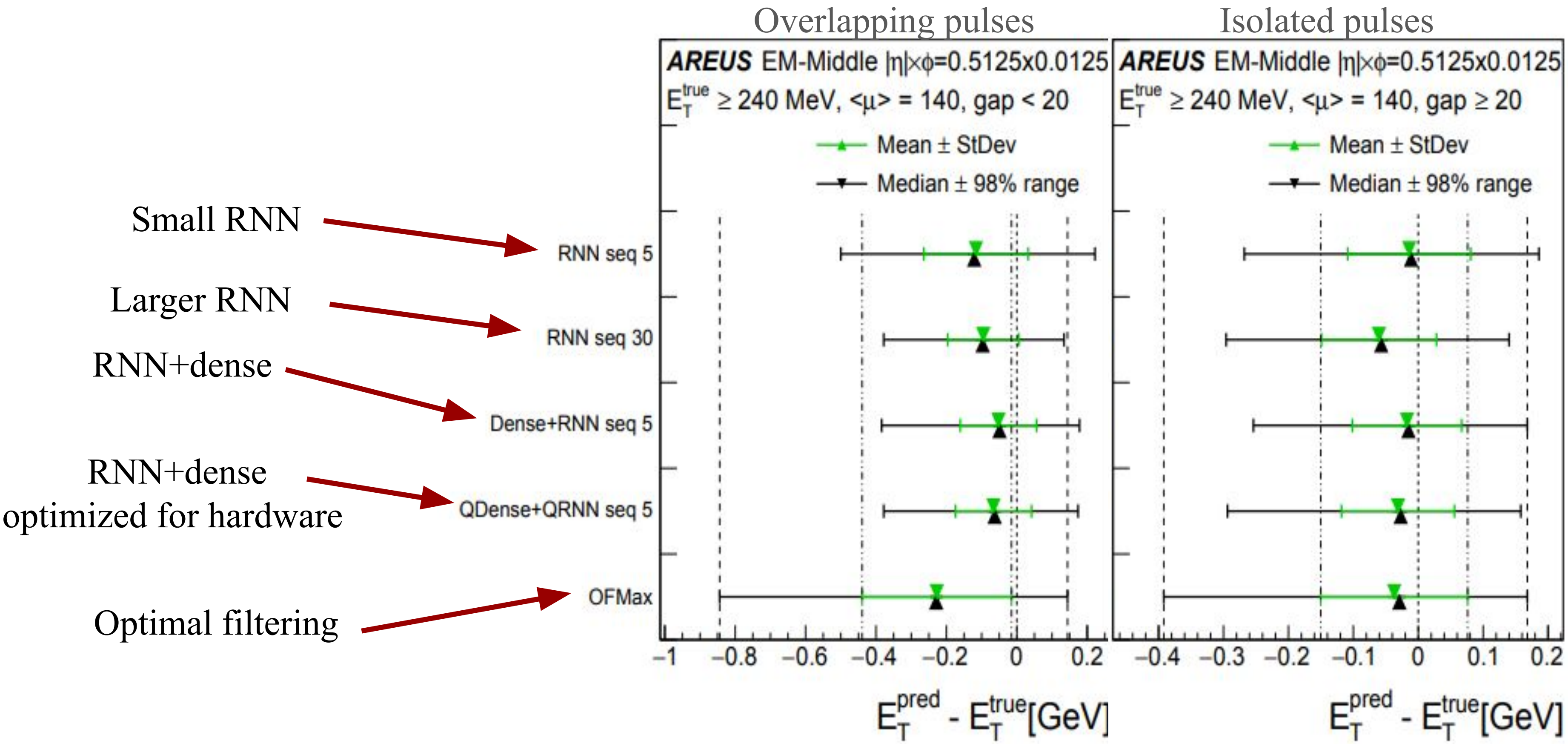


Dense layer : MAC $\propto$ S$\times$N

RNN layers : MAC $\propto$ S$\times$N^2

RNN performance (summary)

- Small RNNs (sequence length 5) can outperform OFMax overall
  - But not in all regions
  - Larger networks needed
- Several optimisation carried out to improve the performance
  - Keeping the network suitable for FPGA processing



Overlapping pulses

Isolated pulses

Small RNN

Larger RNN

RNN+dense

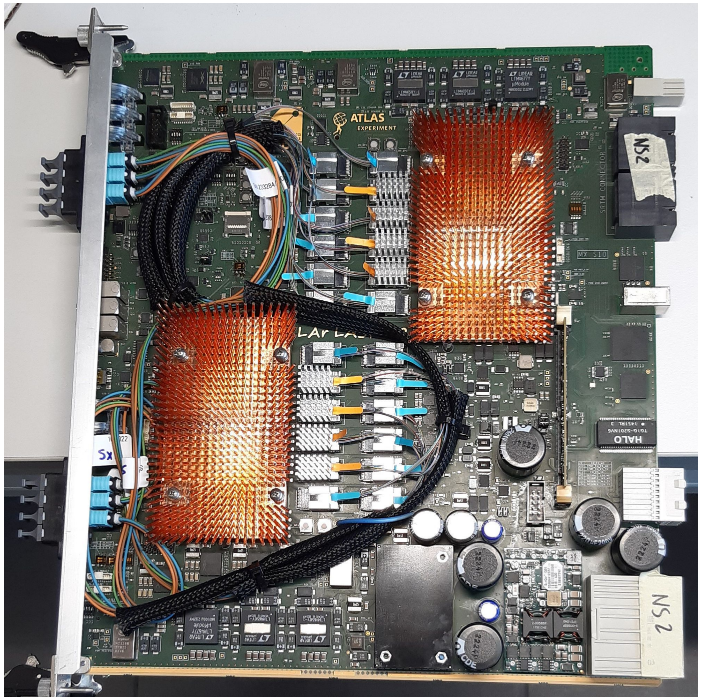
RNN+dense
optimized for hardware

Optimal filtering

Firmware implementation

- Small RNN implemented on Stratix 10 FPGAs
  - Smaller FPGA on a LASP demonstrator board
  - Larger RNN to be implemented on AGILEX FPGAs
- Challenges:
  - 384 channels per FPGA, 125 ns latency
- Optimisations carried out in HLS and VHDL
  - Multiplexing: Reuse of same logic for several channels
  - Quantisation of mathematical operations
  - Architecture of matrix multiplications
- Demonstrated feasibility of running RNNs on the LASP
  - Firmware fits all requirements and tested on the hardware

LASP demonstrator board



*based on experience with the phase-I upgrade

	N networks x multiplexing	ALM	DSP	FMax	latency
Target	**384 channels**	**30%***	**70%***	**Multiplexing x 40 MHz**	**125 ns**
“Naive” HLS	384x1	226%	529%	-	322 ns
HLS optimized	37x10	90%	100%	393 MHz	277 ns
VHDL optimized	28x14	18%	66%	561 MHz	116 ns

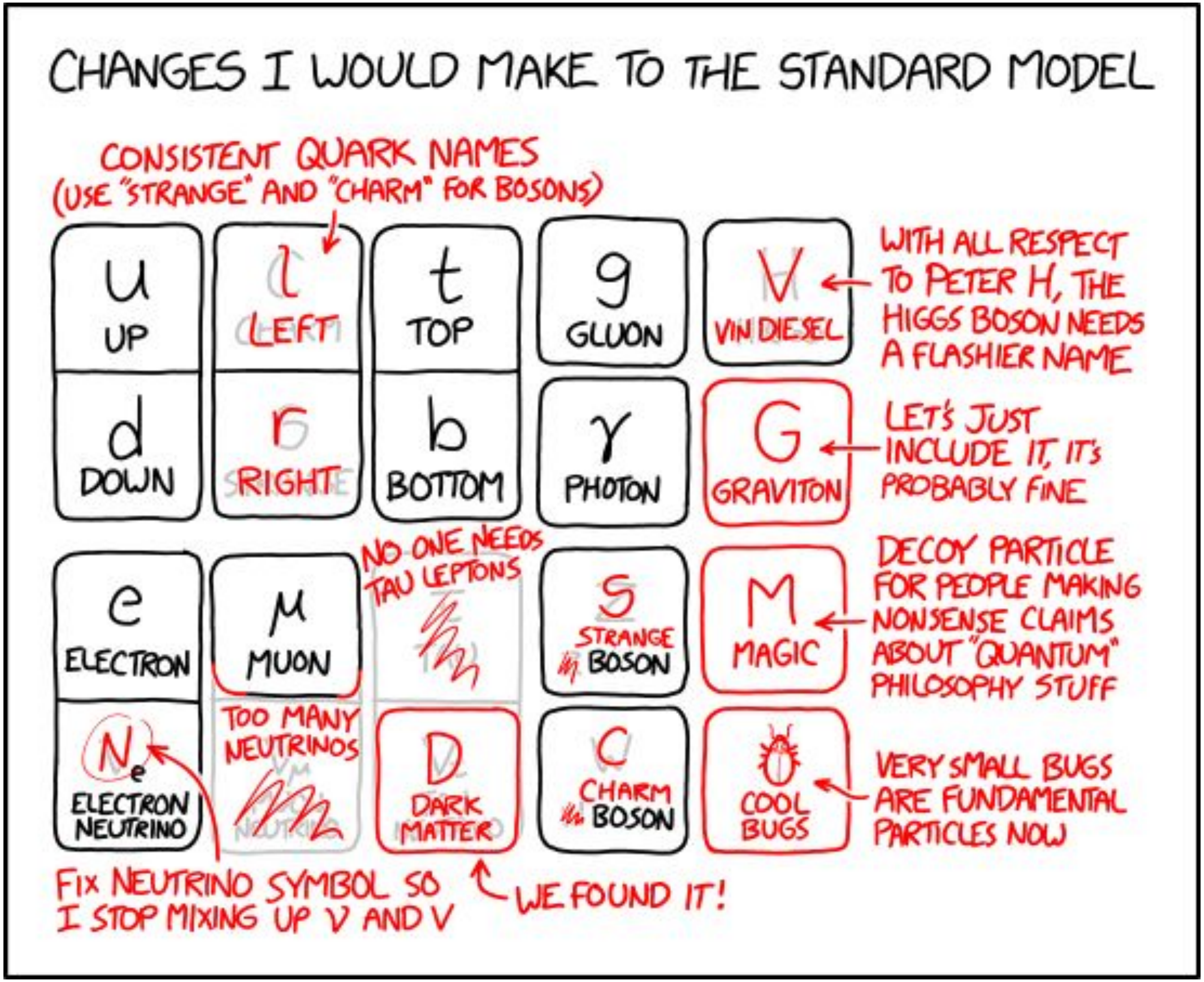
Conclusion(s)

Probing questions

Conclusion

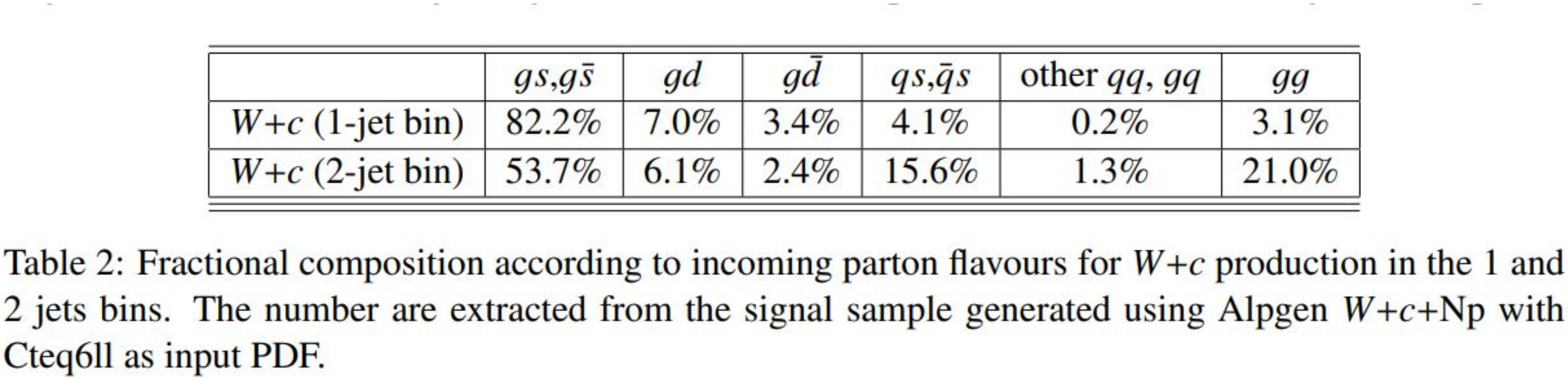
- Wc analysis
  - PDFs are important at the LHC (surprise surprise)
  - Wc analysis can directly probe the badely constrained s-PDF
  - ATLAS data compatible with light-sea quark PDF SU(3) symmetry
- ttH(bb) analysis
  - Coupling to third generation quarks established in run 2
  - ttH(bb) had a (small) contribution to both ttH and H→bb discovery
  - Complex channel with large systematic uncertainties from ttbb background
  - CMS did better (but they forgot few systematics ;-))
- Phase-I upgrade of LAr
  - The new digital trigger system was painfly born (but it was a lot of fun)
  - New system outperform the legacy trigger and is more resilient to pileup
- Phase-II upgrade of LAr
  - Finding new ideas to heat-up the new backend board that is designed at CPPM
  - Neural networks can improve the energy reconstruction in high pileup conditions
  - First neural networks implemented in hardware and fit the specifications
  - Yet to be proven in more realistic (physics) conditions

