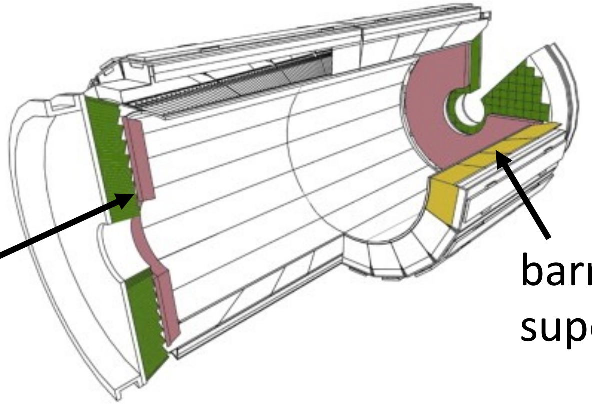
Performance and Upgrade of the CMS Electromagnetic Calorimeter

Charlotte Cooke

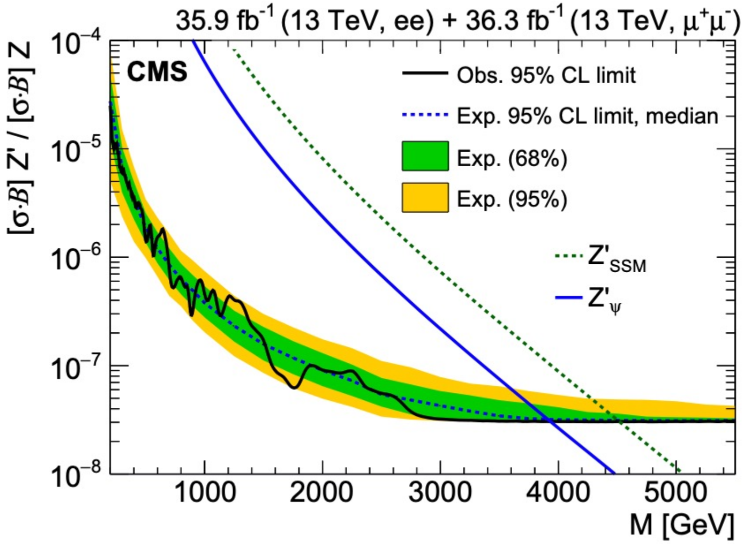
University of Bristol and Rutherford Appleton Laboratory

Introduction

- Lead tungstate ($PbWO_4$) crystal calorimeter
- Provides excellent energy resolution
  - In harsh radiation environment
  - Over a wide range of energies from O(100 MeV) to O(1 TeV)
  - Achieved 1% mass resolution for low-mass Higgs in γγ decay channel



- My work specifically concerns the Z' search – looking for new high mass resonances decaying to two leptons
  - CMS requires electron candidates to pass a set of dedicated high energy electron criteria (HEEP ID)



The upper limits (95% CL) on the product of production cross section and branching fraction for a spin-1 resonance with a width equal to 0.6% of the resonance mass.
From JHEP 06 (2018) 120 arXiv:1803.06292

High Energy Electron Positron (HEEP) ID

- **Aim: check if “ultra legacy” (better calibrated) Run 2 (2015-2018) data improves efficiency and data/MC agreement**
- HEEP ID basics:
  - Simple, robust ID
  - Subdetector based rather than particle flow
    - Uses information from ECAL, HCAL and tracker
  - Requires that the lateral spread of energy deposits in the ECAL is consistent with that of a single electron and that the track is matched to the ECAL deposit

HEEP ID Efficiency Method

Tag and Probe Criteria

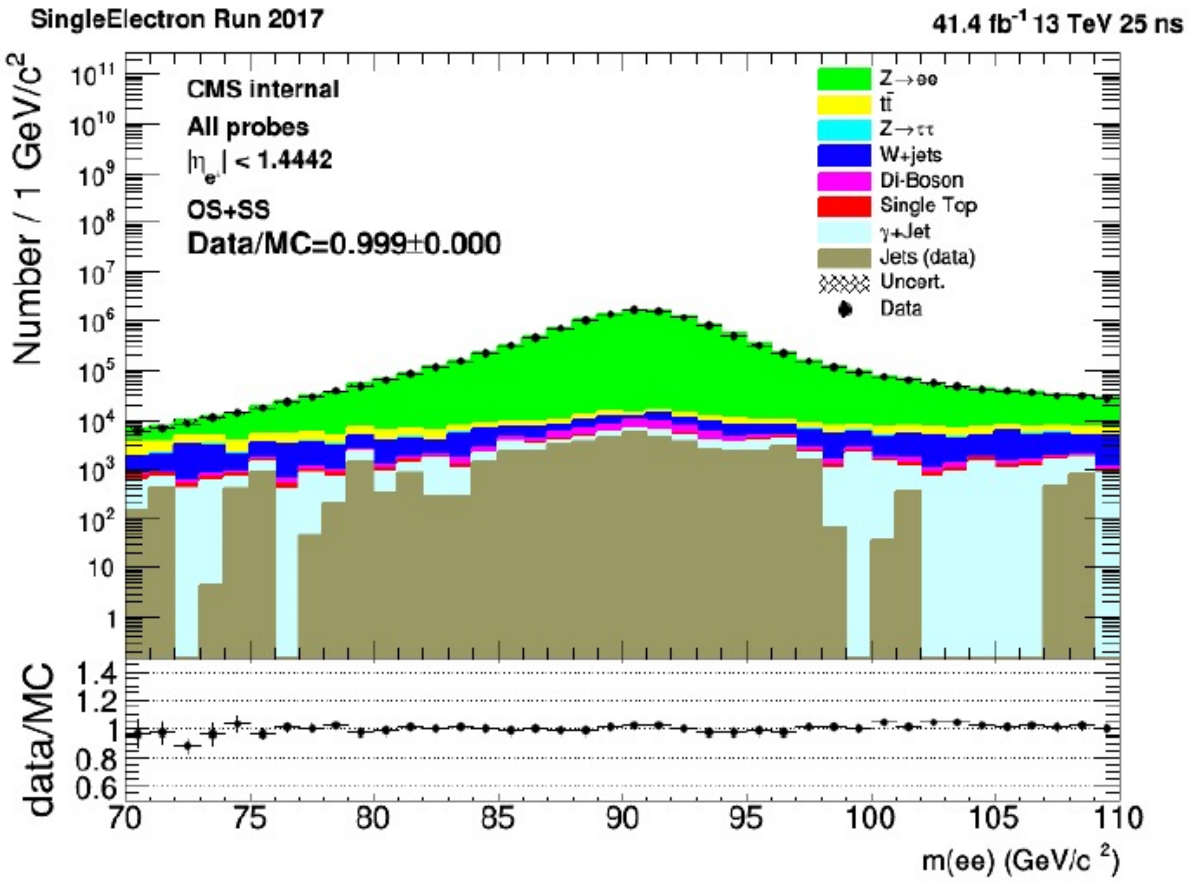
- Using Z->ee events
- Tag passes HEEP ID V7.0
- Tag is a barrel electron
- Tag is matched to a 32 GeV electron trigger
- Probe passes the acceptance criteria (η_{SC} and E_T) of the HEEP ID
- The invariant mass of the tag and probe is in the window 70<M(ee)<110 GeV/c^2
- All possible combinations of tag and probe in an event are selected