Backup



https://xkcd.com/2351/

Wc: PDF contributions to Wc process



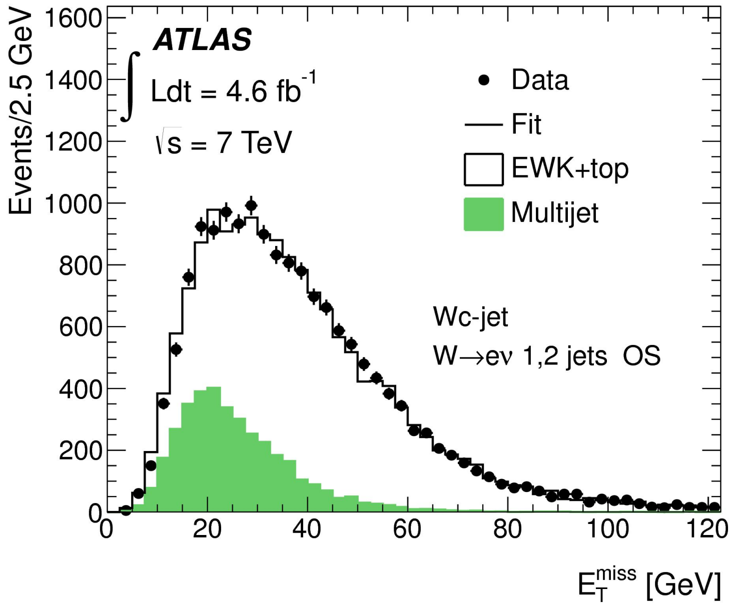
	gs,$g\bar{s}$	gd	$g\bar{d}$	qs,$\bar{q}s$	other qq, gq	gg
W+c (1-jet bin)	82.2%	7.0%	3.4%	4.1%	0.2%	3.1%
W+c (2-jet bin)	53.7%	6.1%	2.4%	15.6%	1.3%	21.0%

Table 2: Fractional composition according to incoming parton flavours for W+c production in the 1 and 2 jets bins. The number are extracted from the signal sample generated using Alpgen W+c+Np with Cteq6ll as input PDF.

Wc: Background estimation electron channel

$$N_{\mathrm{bkg}}^{\mathrm{OS-SS}} = A_{\mathrm{bkg}} \cdot N_{\mathrm{bkg}}^{\mathrm{OS+SS}} = \frac{2 \cdot A_{\mathrm{bkg}}}{1 - A_{\mathrm{bkg}}} N_{\mathrm{bkg}}^{\mathrm{SS}} \quad \Leftarrow \quad A_{\mathrm{bkg}} = N_{\mathrm{bkg}}^{\mathrm{OS-SS}} / N_{\mathrm{bkg}}^{\mathrm{OS+SS}}$$

- Multijet Asymmetry
  - Fraction fit to the MET distribution in OS and SS
- W+jets Asymmetry
  - MC corrected by a data driven factor
  - Extracted from the asymmetry of all tracks in the jet (pretag regions)
- Normalisation
  - Multijet and W+jets normalised in the SS region with a chi2 fit



$$A_{W+\mathrm{light}} = A_{W+\mathrm{light}}^{\mathrm{MC}} \frac{A_{W+\mathrm{light}}^{\mathrm{data,tracks}}}{A_{W+\mathrm{light}}^{\mathrm{MC,tracks}}}$$

Wc: Background estimation muon channel

Matrix Method

SMT rate as function of muon isolation
Extrapolated to signal region

$$N^{\mathrm{OS+SS}}_{\mathrm{multijet}} = N^{\mathrm{pretag}}_{\mathrm{multijet}} \cdot R^{\mathrm{SMT}}_{\mathrm{multijet}}$$

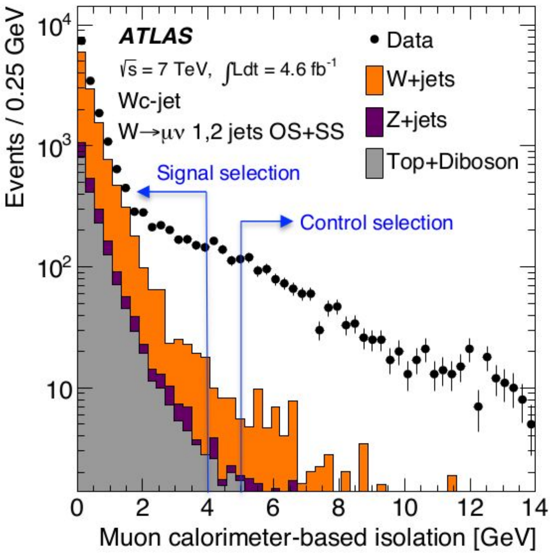
Data pretag - MC

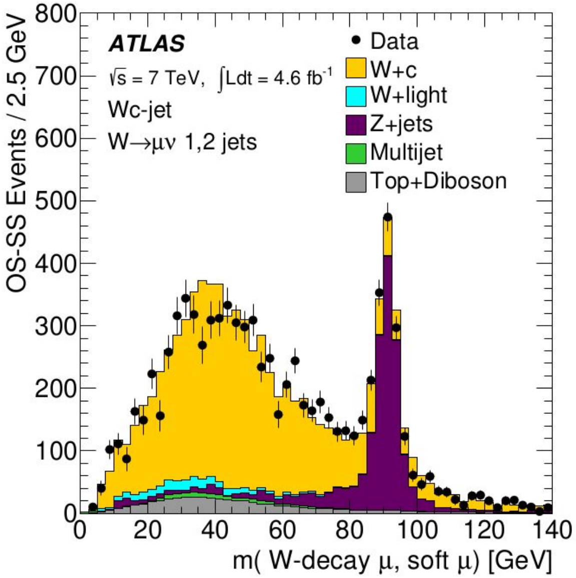
System of equation
using a lifetime tagger

SMT scale factor from
system of equation with
lifetime tagger

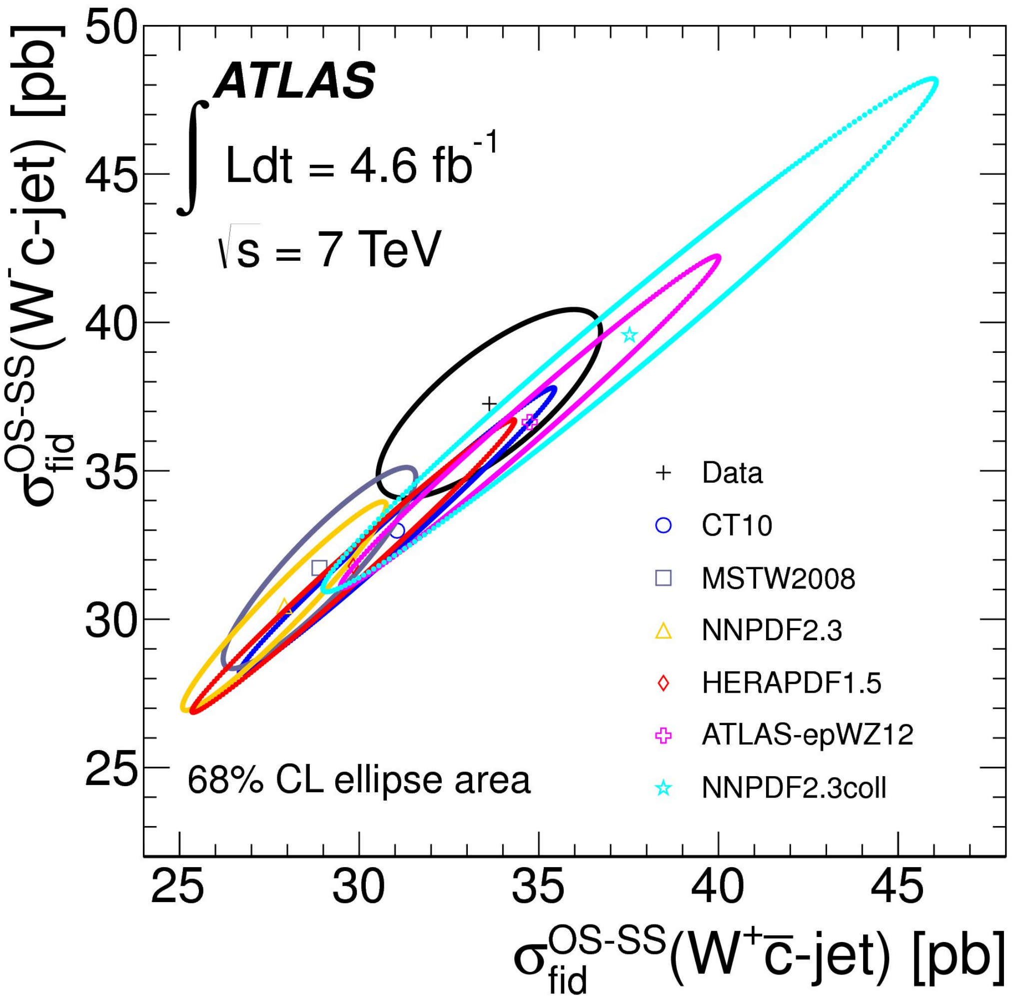
$$N^{\mathrm{OS+SS}}_{W+\mathrm{light}} = N^{\mathrm{pretag}}_{W+\mathrm{jets}} \cdot f_{\mathrm{light}} \cdot R^{\mathrm{SMT}}_{W+\mathrm{light}}$$

- Multijet Asymmetry from control region
- W+jets Asymmetry similar to electron channel
- Z+jets OS-SS normalised under the Z peak

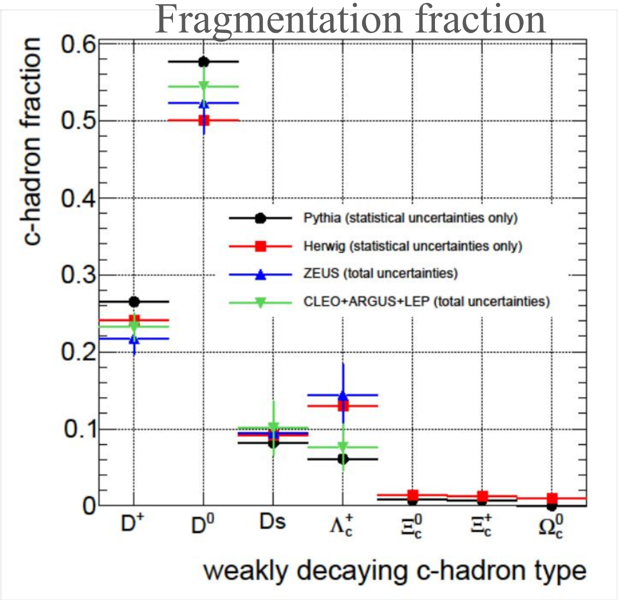


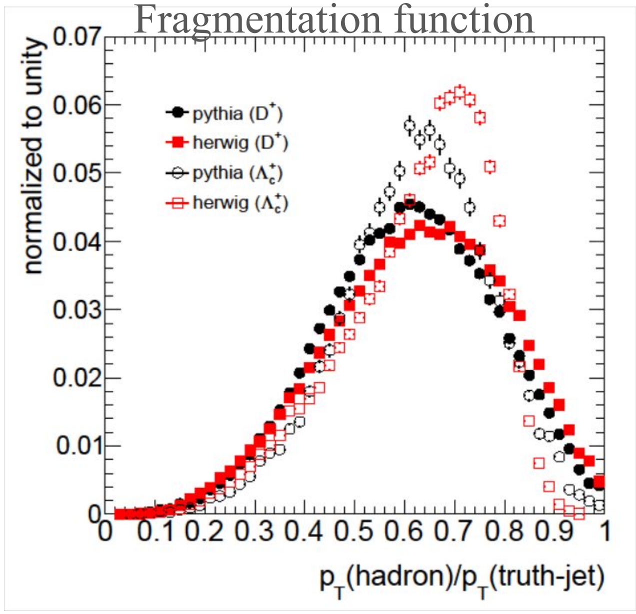


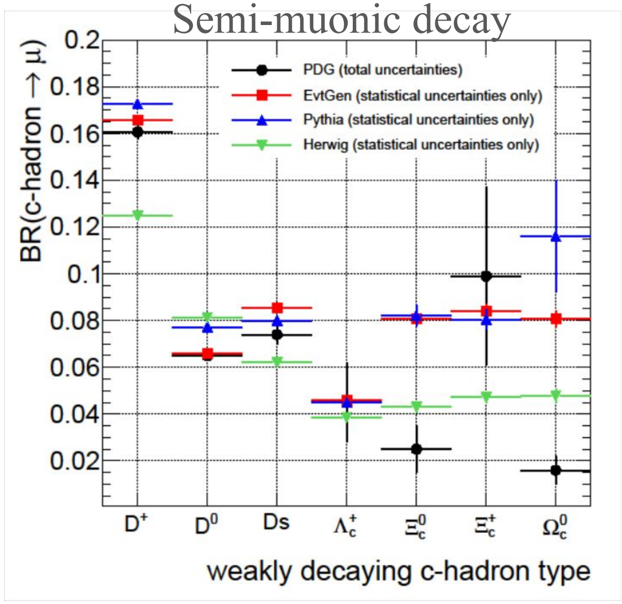
Wc: Correlations between W^+ and W^-

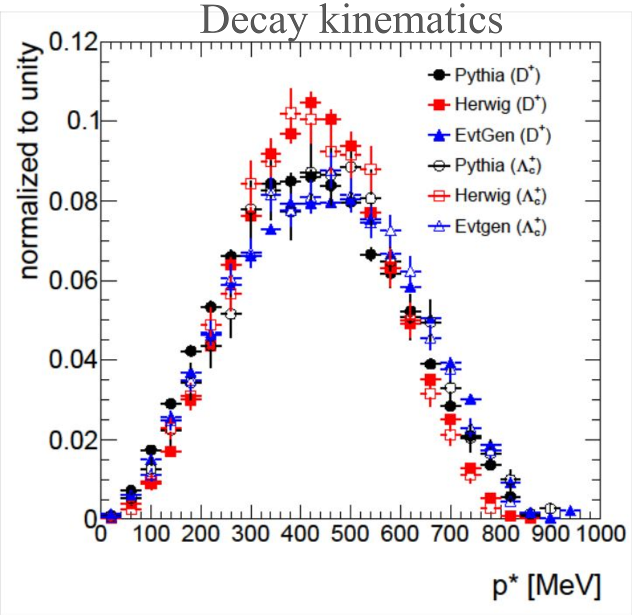


Wc: c fragmentation and decay reweighting


Fragmentation fraction

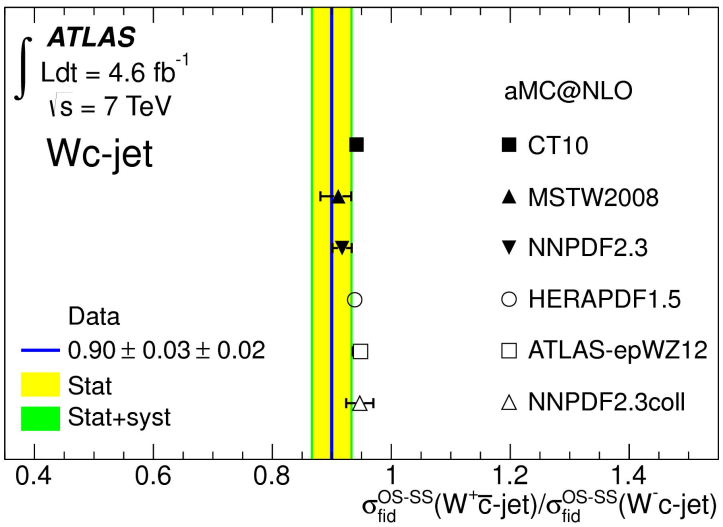

Fragmentation function


Semi-muonic decay


Decay kinematics

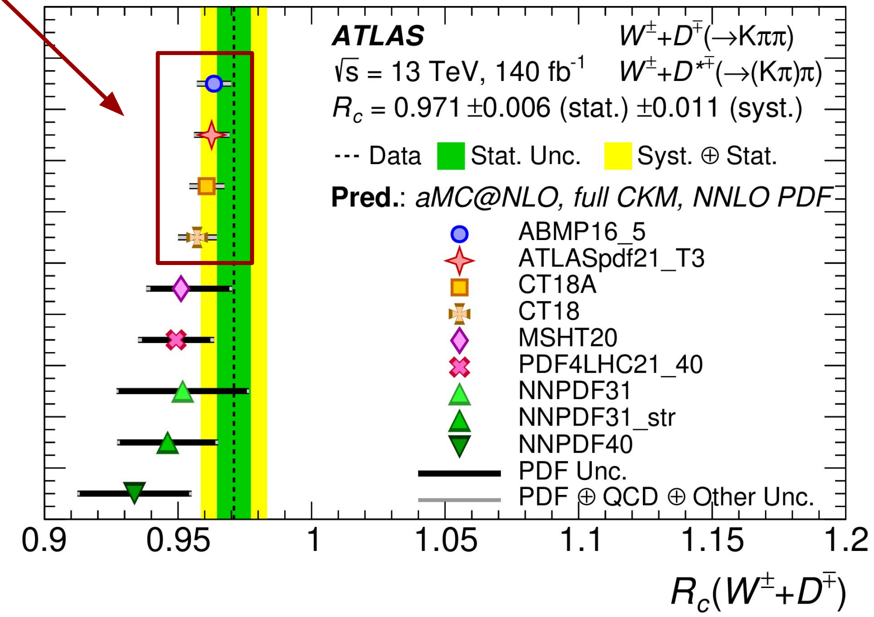
Wc: W^+/W^- compatibility with recent results

- Recent Wc measurement (13 TeV) compatible with PDFs with $s/\bar{s}$ symmetry
- Hard to compare exactly with 7 TeV measurement
  - Not the same centre-of-mass energy to compare directly the results
  - Not the same PDF sets used to compare with simulations