Additional Considerations

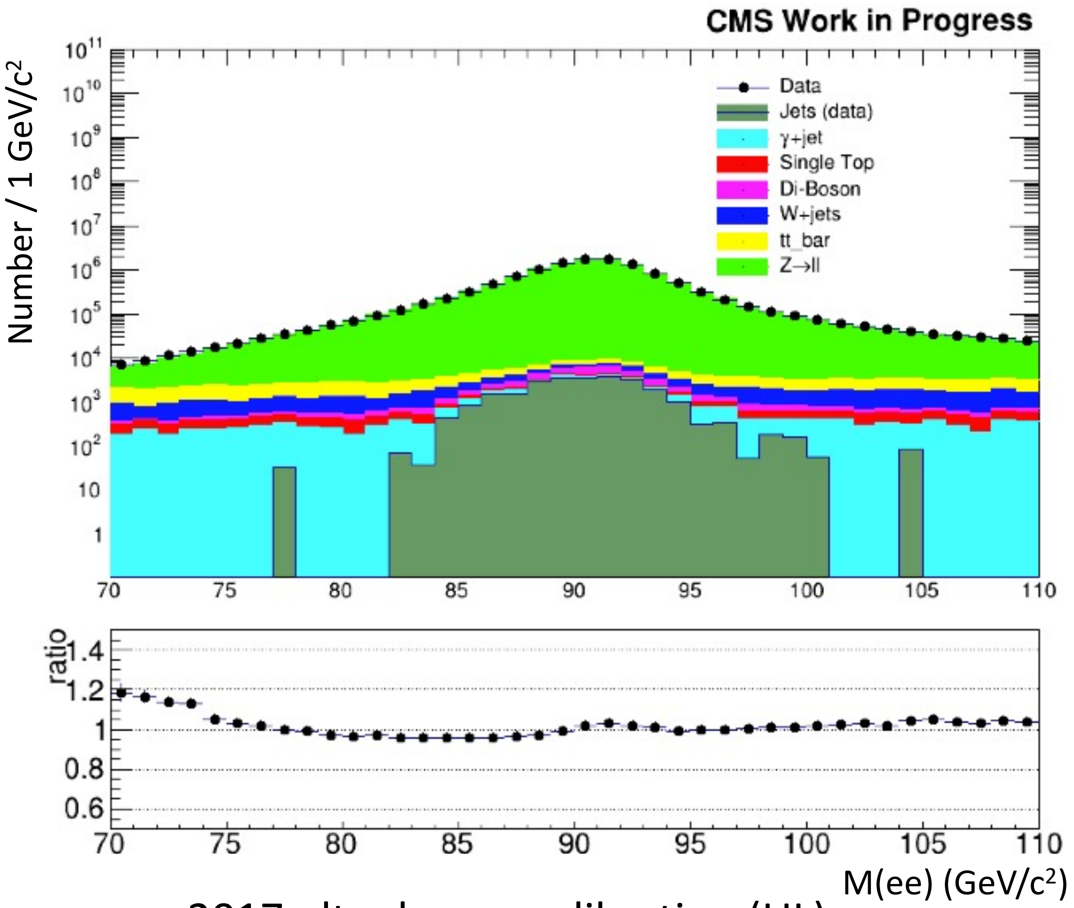
- MC reweighted according to trigger turn on curve, pileup and cross section
- Jets estimated from data looking at same sign pairs failing the HEEP ID
- CMS standard scale and smearing corrections are applied to match MC to data
- Efficiency defined as

$$\varepsilon = \frac{N_{passing\ probes}}{N_{all\ probes}}$$

HEEP ID using Ultra Legacy Data (Barrel)

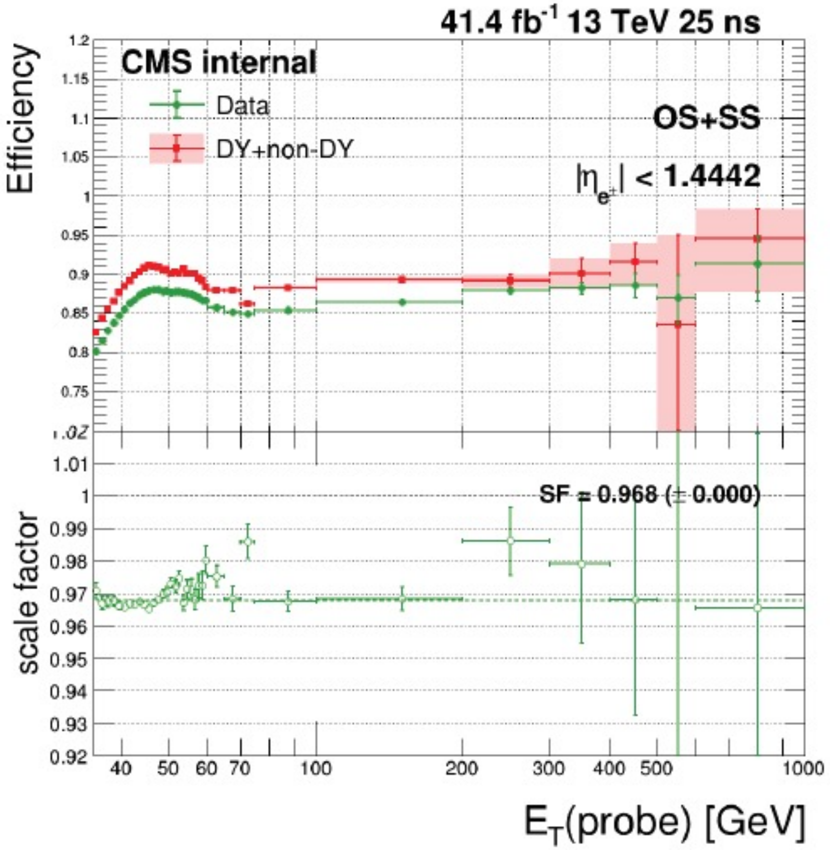


2017 end of year calibration (EOY)

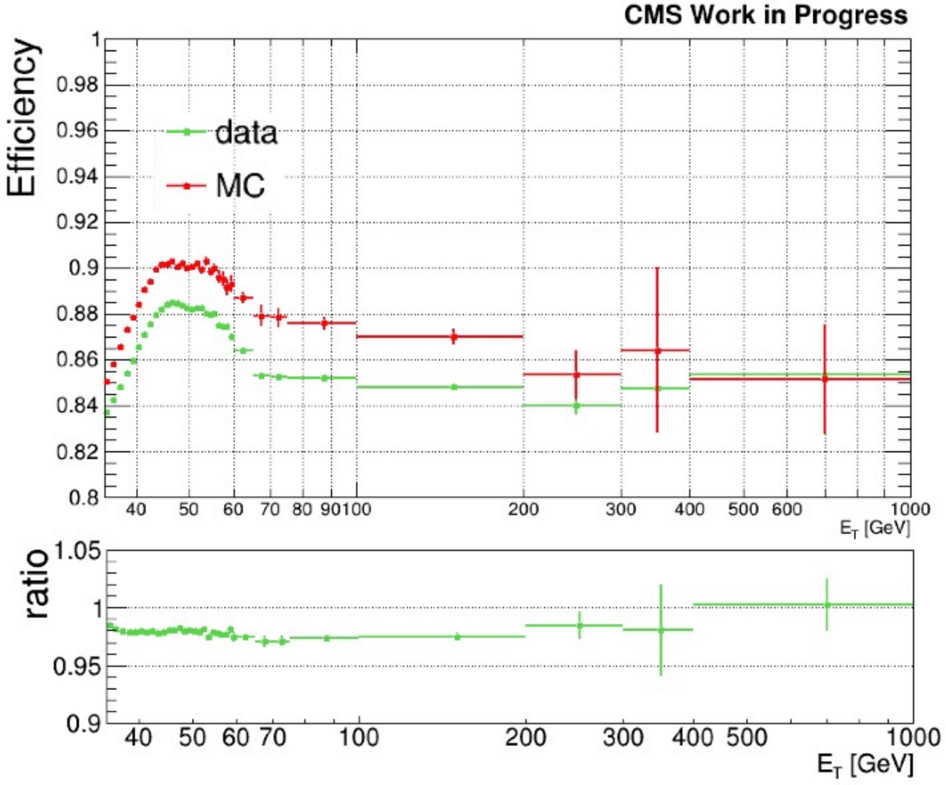


2017 ultra legacy calibration (UL)
Work in progress

HEEP ID using Ultra Legacy Data (Barrel)