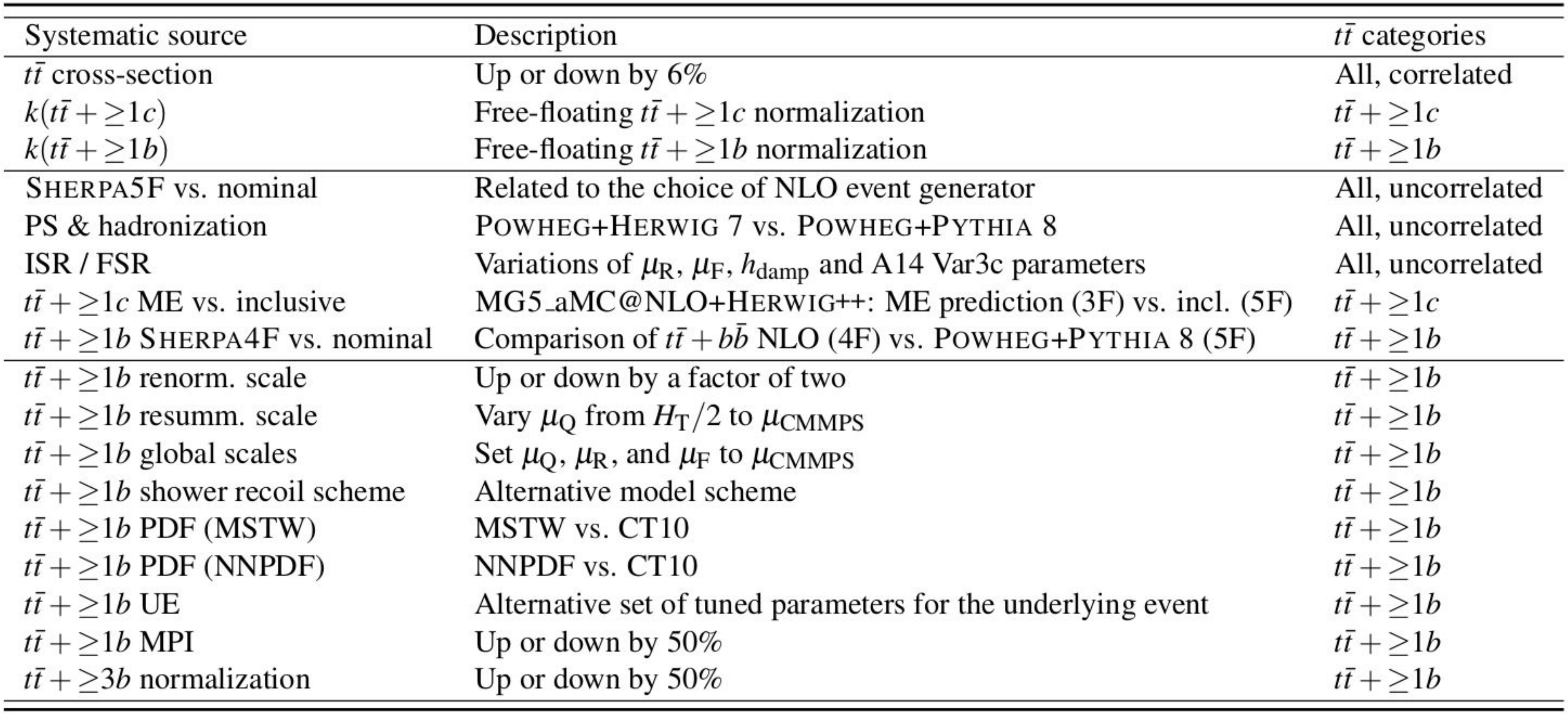


$s/\bar{s}$ symmetry imposed

$$R_c^{\pm} \equiv \sigma(W^+ + \bar{c})/\sigma(W^- + c)$$



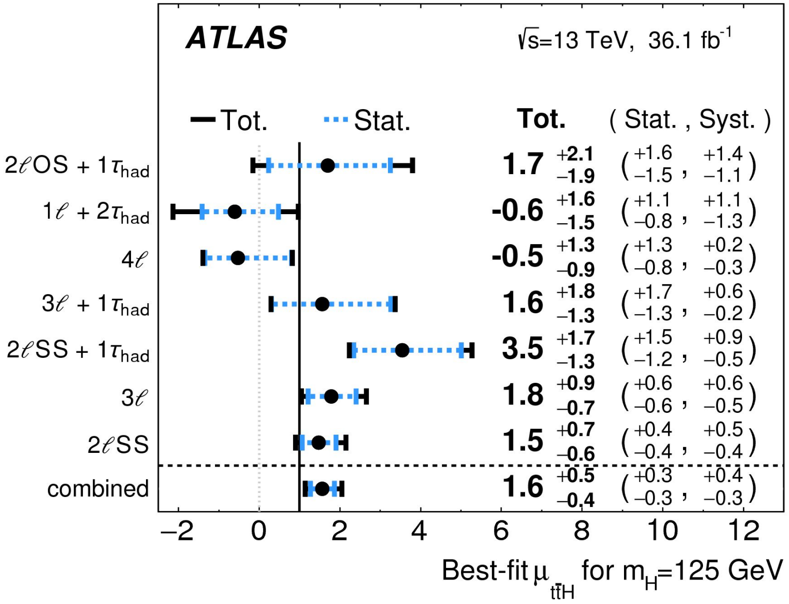
ttH(bb): ttbar systematics

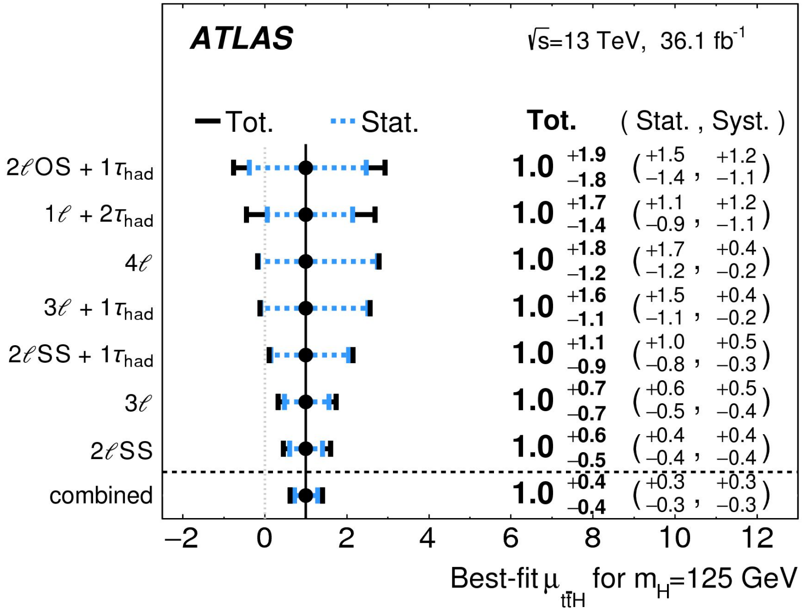


Systematic source	Description	$t\bar{t}$ categories
$t\bar{t}$ cross-section	Up or down by 6%	All, correlated
$k(t\bar{t}+\geq 1c)$	Free-floating $t\bar{t}+\geq 1c$ normalization	$t\bar{t}+\geq 1c$
$k(t\bar{t}+\geq 1b)$	Free-floating $t\bar{t}+\geq 1b$ normalization	$t\bar{t}+\geq 1b$
SHERPA5F vs. nominal	Related to the choice of NLO event generator	All, uncorrelated
PS & hadronization	POWHEG+HERWIG 7 vs. POWHEG+PYTHIA 8	All, uncorrelated
ISR / FSR	Variations of μ_R, μ_F, h_{damp} and A14 Var3c parameters	All, uncorrelated
$t\bar{t}+\geq 1c$ ME vs. inclusive	MG5_aMC@NLO+HERWIG++: ME prediction (3F) vs. incl. (5F)	$t\bar{t}+\geq 1c$
$t\bar{t}+\geq 1b$ SHERPA4F vs. nominal	Comparison of $t\bar{t}+b\bar{b}$ NLO (4F) vs. POWHEG+PYTHIA 8 (5F)	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ renorm. scale	Up or down by a factor of two	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ resumm. scale	Vary μ_Q from $H_T/2$ to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ global scales	Set μ_Q, μ_R, and μ_F to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ shower recoil scheme	Alternative model scheme	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF (MSTW)	MSTW vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF (NNPDF)	NNPDF vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ UE	Alternative set of tuned parameters for the underlying event	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ MPI	Up or down by 50%	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 3b$ normalization	Up or down by 50%	$t\bar{t}+\geq 1b$

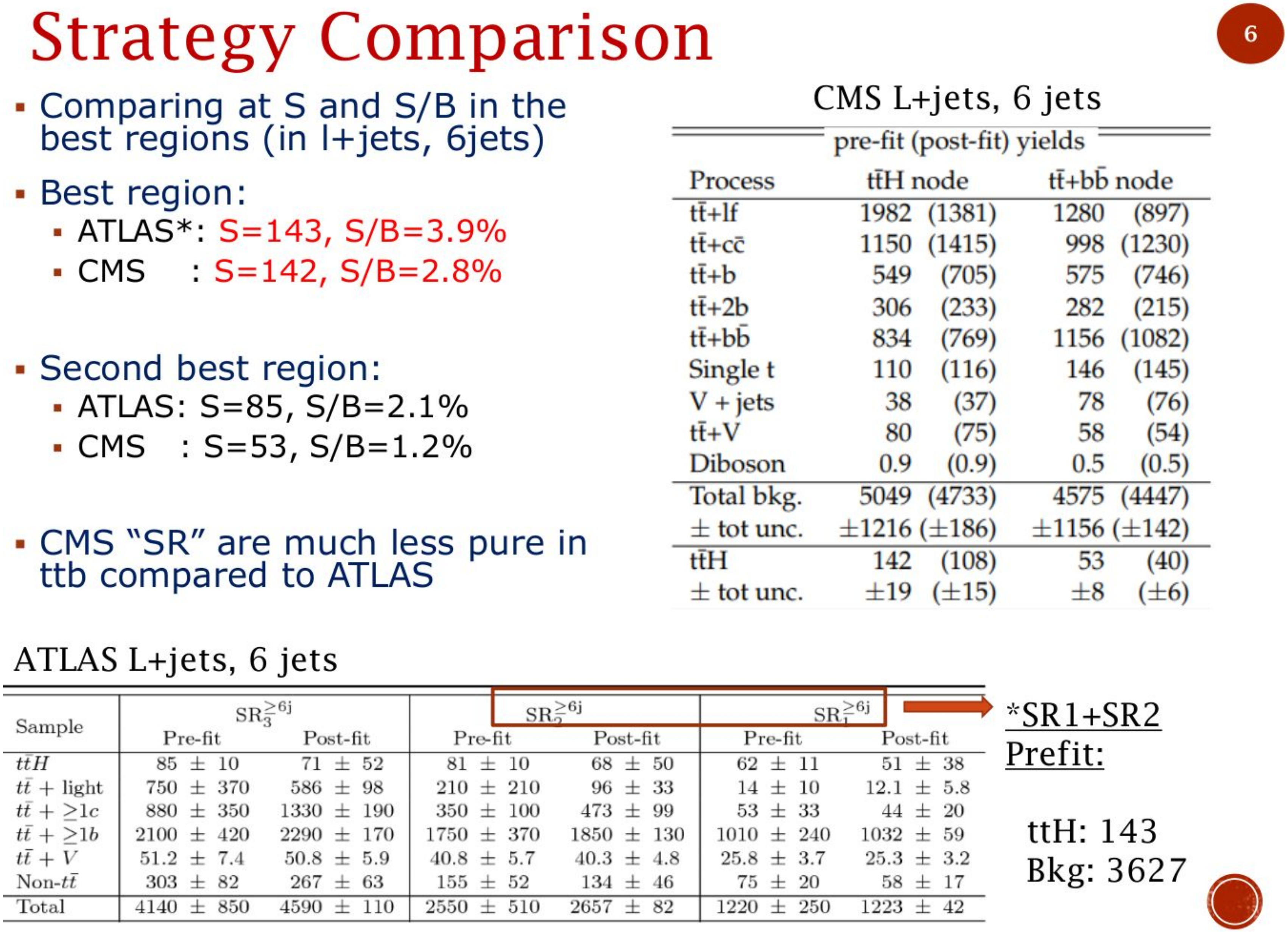
ttH(bb): Combination

Channel	Best-fit μ		Significance	
	Observed	Expected	Observed	Expected
Multilepton	$1.6\,^{+0.5}_{-0.4}$	$1.0\,^{+0.4}_{-0.4}$	4.1σ	2.8σ
$H \to b\bar{b}$	$0.8\,^{+0.6}_{-0.6}$	$1.0\,^{+0.6}_{-0.6}$	1.4σ	1.6σ
$H \to \gamma\gamma$	$0.6\,^{+0.7}_{-0.6}$	$1.0\,^{+0.8}_{-0.6}$	0.9σ	1.7σ
$H \to 4\ell$	< 1.9	$1.0\,^{+3.2}_{-1.0}$	—	0.6σ
Combined	$1.2\,^{+0.3}_{-0.3}$	$1.0\,^{+0.3}_{-0.3}$	4.2σ	3.8σ





ttH(bb): Comparison with CMS



Strategy Comparison

- Comparing at S and S/B in the best regions (in l+jets, 6jets)
- Best region:
  - ATLAS*: S=143, S/B=3.9%
  - CMS : S=142, S/B=2.8%
- Second best region:
  - ATLAS: S=85, S/B=2.1%
  - CMS : S=53, S/B=1.2%
- CMS "SR" are much less pure in ttb compared to ATLAS

CMS L+jets, 6 jets

	pre-fit (post-fit) yields	
Process	$t\bar{t}$H node	$t\bar{t}$+$b\bar{b}$ node
$t\bar{t}$+lf	1982 (1381)	1280 (897)
$t\bar{t}$+$c\bar{c}$	1150 (1415)	998 (1230)
$t\bar{t}$+b	549 (705)	575 (746)
$t\bar{t}$+2b	306 (233)	282 (215)
$t\bar{t}$+$b\bar{b}$	834 (769)	1156 (1082)
Single t	110 (116)	146 (145)
V + jets	38 (37)	78 (76)
$t\bar{t}$+V	80 (75)	58 (54)
Diboson	0.9 (0.9)	0.5 (0.5)
Total bkg.	5049 (4733)	4575 (4447)
± tot unc.	±1216 (±186)	±1156 (±142)
$t\bar{t}$H	142 (108)	53 (40)
± tot unc.	±19 (±15)	±8 (±6)

ATLAS L+jets, 6 jets

Sample	$\mathrm{SR}_3^{\geq 6j}$		$\mathrm{SR}_2^{\geq 6j}$		$\mathrm{SR}_1^{\geq 6j}$	
	Pre-fit	Post-fit	Pre-fit	Post-fit	Pre-fit	Post-fit
$t\bar{t}H$	85 ± 10	71 ± 52	81 ± 10	68 ± 50	62 ± 11	51 ± 38
$t\bar{t}$ + light	750 ± 370	586 ± 98	210 ± 210	96 ± 33	14 ± 10	12.1 ± 5.8
$t\bar{t} + \geq 1c$	880 ± 350	1330 ± 190	350 ± 100	473 ± 99	53 ± 33	44 ± 20
$t\bar{t} + \geq 1b$	2100 ± 420	2290 ± 170	1750 ± 370	1850 ± 130	1010 ± 240	1032 ± 59
$t\bar{t} + V$	51.2 ± 7.4	50.8 ± 5.9	40.8 ± 5.7	40.3 ± 4.8	25.8 ± 3.7	25.3 ± 3.2
Non-$t\bar{t}$	303 ± 82	267 ± 63	155 ± 52	134 ± 46	75 ± 20	58 ± 17
Total	4140 ± 850	4590 ± 110	2550 ± 510	2657 ± 82	1220 ± 250	1223 ± 42

*SR1+SR2

Prefit:

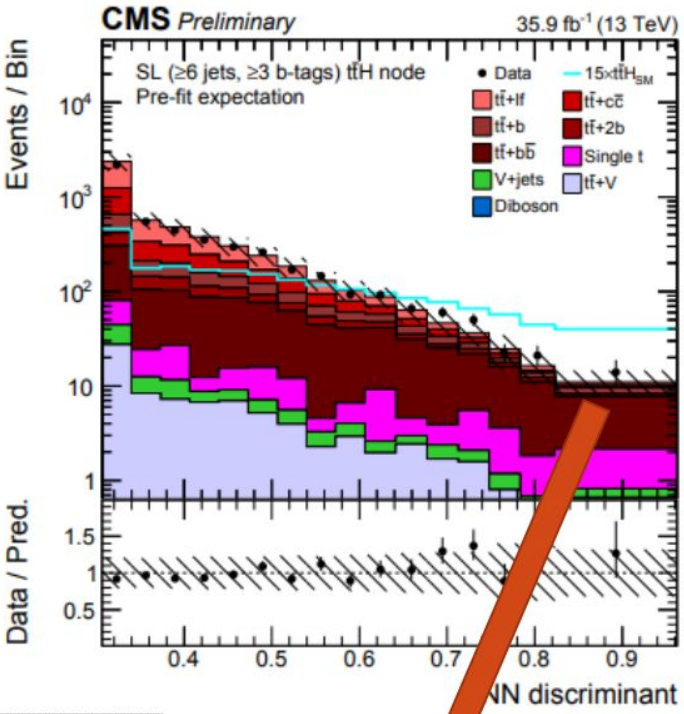
ttH: 143

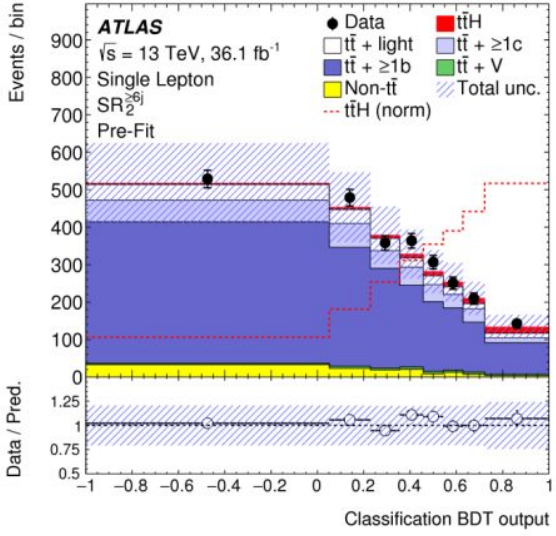
Bkg: 3627

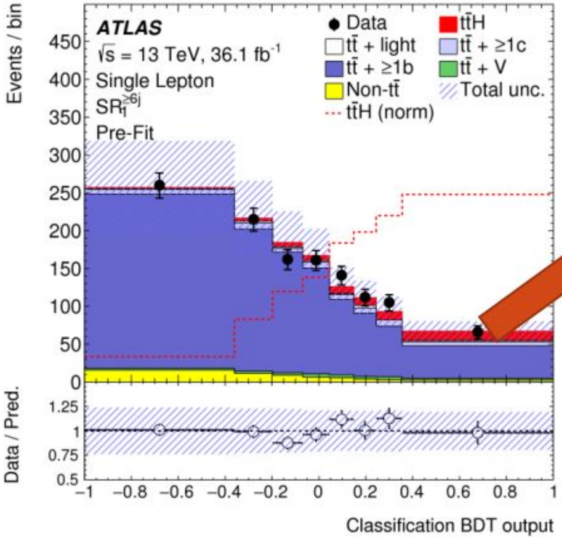
ttH(bb): Comparison with CMS

MVA Comparison

- Check MVA separation in best regions
  - SR1+SR2 for ATLAS
  - 16 bins from both CMS and ATLAS
- Very hard to check by eye
  - Alternatively check purity in last bin
- It seems that ATLAS separation is larger
  - But maybe CMS signal events are distributed differently







S/B=25%, S=2.5

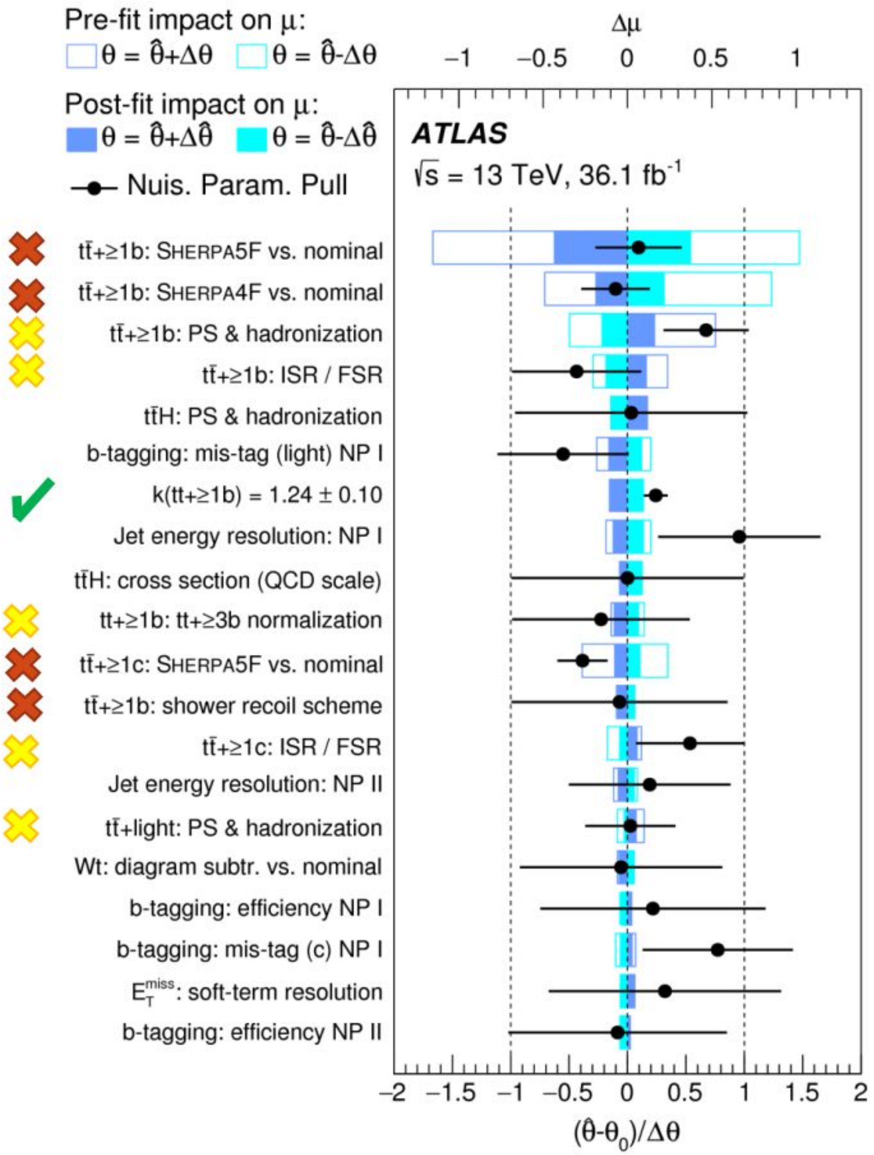
S/B=25%, S=12.5

Disclaimer:
Numbers are
checked by eye

ttH(bb): Comparison with CMS

ttbar syst legend

- ✓ CMS has it
- ✖ CMS partially has it
- ✖ CMS does not have it

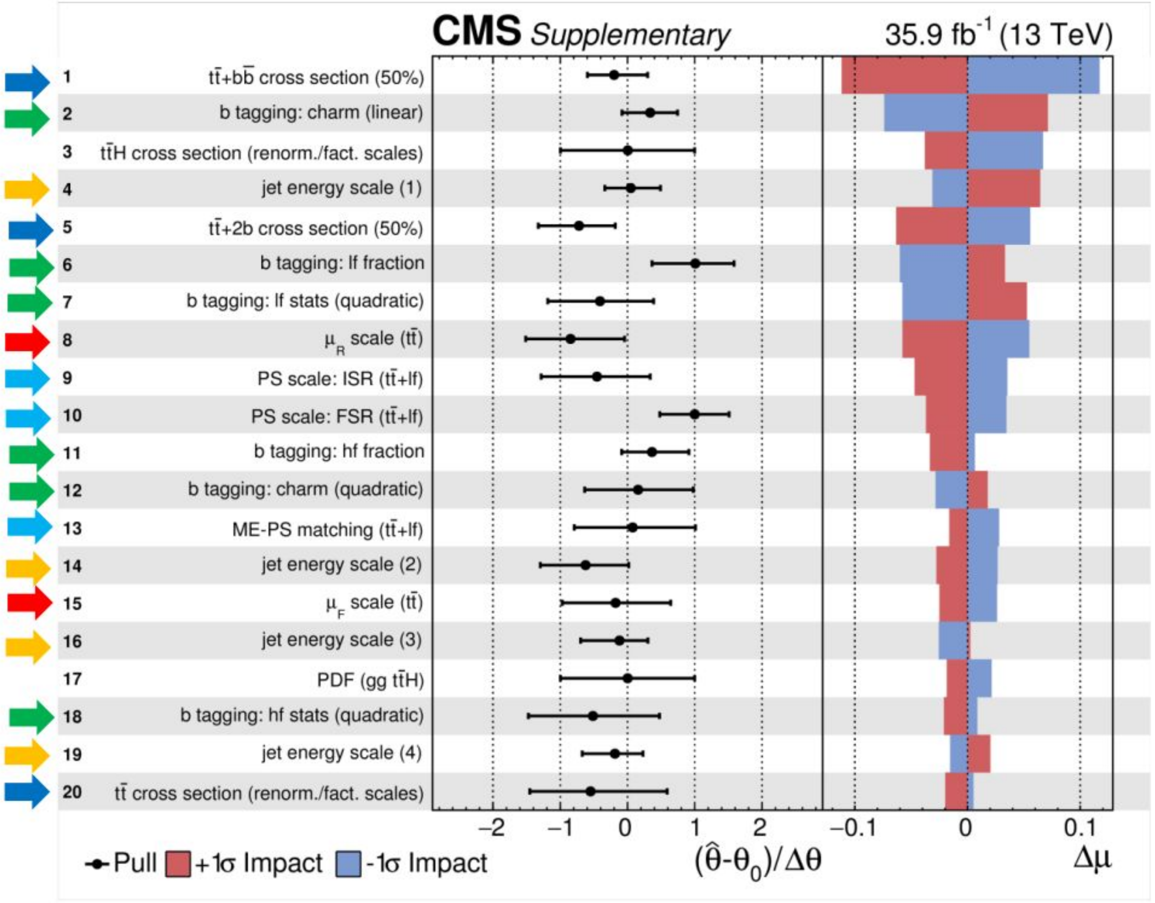


ttH(bb): Comparison with CMS

Syst legend

- ➡ ttbar norm only
- ➡ ttbar norm but can affect njets shape
- ➡ ttbar norm+shape (can affect MVA shape)
- ➡ btagging
- ➡ JES/JER

Will come to pulls and constraints in a minute

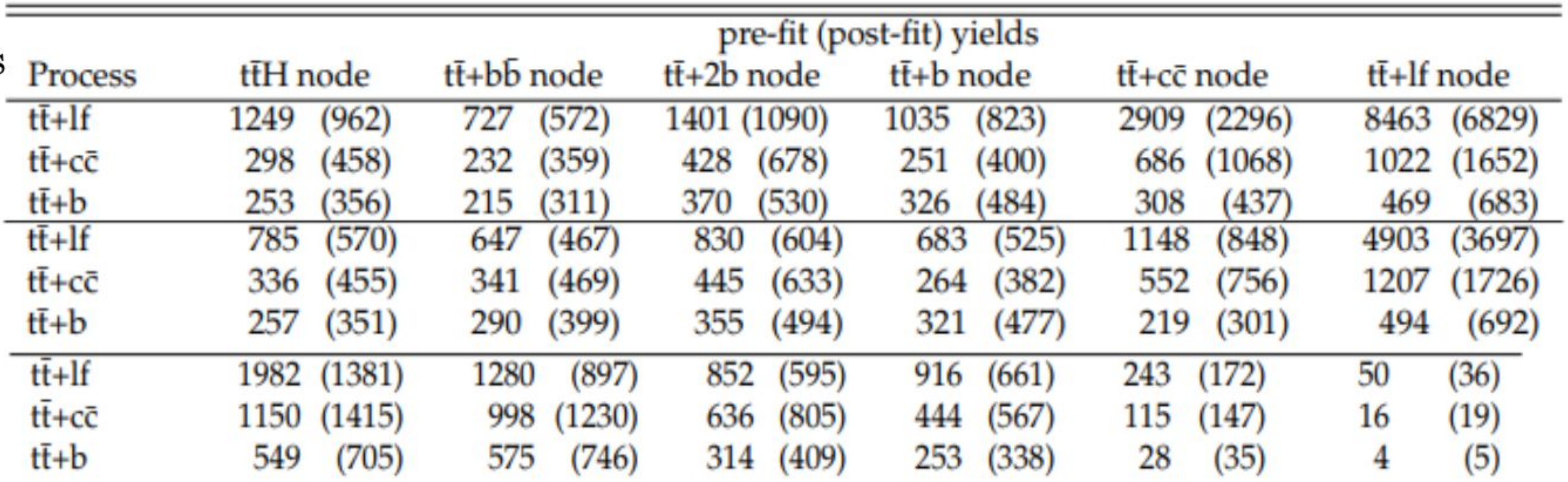


We know shape effect is negligible for both of these

ttH(bb): Comparison with CMS

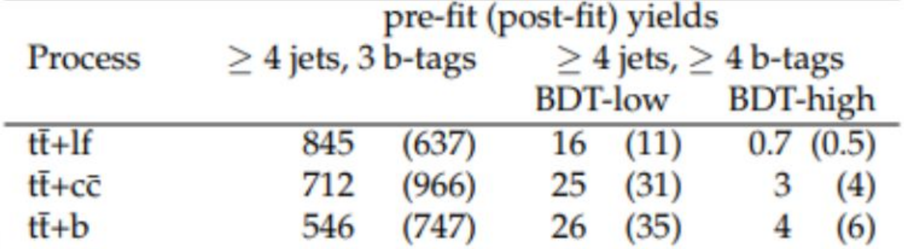
- But is it true that CMS modeling prefit is great?
- Why the systematic increase (decrease) of ttcc/b (tt+lf) yields?
- What are the pulls that decrease tt+lf?
- Are ttcc/b yields used to correct the shapes?
  - We have seen this several times in our fits