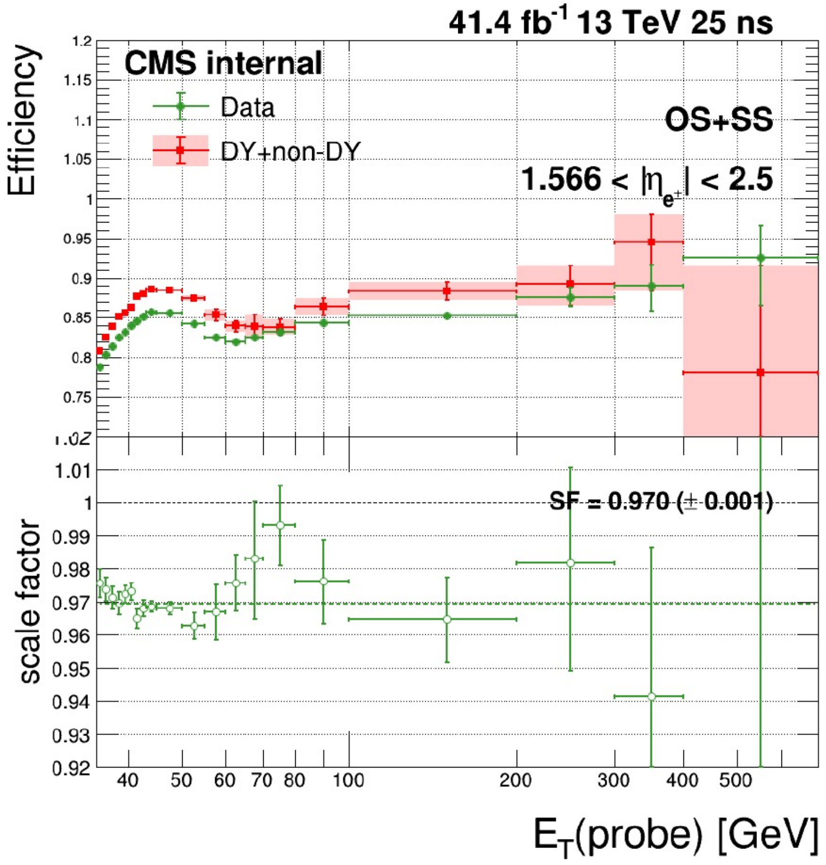


2017 end of year calibration (EOY)

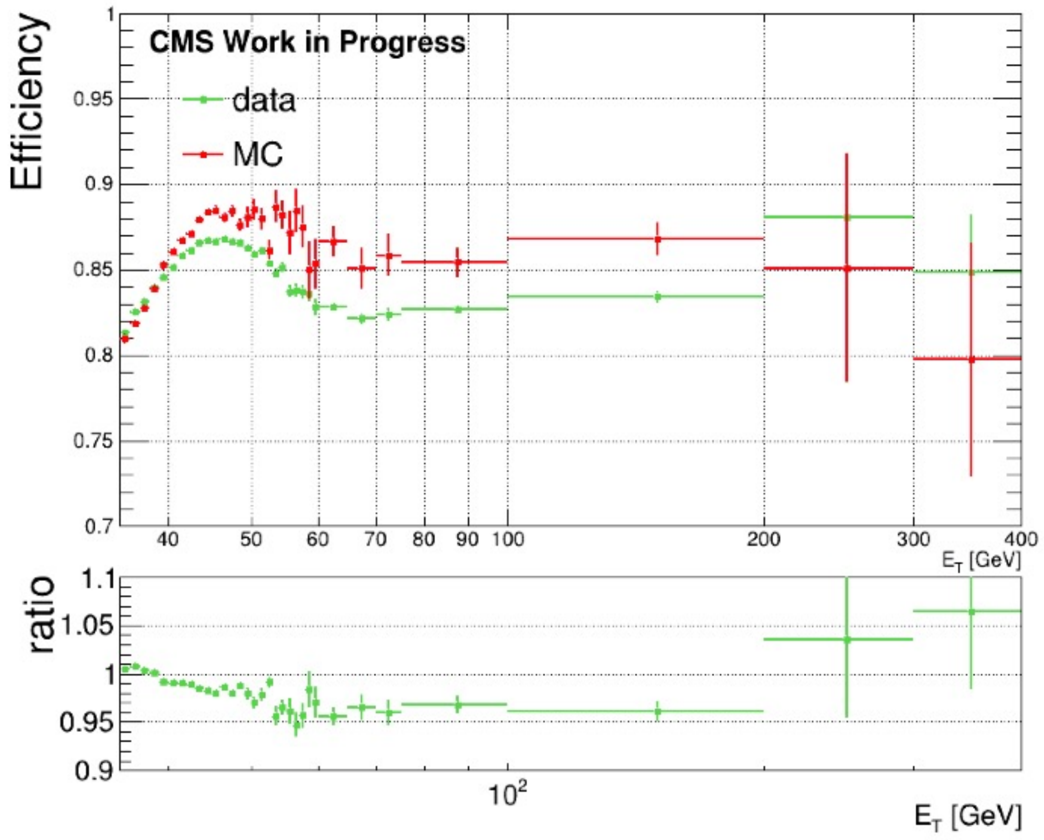


2017 ultra legacy calibration (UL)
Work in progress

HEEP ID using Ultra Legacy Data (Endcap)



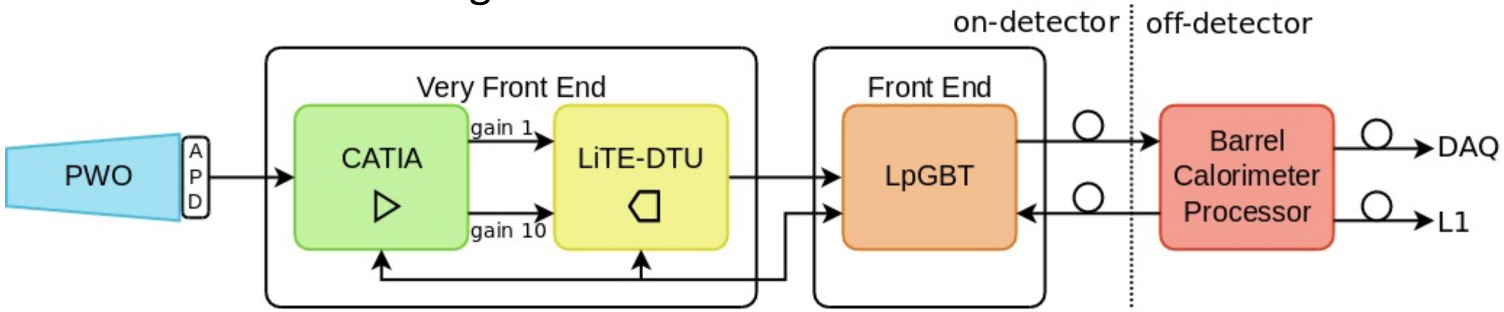
2017 end of year calibration (EOY)



2017 ultra legacy calibration (UL)
Work in progress

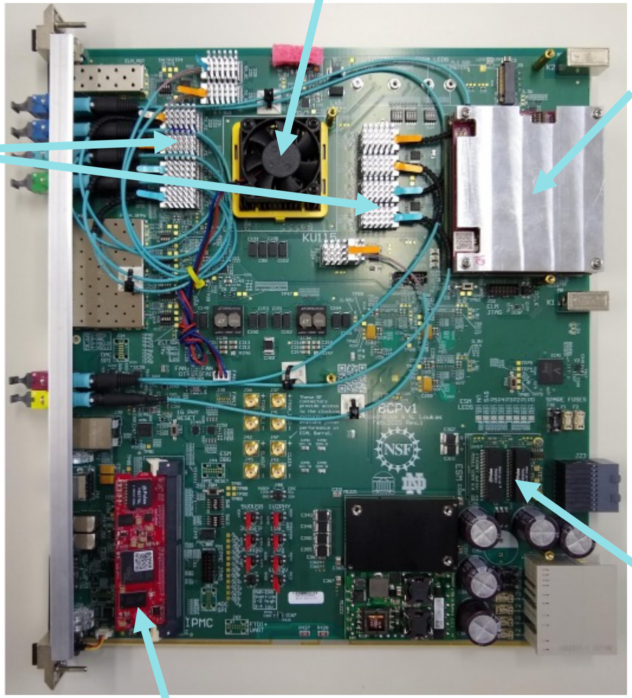
Phase II ECAL Upgrade Overview

- Refurbish ECAL barrel supermodules during Long Shutdown 3 (2026-2028)
- Keep the lead tungstate crystals in the barrel
  - Reduce temperature from 18^{O}C to 9^{O}C to keep noise below 250 MeV
- Replace the on and off detector electronics
  - Maintain performance from Phase I and meet trigger requirements
  - Use new radiation hard ASICS
    - Factor of 4 increase in sampling rate (160 MHz) to meet Phase II trigger requirements
    - ~30 ps timing resolution
    - Single crystal info to trigger
    - More advanced algorithms in off detector FPGAs



Trigger Primitive Algorithms

- Algorithms include:
  - **Spike rejection**
    - **Pulse shape and Swiss-cross**
  - Conversion of digitized pulse data into transverse energy
  - Precise timing measurement
  - **Basic clustering of localised energy**
  - TP encoding and shipping to L1 trigger system
  - Generation and transmission of clock and control signals to the FE
- I work on the algorithms in bold which are being developed in high level synthesis (HLS)

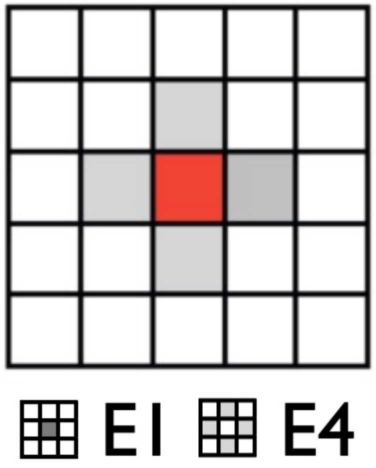


BCP V1

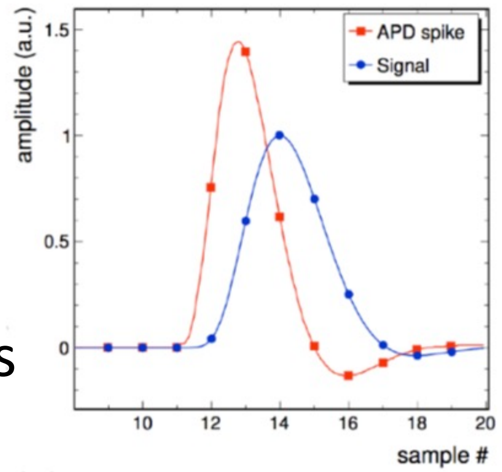
Swiss Cross and Pulse Shape

- Spike: anomalous APD signal caused by direct impact on photodiode

- Swiss cross:
  - Previously used offline and in high level trigger
  - Uses the swiss cross variable $(1 - E_4/E_1)$
  - Cut placed at 0.95 separates spikes and EM showers with high efficiency



- Pulse Shape:
  - Uses a linear discriminant variable to distinguish between EM showers and spikes
  - Tested 300 channels in software and 72 channels on hardware



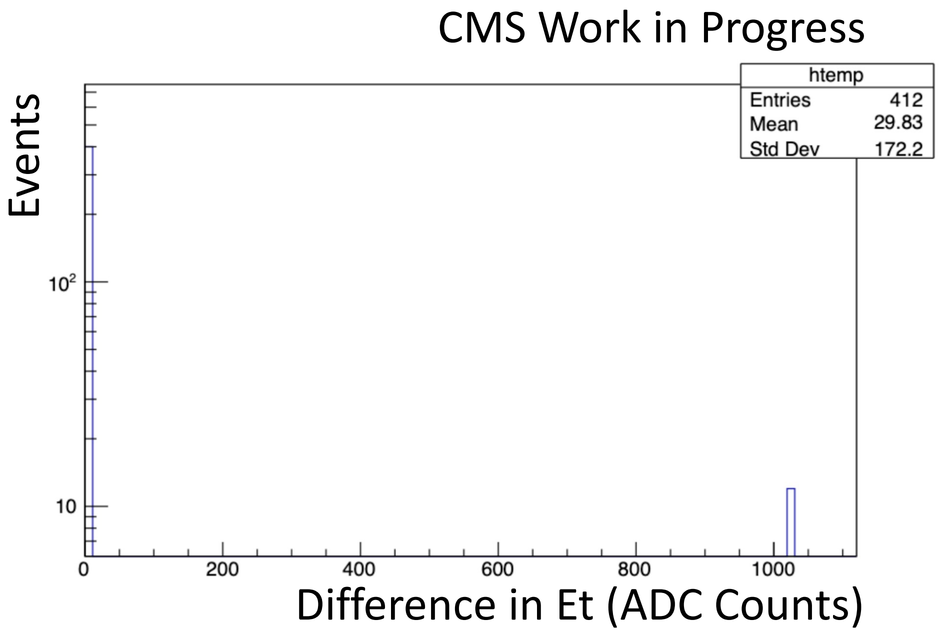
From: the Phase-2 Upgrade of the CMS Barrel Calorimeters TDR

	DSP	FF	LUT
Utilisation (%)	16	8	13

Resource estimates for KU115 FGPA

Algorithm Testing

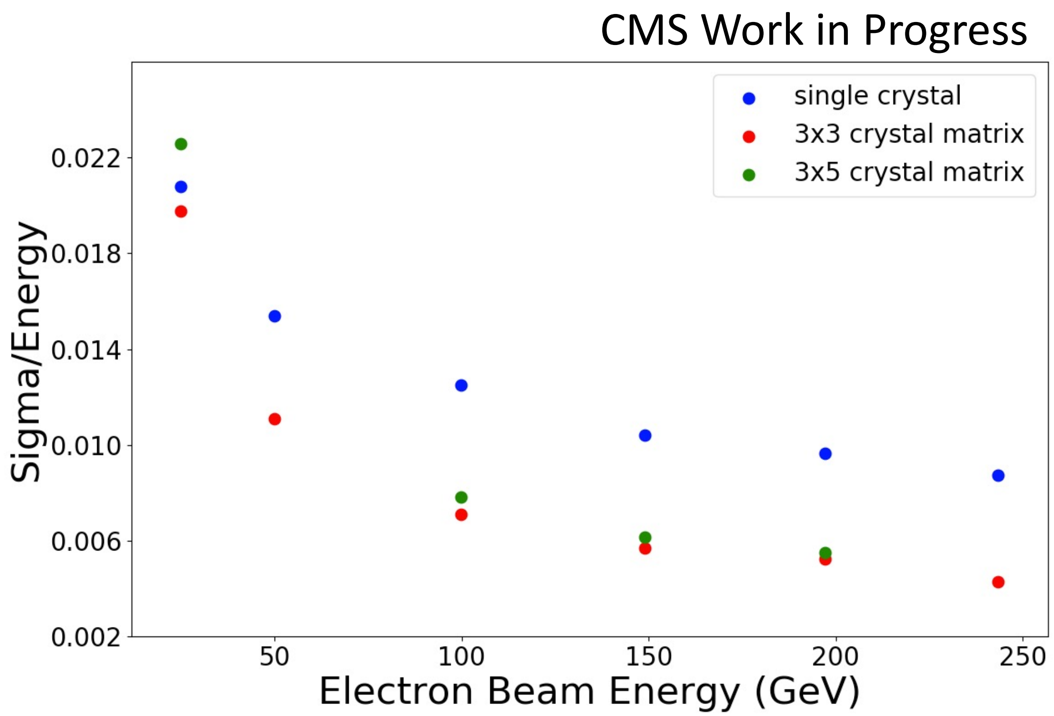
- We have developed a software framework to test HLS algorithms in CMS simulation and reconstruction framework against C++ counterparts
- Using test beam data with prototype Phase II electronics
- Validation plots allow for a quick and easy comparison of the two algorithms
  - Variables checked are:
    - Pulse shape spike flag
    - Swiss cross spike flag
    - Number of clusters
    - Cluster eta
    - Cluster phi
    - Cluster Et
    - Number of crystals in a cluster