L+jets



		pre-fit (post-fit) yields					
	Process	$t\bar{t}H$ node	$t\bar{t}+b\bar{b}$ node	$t\bar{t}$+2b node	$t\bar{t}$+b node	$t\bar{t}+c\bar{c}$ node	$t\bar{t}$+lf node
4jets	$t\bar{t}$+lf	1249 (962)	727 (572)	1401 (1090)	1035 (823)	2909 (2296)	8463 (6829)
	$t\bar{t}+c\bar{c}$	298 (458)	232 (359)	428 (678)	251 (400)	686 (1068)	1022 (1652)
	$t\bar{t}$+b	253 (356)	215 (311)	370 (530)	326 (484)	308 (437)	469 (683)
5jets	$t\bar{t}$+lf	785 (570)	647 (467)	830 (604)	683 (525)	1148 (848)	4903 (3697)
	$t\bar{t}+c\bar{c}$	336 (455)	341 (469)	445 (633)	264 (382)	552 (756)	1207 (1726)
	$t\bar{t}$+b	257 (351)	290 (399)	355 (494)	321 (477)	219 (301)	494 (692)
6jets	$t\bar{t}$+lf	1982 (1381)	1280 (897)	852 (595)	916 (661)	243 (172)	50 (36)
	$t\bar{t}+c\bar{c}$	1150 (1415)	998 (1230)	636 (805)	444 (567)	115 (147)	16 (19)
	$t\bar{t}$+b	549 (705)	575 (746)	314 (409)	253 (338)	28 (35)	4 (5)

Dilep



	pre-fit (post-fit) yields		
Process	≥ 4 jets, 3 b-tags	≥ 4 jets, ≥ 4 b-tags	
		BDT-low	BDT-high
$t\bar{t}$+lf	845 (637)	16 (11)	0.7 (0.5)
$t\bar{t}+c\bar{c}$	712 (966)	25 (31)	3 (4)
$t\bar{t}$+b	546 (747)	26 (35)	4 (6)

tt+lf reduced by ~ 20% - 30%
ttcc and ttb increased by
around 50%

ttH(bb): Comparison with CMS

ttb norm factors (at least the two appearing in the plot) are pulled down

Opposite direction to the 50% yields change

Syst legend

- ➡ ttbar norm only
- ➡ ttbar norm but can affect njets shape
- ➡ ttbar norm+shape (can affect MVA shape)
- ➡ btagging
- ➡ JES/JER

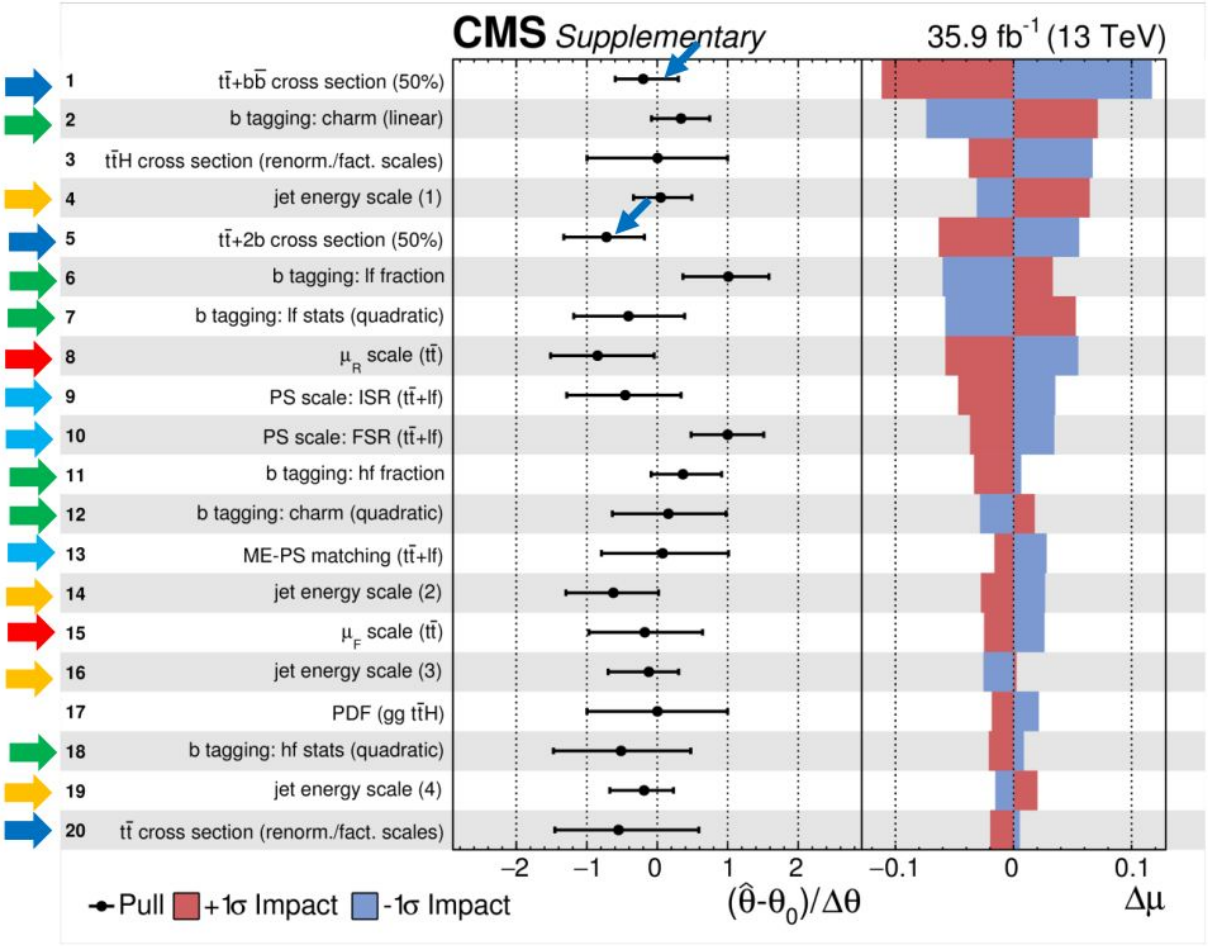
CMS answer was that it is b-tagging that pulls back up the yeilds

However they do not see this as a problem

As suspected it seems that they are using heavily the ttbar yields to correct the shapes

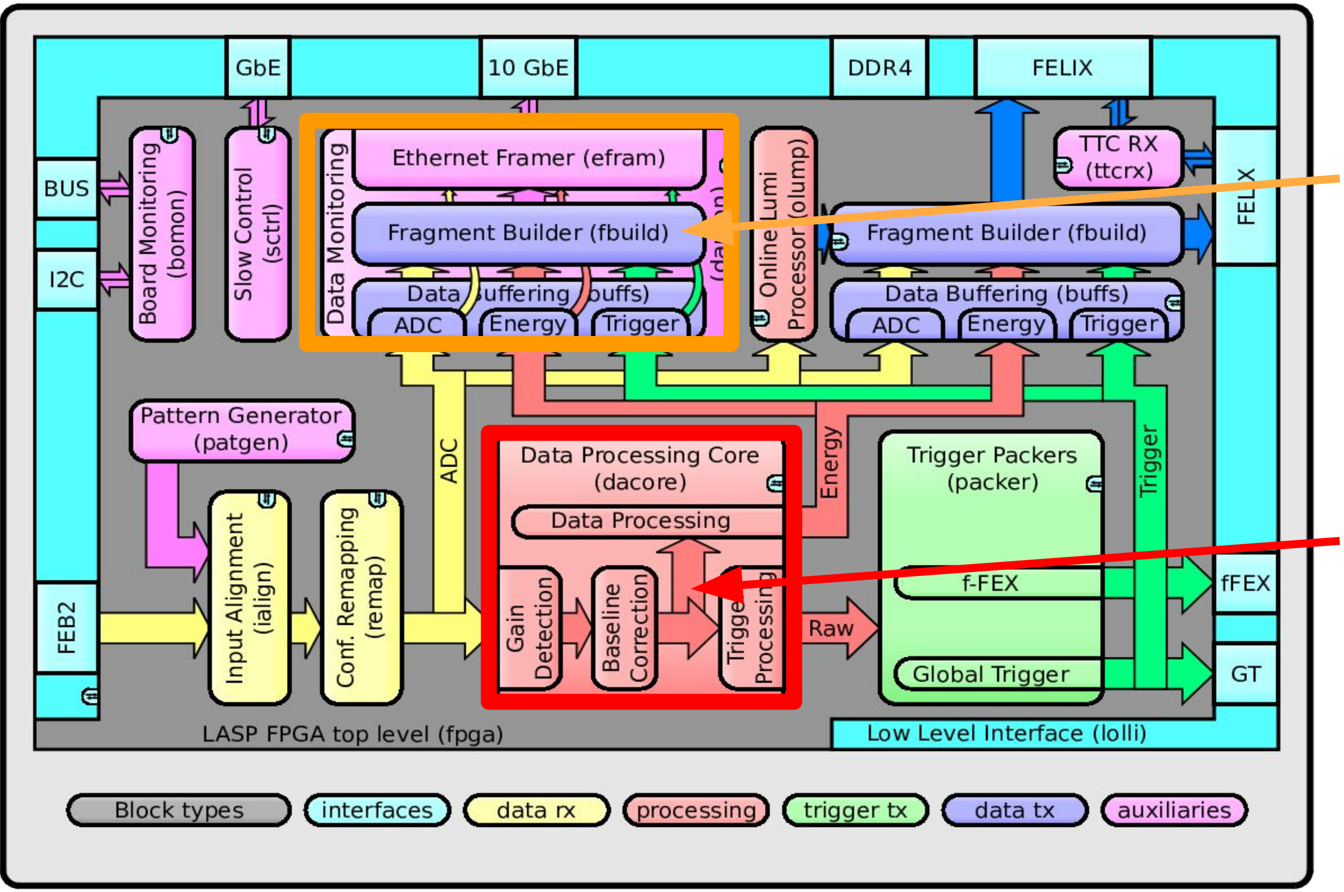
Then use detector syst to corrected back the yields

A correlation plot would help



LASP Firmware

- LASP board containing 2 processing units based on INTEL FPGAs
  - Demonstrator board available with stratix 10 FPGAs
  - Final board will be equipped with Agilex FPGAs
- One FPGA should process 384 channels
  - About 125 ns allocated latency for energy computation