Plot showing difference in Et between HLS and C++ due to overflow in HLS algorithm (now fixed)

Clustering Algorithm

- Test beam studies found a 3x3 cluster provides good resolution for EM showers
  - Can use same 3x3 crystal region as swiss cross
- Clustering requirements:
  - Seed has a greater Et than any crystals it shares a side with (corners not considered)
  - Seed Et > 0.25 GeV
  - Seed is not an pulse shape spike

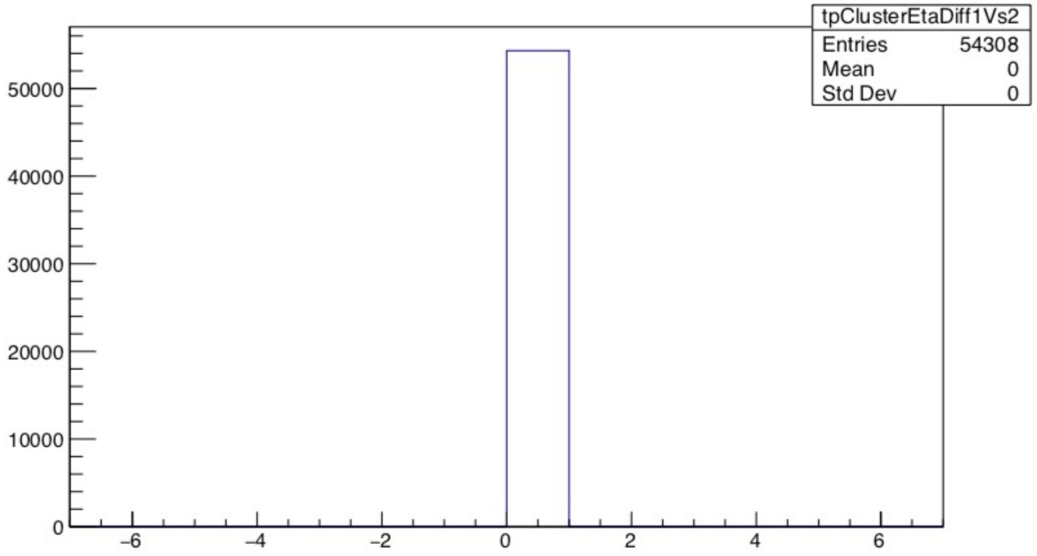


2018 ECAL Test Beam

Clustering Algorithm

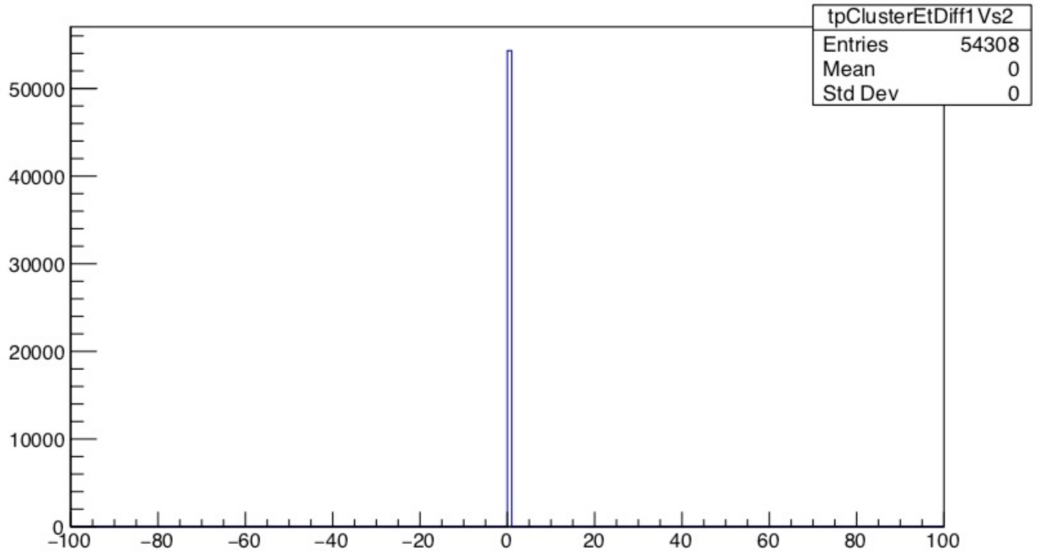
- Tested in software with no discrepancies in over 50000 events from 2018 ECAL test beam
- Hardware test (with swiss cross) planned for this summer

CMS Work in Progress



Difference in Eta (crystals)

CMS Work in Progress



Difference in Et

Conclusion and Next Steps

- HEEP ID
  - UL performance similar to that seen in EOY
    - Scale factor observed to be flat vs Et
  - ID needs to be verified at start of run 3
    - Due to upgraded HCAL and pixel detectors, new calibrations and alignment
- Algorithms
  - First versions of algorithms tested and working
  - Further software and hardware tests of HLS algorithms, including the use of more recent test beam data
  - Further optimisation to reduce resource usage

Backup Slides

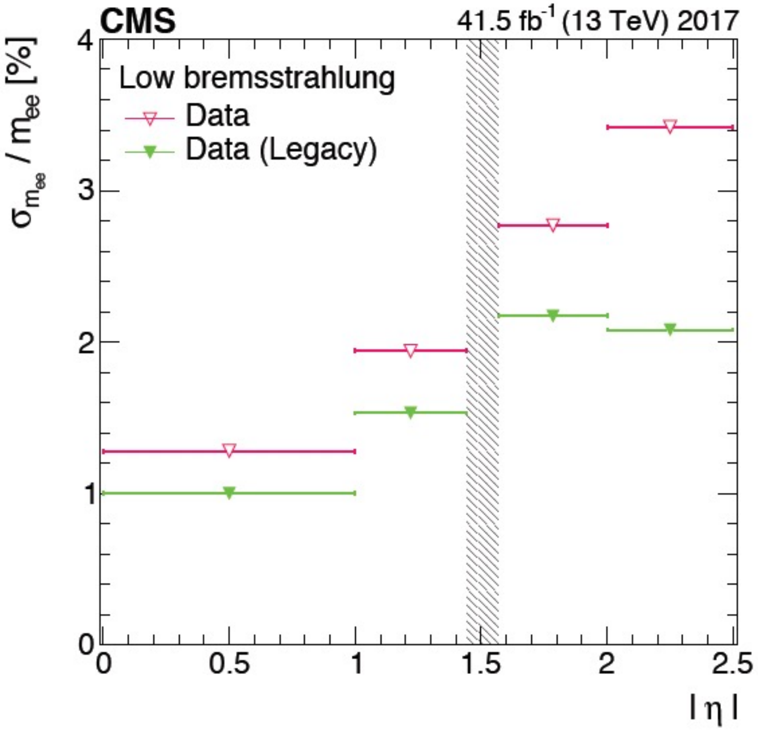
High Energy Electron Positron (HEEP) ID

Variable	Barrel	Endcap
E_T	> 35 GeV	> 35 GeV
η range	$\|\eta_{sc}\|< 1.4442$	$1.566<\|\eta_{sc}\|< 2.5$
isEcalDriven	=1	=1
$\|\Delta\eta_{in}^{seed}\|$	< 0.004	< 0.006
$\|\Delta\phi_{in}\|$	< 0.06	< 0.06
H/E	< 1/E + 0.05	< 5/E + 0.05
full 5x5 $\sigma_{i\eta i\eta}$	n/a	<0.03
full 5x5 E^{2x5}/E^{5x5}	> 0.94 OR E^{1x5}/E^{5x5} > 0.83	n/a
EM + Had Depth 1 Isolation	<2+0.03*Et +0.28*rho	< 2.5 +0.28*rho for Et<50 else < 2.5+0.03*(Et-50) +0.28*rho
Track Isol: Trk Pt	<5	<5
Inner Layer Lost Hits	<=1	<=1
|dxy|	<0.02	<0.05

- **Aim: check if "ultra legacy" (better calibrated) Run 2 (2015-2018) data improves efficiency and data/MC agreement**
- HEEP ID basics:
  - Simple, robust ID
  - Subdetector based rather than particle flow
    - Uses information from ECAL, HCAL and tracker
  - Requires that the lateral spread of energy deposits in the ECAL is consistent with that of a single electron and that the track is matched to the ECAL deposit

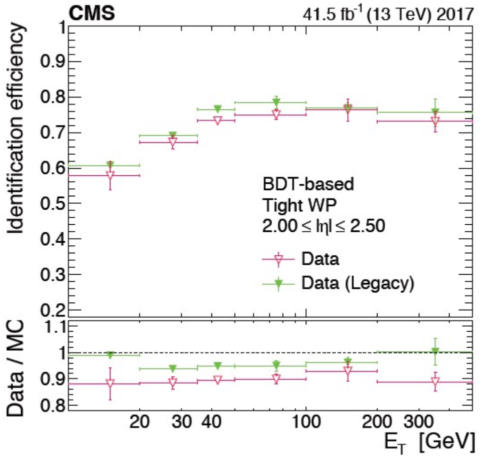
Ultra Legacy Data

- Recalibrated ECAL using Run 2 data
- Well known resonances of π_0 and Z were used calibrate crystals on a per year basis
- Evaluating impact of UL calibration on Z' analysis:
  - UL provides better electron energy resolution, especially in forward region
  - UL may also improve HEEP ID efficiency and data/MC agreement

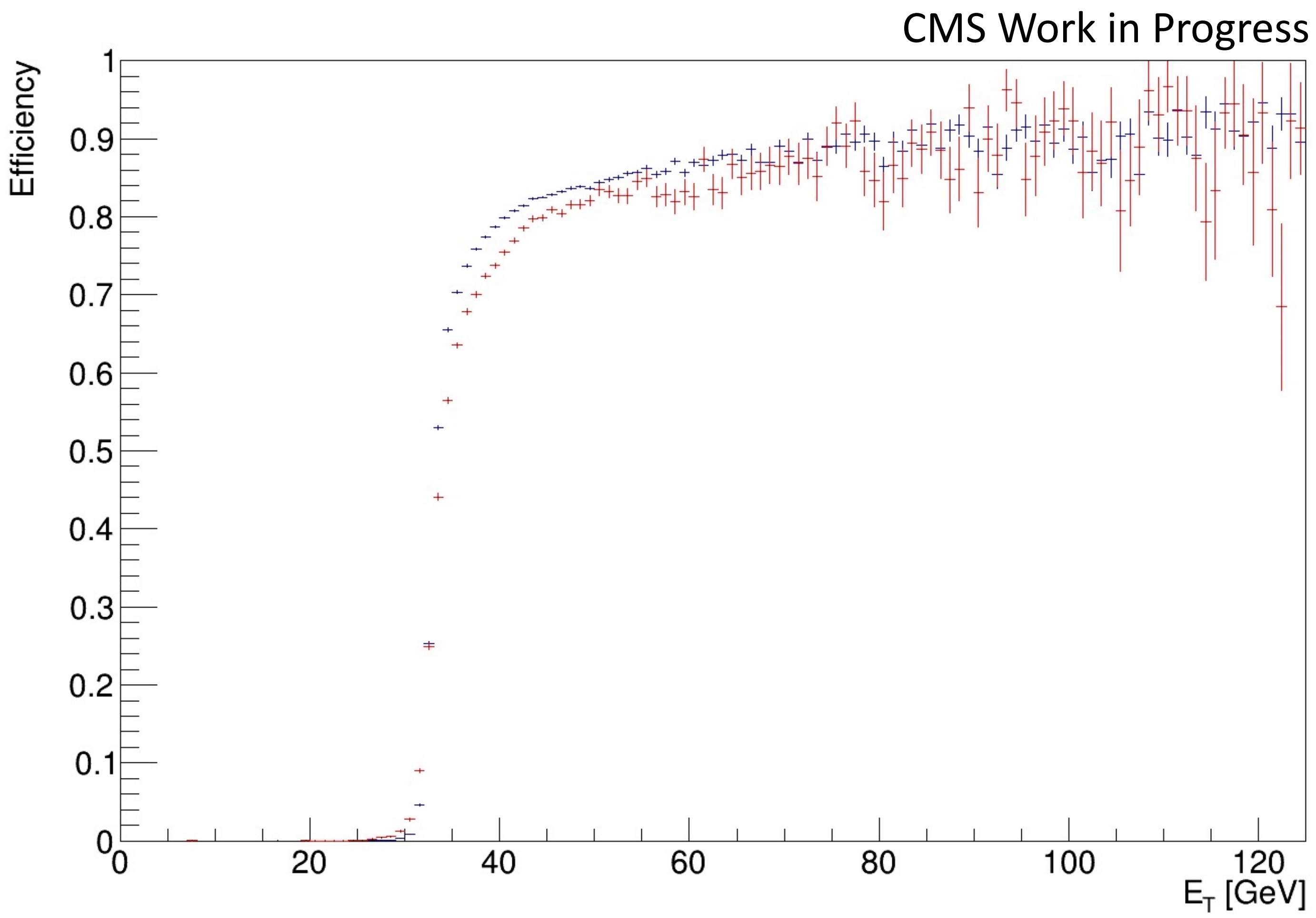


Dielectron mass resolution from Z -> ee events
From [arXiv:2012.06888]

Tight BDT-based electron identification efficiency (upper panel) and data-to simulation correction factors (lower panel)
From [arXiv:2012.06888]



32 GeV Electron High Level Trigger Turn on Curve for Run D



- Barrel
- Endcap

Jets

- Jets can be misidentified as electrons or protons
- This background is calculated using a data driven method
- P(jet misIDed as electron) $\approx$ P(jet misIDed as positron)
- Same sign and opposite sign have same number of events and topographical features
- Shape and normalisation of SS region can be used to calculate the shape and normalisation of the SS+OS region

Jets – SS Method

- Start from tag and probe pairs in data with the same requirements as detailed in the section above
- Require same sign pairs
- Require that probe fails the HEEP ID
- Subtract all MC from the data, with the MC also having the same sign and probe failing HEEP ID requirement. In any cases where the MC is greater than the data, set that bin value to 0.
- Multiply this result by two to account for both the same sign and opposite region to jets
- Attribute this result to jets