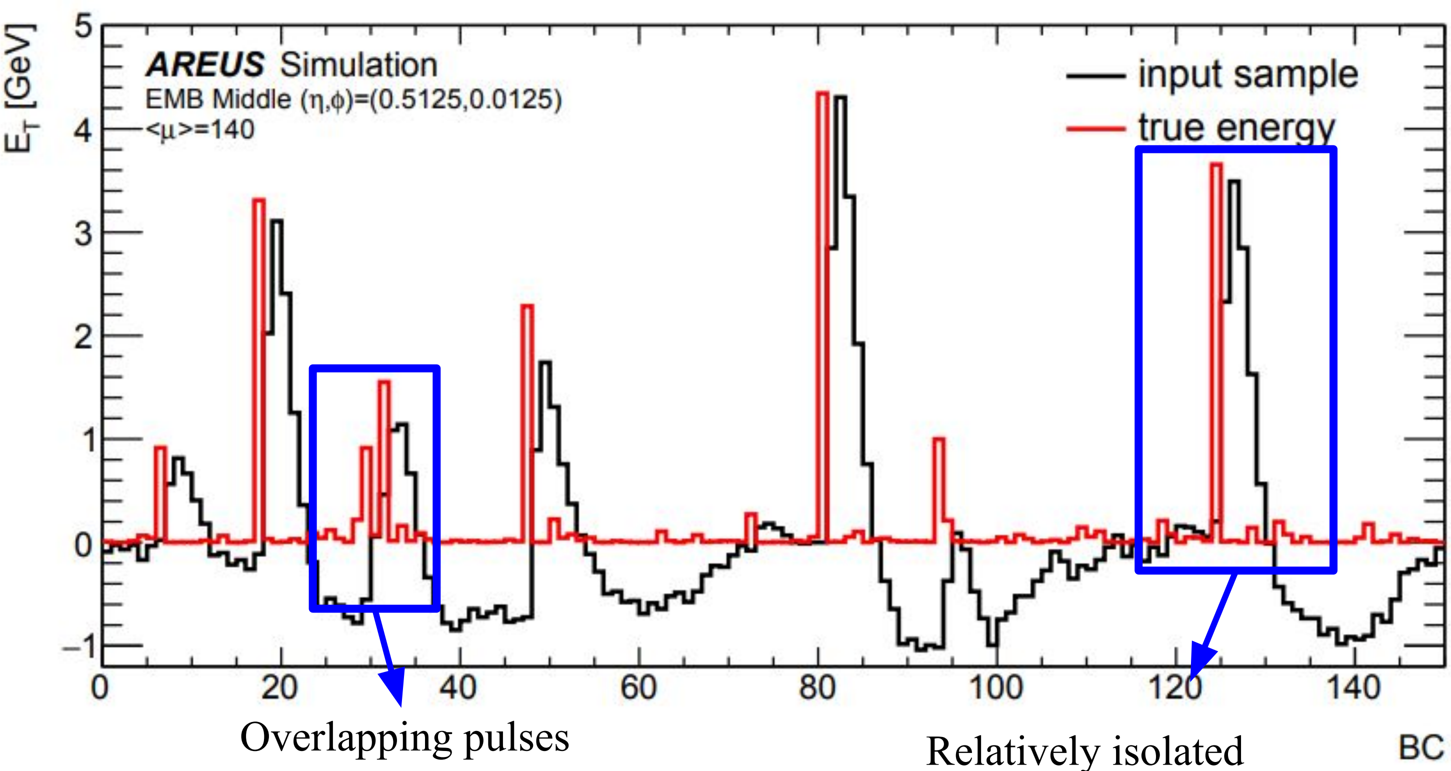


Buffering waiting for L1A
Could be used to refine energy computation with less stringent latency requirements

Compute energy at 40 MHz
Assign the energy to the correct bunch crossing (collision time)

Simulated LAr pulse chain

- Single cell pulse sequence
  - Using measured pulse shape
  - Minimum bias data with a specific pileup
  - Overlay regular high energy pulses with variable gap

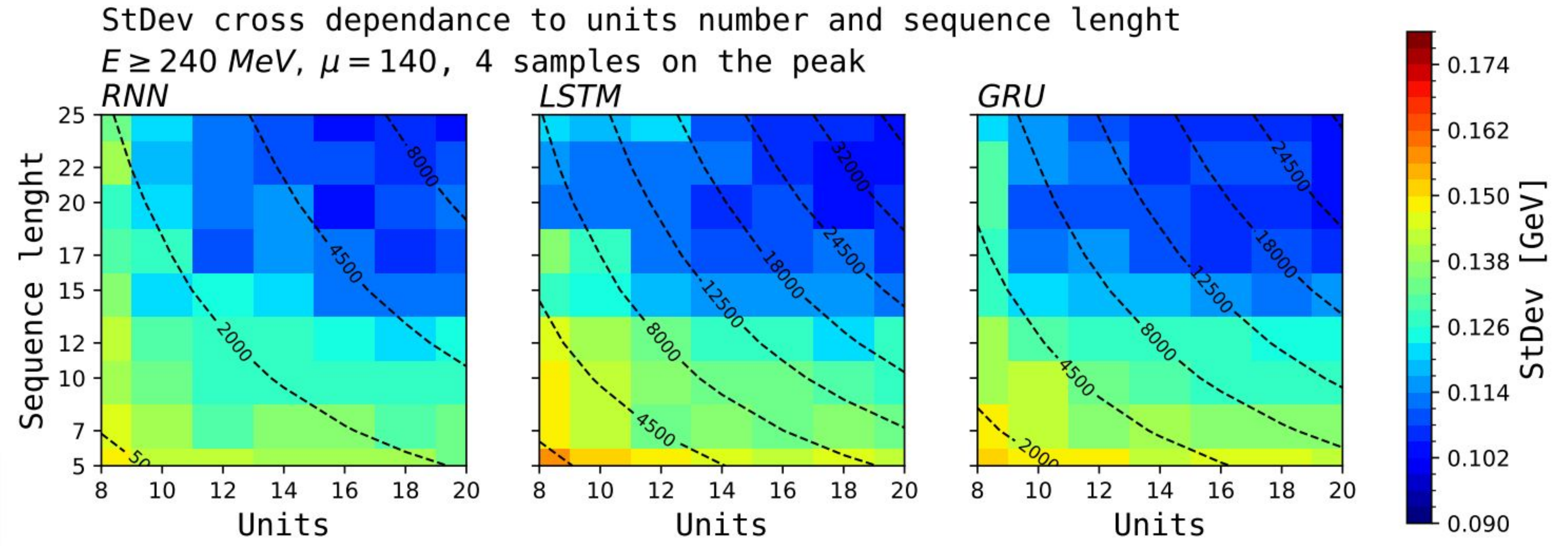


Overlapping pulses
Distorted pulse shape

Relatively isolated pulse

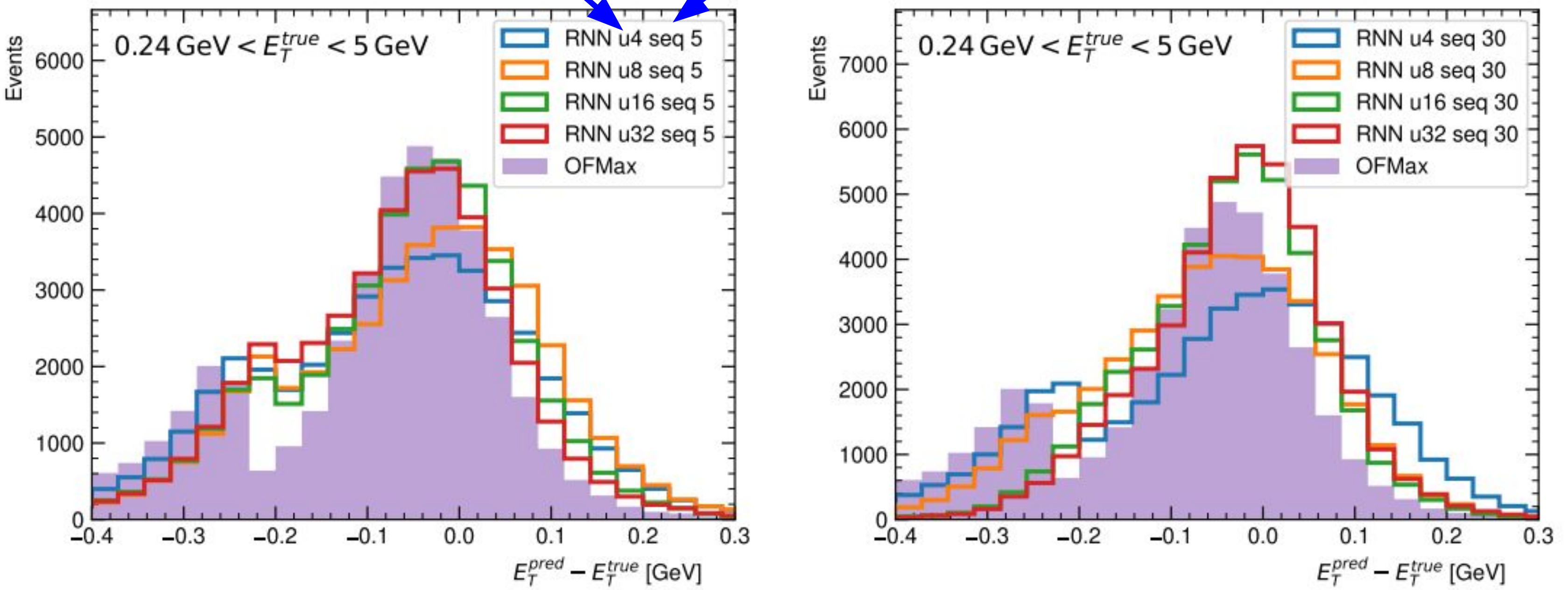
RNN Performance vs RNN Cell Type

- Checking performance of Vanilla-RNN, GRU and LSTM
    - Increased NN size by increasing sequence length and number of units
- Network size probed by number of multiplications (MAC units)
    - Dashed lines in the plots
- Vanilla-RNN can reach the same performance of GRU and LSTM with much less required MACs
    - Best adapted to fit in FPGAs



RNN Performance

- Compare energy resolution between RNNs and OFMax
  - RNNs with increased size
  - Keep size under control to fit FPGAs
- Second peak in resolution due to overlapping events
- Use Std. Dev. as metric (although the shape is not very gaussian)

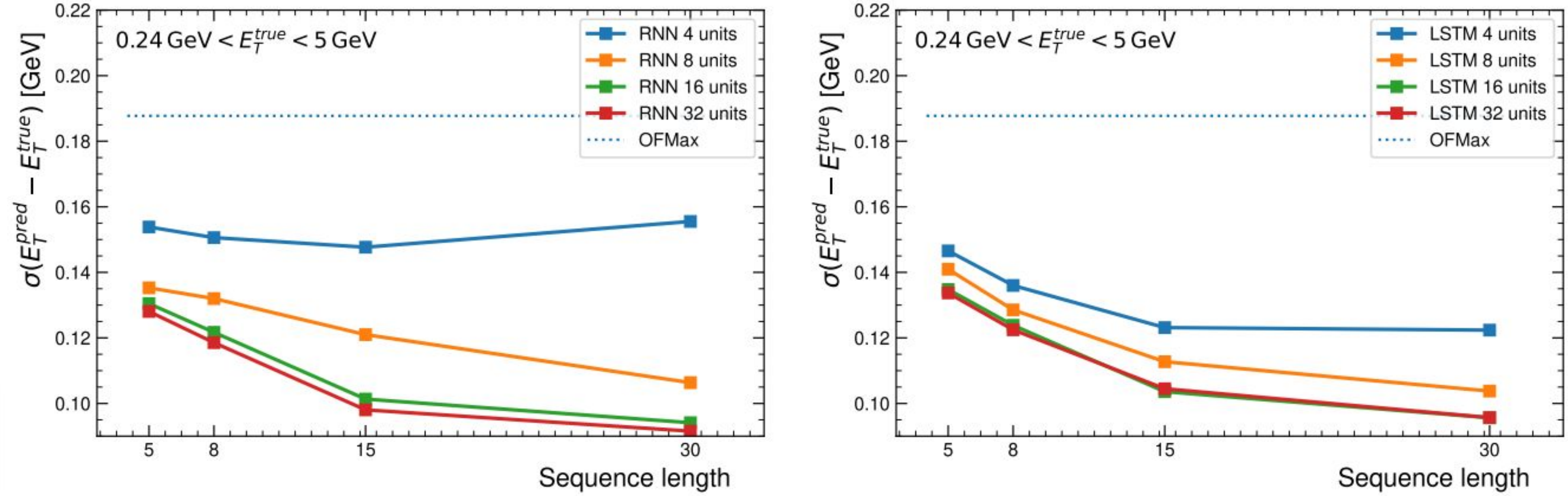


Number of units (u)

Sequence length (seq)

RNN Performance

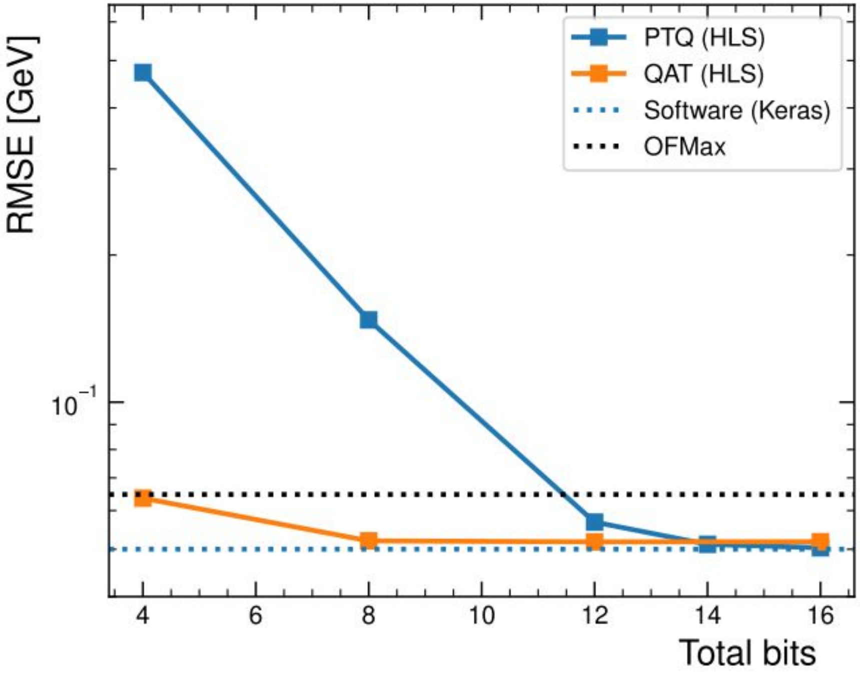
- Compare energy resolution between RNNs and OFMax
  - RNNs with increased size
  - Keep size under control to fit FPGAs
- Second peak in resolution due to overlapping events
- Use Std. Dev. as metric (although the shape is not very gaussian)