Scales and Smearings

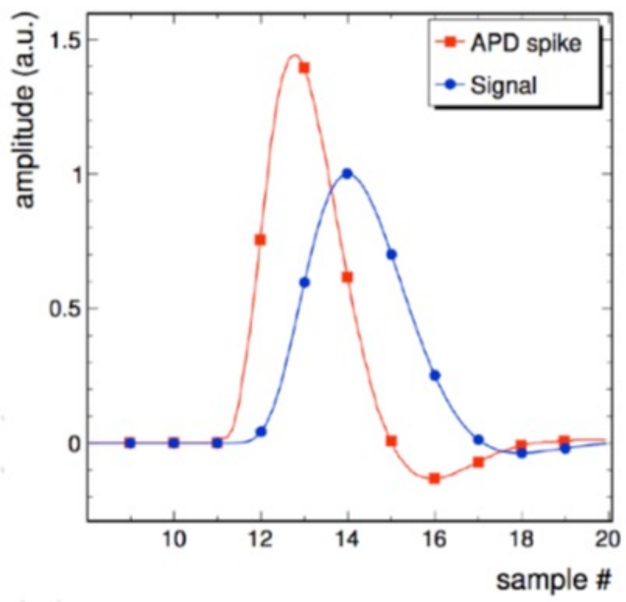
- Due to imperfections in the MC simulation, there are small differences between the data and MC for both their resolution and response.
- The data also has conditions that vary over time, causing a slight change in response while the MC conditions are static. To correct for both of these effects, energy corrections known as scale and smearing are applied.
- Scaling:
  - Data scaled to match MC response
  - Dependent on run, eta and shower shape
- Smearing:
  - MC smeared to match data resolution
  - Dependent on eta and shower shape

Data and MC Samples

Physics Process	Generator	Cross section	Number of Events
Drell Yan	MC@NLO	6077	200000000
W+Jets to LNu	madgraphMLM-pythia8	53550	78981243
tt_bar to Semi Leptonic	powheg-pythia8	364.4	106724000
tt_bar to 2L2Nu	powheg-pythia8	87.3	355332000
ST top	powheg-pythia8	35.6	8507203
ST anti top	powheg-pythia8	35.6	8433998
GJets HT 100 to 200	madgraphMLM-pythia8	5036	10034997
GJets HT 200 to 400	madgraphMLM-pythia8	1128	33884844
GJets HT 400 to 600	madgraphMLM-pythia8	126.4	9022800
GJets HT 600 to inf	madgraphMLM-pythia8	41.24	8330226
WW	pythia8	118.7	15634000
WZ	pythia8	27.56	7889000
ZZ	pythia8	12.14	2706000

Pulse Shape (LD) Algorithm

- Uses a second order polynomial to distinguish between EM showers and spikes
- Tested 300 channels in software and 72 channels on hardware using MP7
- Small discrepancy between C++ and HLS implementations when the LD value is at the threshold
  - Occurs approximately 1 in 2650 events
  - Should still meet required >99% spike rejection

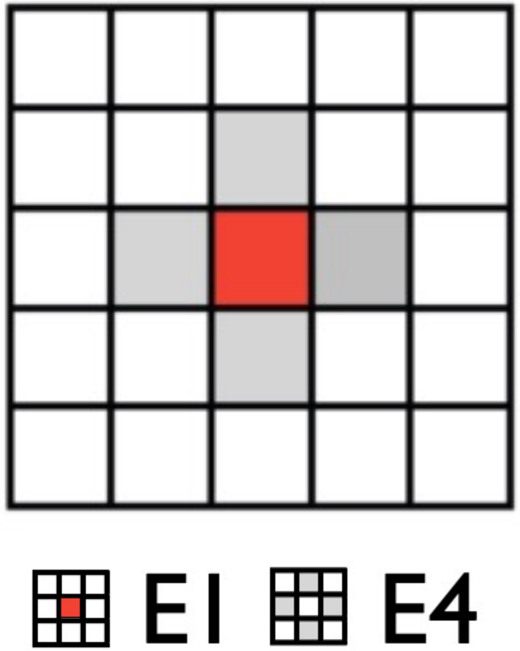


	DSP	FF	LUT
Utilisation (%)	16	8	13

Resource estimates for 300 channels on 1xKU115

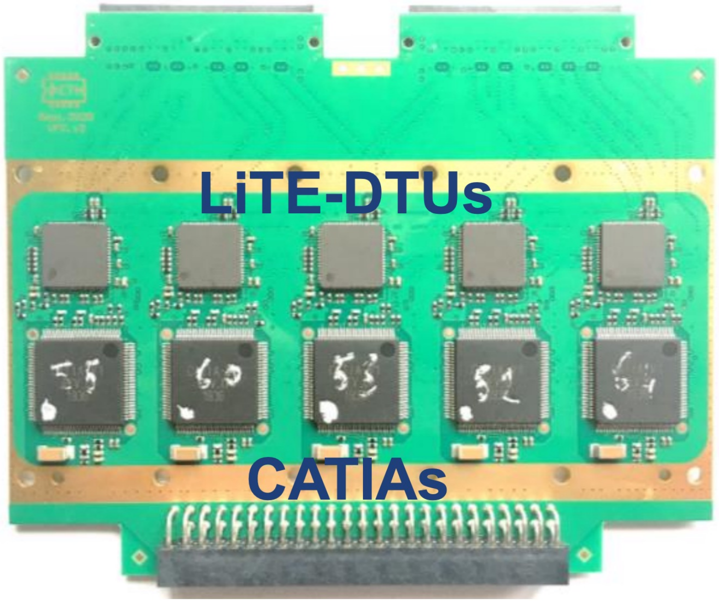
Swiss Cross Algorithm

- Previously used offline and in HLT
- Can now use in L1 due to single crystal granularity for trigger
- Uses the swiss cross variable $(1 - E_4/E_1)$
  - Cut placed at 0.95 separates spikes and EM showers with high efficiency
- Tested in software with no discrepancies over 2800 events
- HLS implementation currently being reworked to reduce resource usage



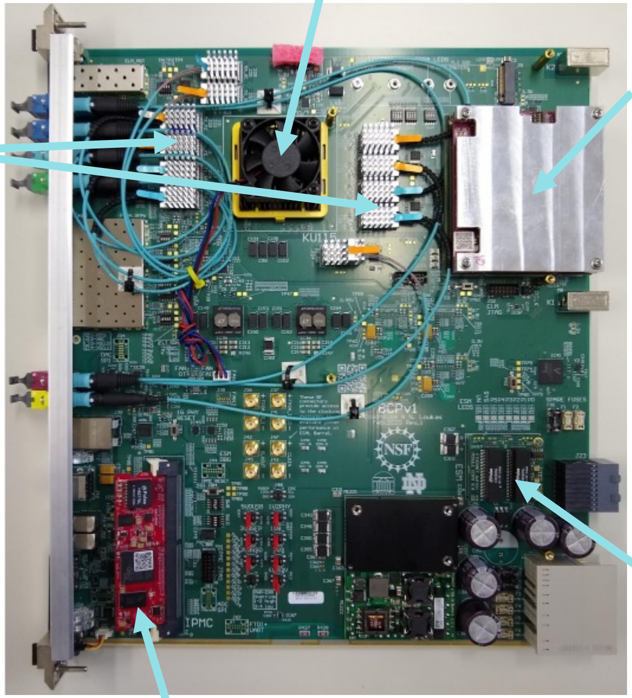
Frontend Electronics

- CATIA
  - Pre-amplifier ASIC
    - Minimal shaping to obtain best time resolution and spike rejection
  - Beam test to tune absolute gain in 2021
  - v1r4 submitted in November 2020
  - v2r0 submission foreseen in June
- LiTE-DTU
  - Data conversion, compression and transmission ASIC
    - ADC Sampling at 160 MHz
  - v1.2 being tested with CATIA v1r0 on VFE
  - v2.0 submission planned in July
    - Feature list frozen and implementation underway



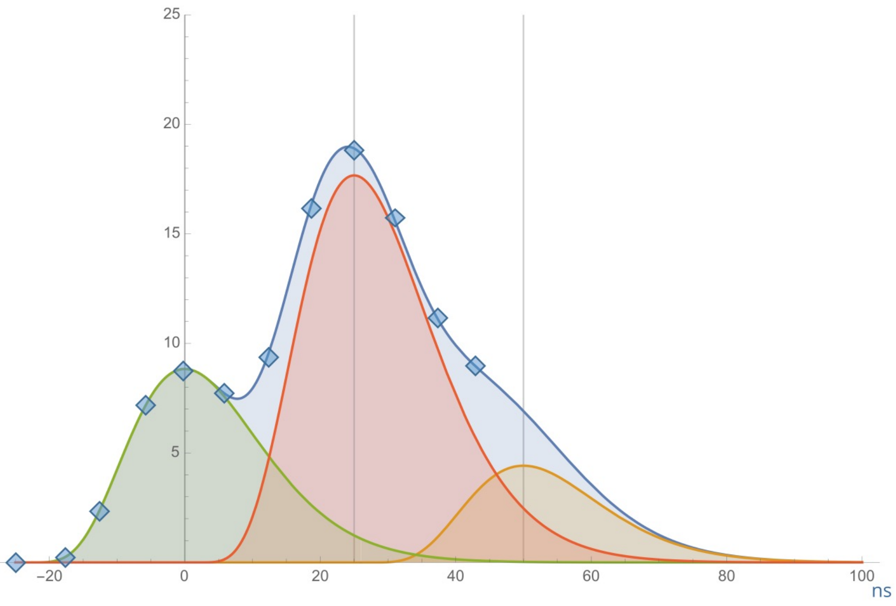
Backend Electronics

- Barrel calorimeter processor (BCP)
  - Each board handles signals from 600 crystals
- BCPv1 is currently being tested
  - Uses one KU115 FPGA
  - Integration tests largely finished
  - Currently working on simulation of ECAL Barrel synchronisation
- BCPv2 will use one VU13P instead of 2xKU115 originally envisioned
  - Provides nearly 3 times the memory, 30% more logic cells and 11% more digital signal processing



Other algorithms

- Multifit algorithm
  - See CMS DN-2020/012 for full details
  - Does both amplitude and timing reconstruction
  - Template fit with multiple components contributing to the signal
  - Method has been used in Phase I for HLT and offline reconstruction
  - Successful test of HLS using CMSSW

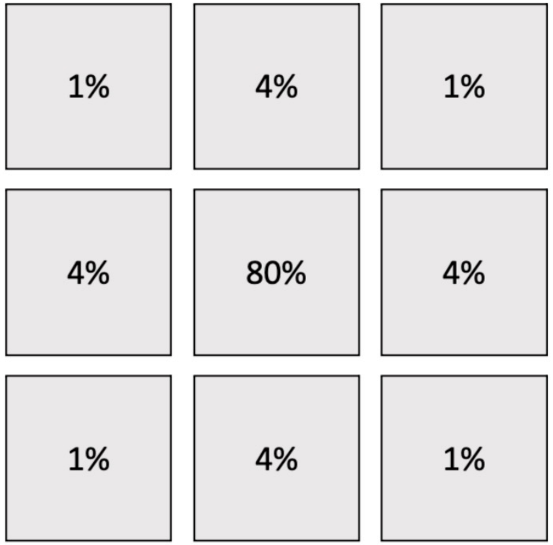


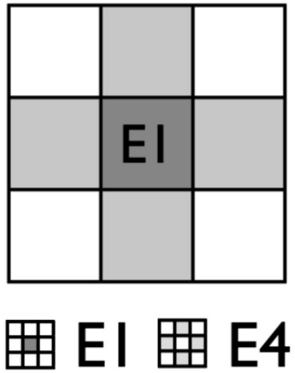
	DSP	FF	LUT
Utilisation (%)	54	21	35

Resource estimates for 300 channels on 1xKU115

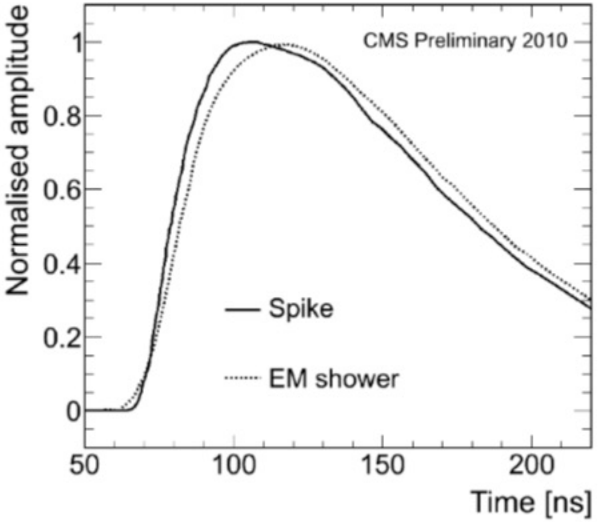
Spikes

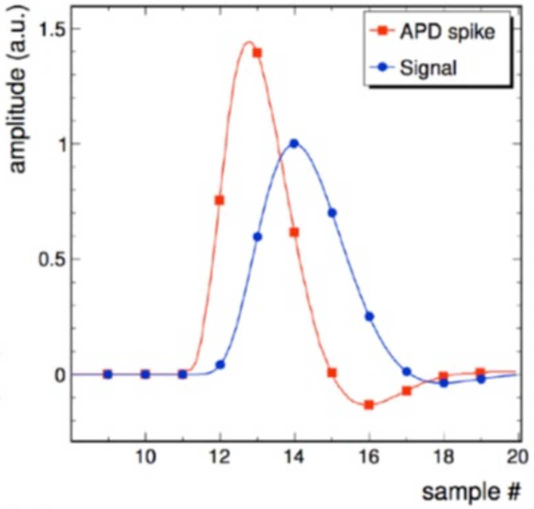
- Spikes occur when particles impact directly on the APDs in the ECAL barrel giving a large signal
  - Presents issues for triggering
  - Need >99% spike rejection in Phase II





Swiss-cross = 1-(E4/E1)





LD Algorithm Formulae

$$LD = R_+ - \sum_{i=0}^{O} p_i R_-^i$$

Where O is the order of the polynomial, pi is the polynomial weights and

$$R_+ = \frac{a_{n+1}}{a_n}$$

$$R_- = \frac{a_{n-1}}{a_n}$$

The a are the three consecutive sample, with an being the peak. The weights are chosen such that the scintillation pulses are set around 0, which gives spikes an LD value around -0.5 due to the difference in shape.

$$100a_{n+1}a_n - LD_{thr}a_n^2 - p_0 a_n^2 < p_1 a_{n-1}a_n + p_2 a_{n-1}^2$$

Example Clustering Algorithm Output

Event from H4 test beam (5x5 matrix) mapped on to BCPv1 (15x20 xtal) region

```
Crystal Et matrix with neighbours. Zero index in top left corner for each block
               0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
               0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0

   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   1   1   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   1   9  16   3   1   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   2  35 387   6   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   1   6  92   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   1   1   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0
   0   0       0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0       0   0

               0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
               0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
```

Output from 3x3 clustering algo

```
Processing TPClusterAlgoV1
This TP collection has size: 21
Adding TP cluster et=557, ieta=3, iphi=3, number of crystals=9, spike=0, swiss cross=0.614987
```