RNN Quantization

- RNNs need quantization to fit on FPGAs
    - Full floating point arithmetics takes a lot of FPGA resources
    - Need to use fixed points representations with small number of bits
- Quantize weights post training (PTQ)
    - Reduced resolution due to truncation/rounding
    - Can reach float precision with 16 bits
- Quantized Aware Training (using qKeras)
    - Optimize weights that are already quantized
    - Can reach float precision with 8 bits

- Stratix 10 considerations
    - One floating point multiplication per DSP
    - Two fixed point multiplication with 18x19 bits
- Agilex considerations
    - Additional DSP mode with four 9x9 bits multiplications
- Reduced number of bits allow to use more multiplications
    - Also matters for additions and timing closer



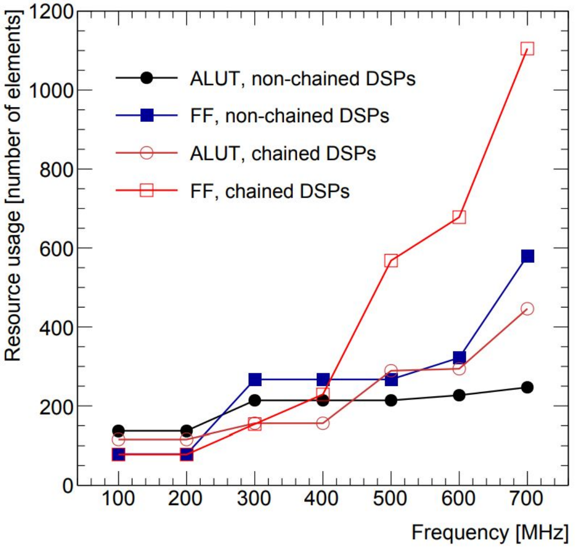
HLS optimisations

- Optimisation needed to fit RNNs within resource and latency limitations
  - Impossible to fit 384 NNs in the FPGAs, need to serialize (time multiplexing)
  - Need to go to high frequency
- Several optimisations are performed
  - Activation functions in LUT (only for LSTM)
  - Number of bits in fixed point representation (18x19 to match Stratix 10 DSP)
  - Rounding and truncation in arithmetic operations
  - Implementation of vector/matrix multiplication (Dot product)
- Dot product implementations
  - Naive C++: let HLS do it all
  - ACC37: accumulate (sum) in DSPs by chaining them
  - ACC19: ACC in ALUT

$$A.B = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_8 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_8 \end{bmatrix} = \sum_{i=0}^{7} a_i.b_i$$

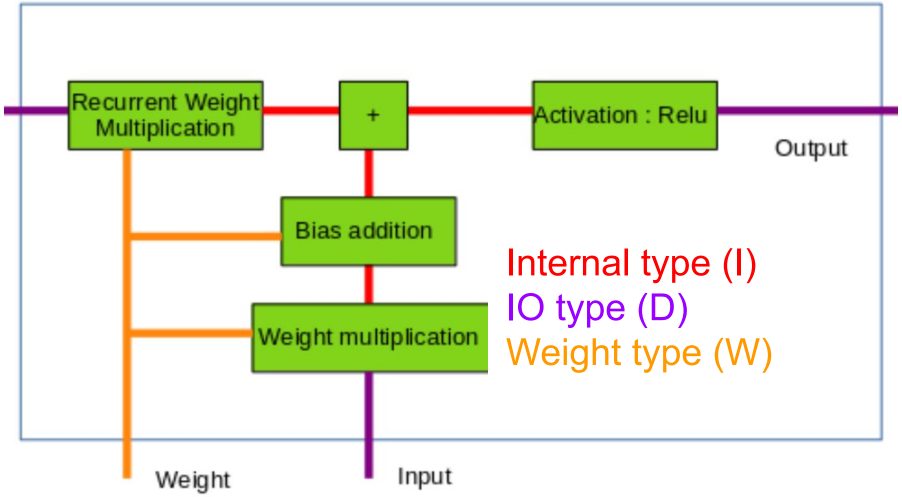
@100 MHz

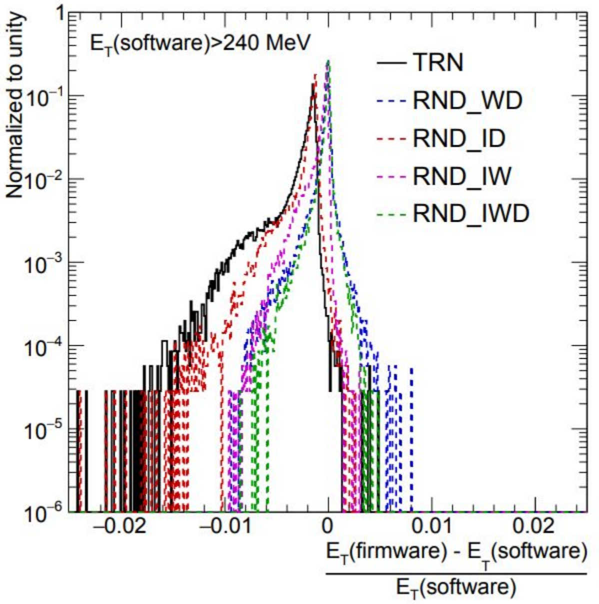
Implementation	ALUTs	FF	DSP
C++ style	709	222	8
ACC37	116	79	4
ACC19	137	78	4

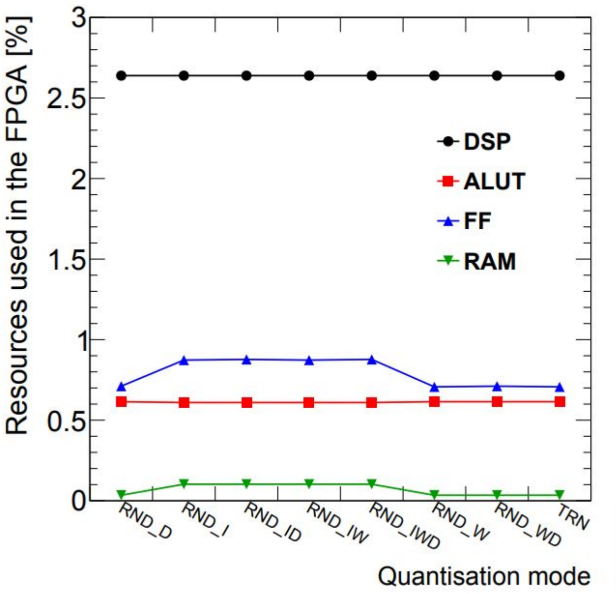


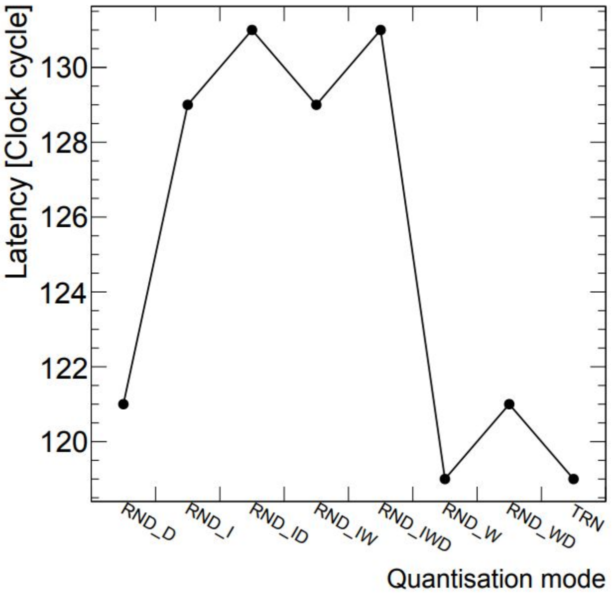
Rounding vs Truncation

- Compromise between resolution and resource usage and latency
    - Truncation of IO and Internal types leads to important reduction of latency with small impact on energy resolution
    - Weight type rounded in software
        - No impact on latency
- Use truncation in the firmware
    - Will become less relevant with QAT









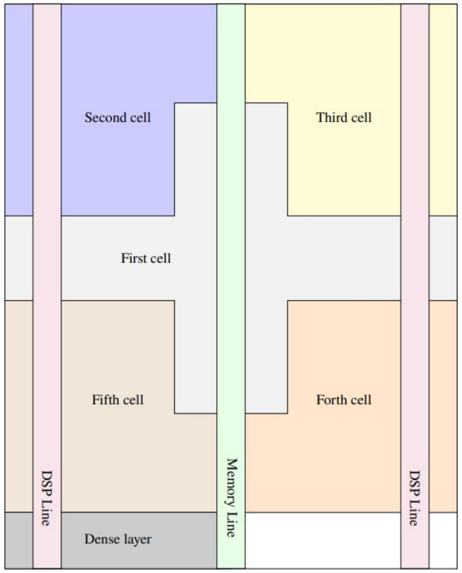
VHDL implementation of Vanilla RNN

- HLS does not allow to reach the target frequency and resource usage
  - Increase of the RNN ALM resources and reduction of FMax as we add networks to the FPGA
- Move to VHDL for the final fine tuning
- Force placement of the RNN components
  - Allow to better tackle timing violations and improve FMax
- Use incremental compilation
  - Keep networks with no timing violations and recompile only the rest

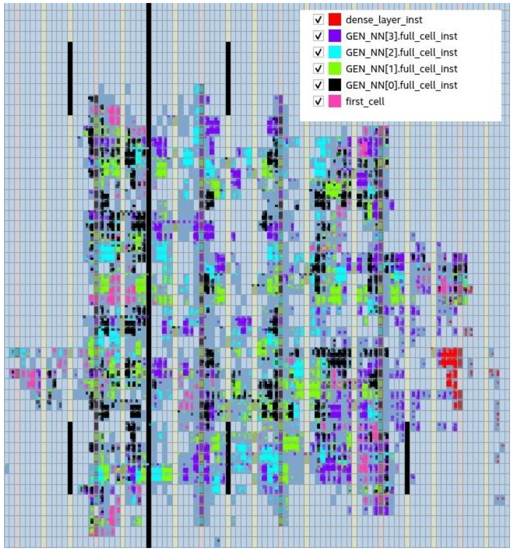
Optimized placement of RNN cells

First cells in the middle and connected to all cells (common computations done only in first cell and propagated to the others

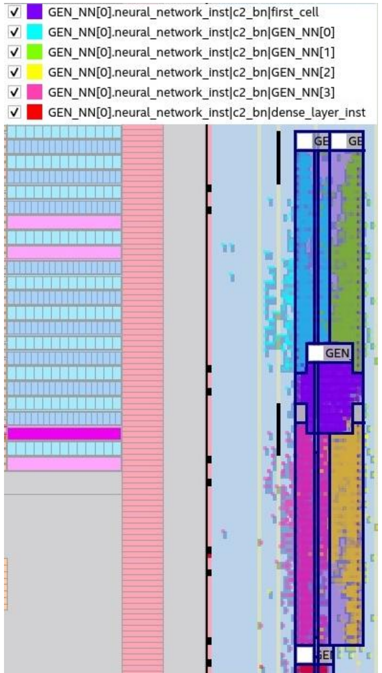
Dense layer next to last cell



HLS placement



VHDL forced placement



Testing on hardware

- VHDL implementation tested on Startix 10 DevKit
- Test firmware to inject input and weights and collect the output is built
    - ○ Data extraction using a JTAG-UART connection with a NIOS
- Data match firmware simulation bit-by-bit
- Firmware resolution < 0.1% as expected from simulation