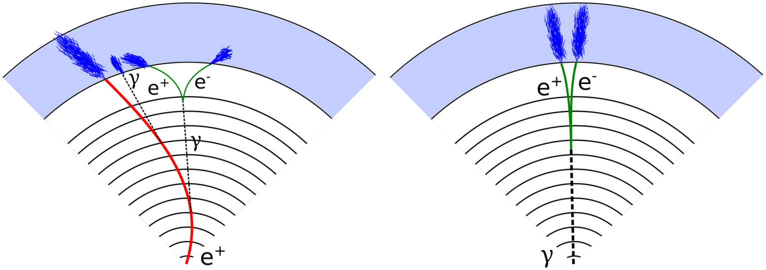
CMS Electron and Photon Performance at 13 TeV

Jonas Rembser - On behalf of the CMS Collaboration

CALOR 2018 - 18th International Conference on Calorimetry in Particle Physics
May 21-25, Eugene, USA

CMS Electron and Photon Performance at 13 TeV



Typical **electron** and **photon patterns** in a particle detector like CMS.

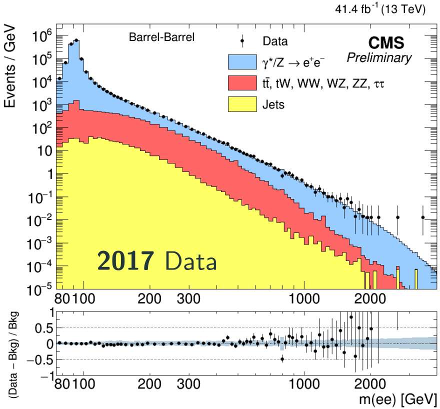
Electrons and **photons** essential for CMS physics, e.g.:

- Higgs decays to 4 leptons or two photons,
- BSM physics (heavy resonances, SUSY decays),
- SM measurements (top physics, multibosons, ...).

Goals of this talk:

- Overview on electrons and photons in CMS:
  - Energy cluster plus track **reconstruction**,
  - **Identification** techniques,
- Show 2017 performance and 2016 performance with improved calibrations.

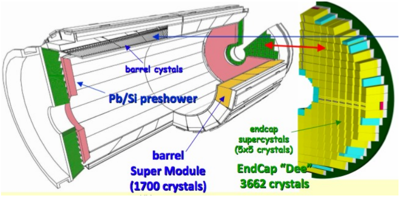
First Run II results for CALOR and first 2017 performance plots in general.



Seach for Z' in e^+e^- final state [1].

Electromagnetic calorimeter (ECAL) [2]

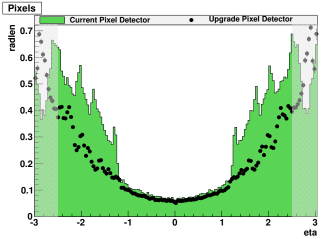
- 75848 quasi-projective scintillating PbWO4 crystals inside the 3.8 T magnet.
  - Barrel: 25.8 radiation lengths, 2.2 x 2.2 cm^2.
  - Endcaps: 24.7 radiation lengths, 2.86 x 2.86 cm^2.
- $\sim$3 X_0 lead scintillator preshower in endcaps.
- Up to $|\eta| < 3.0$, barrel-endcap transition at $|\eta| = 1.479$.



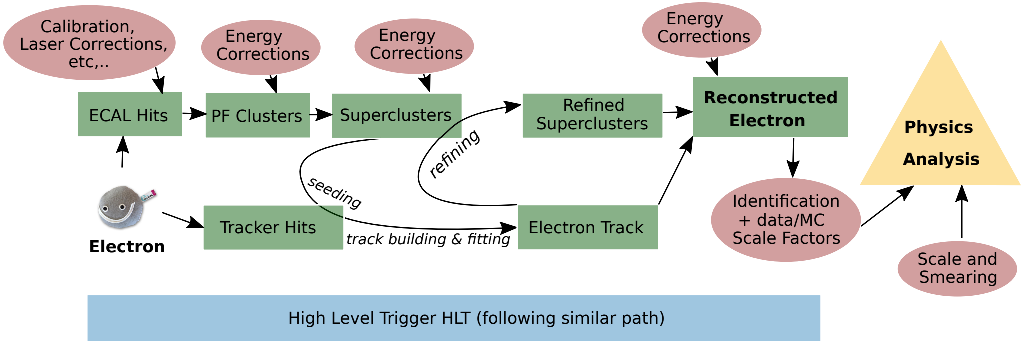
CMS ECAL.

Silicon tracker [3]

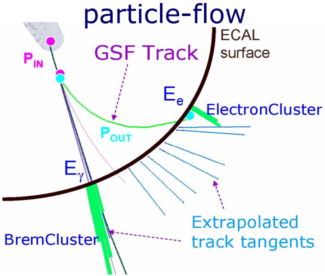
- **New pixel detector** in 2017 with added 4th layer and reduced material budget in the endcaps.
- Microstrip detector with 14 (12) layers in the barrel (endcap).
- Coverage up to $|\eta| < 2.5$, defining fiducial region for e/γ objects.



Material budget (X_0) of old and new phase I pixel detector [4].

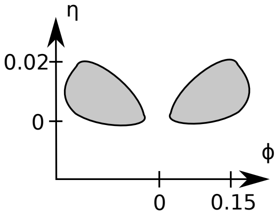


- Picture is similar for photons.
- Also seeding from Tracker hits matching clusters (complementing at low p_T, called "tracker driven").
- *Refining:* adding clusters matching extrapolated tangents at tracker layers, recovering bremsstrahlung or conversions.



Clustering

- Collect single particle-like clusters to clusters that look like EM clusters.

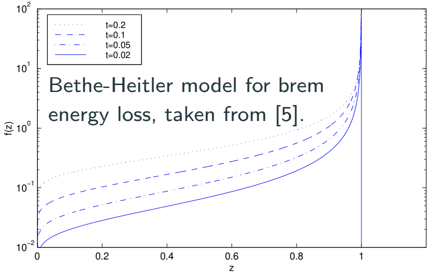


The ECAL hits follow a "mustache" pattern due to magnetic field.

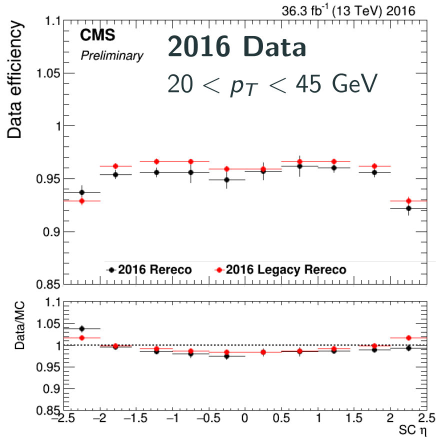
Interplay between the clustering and tracking algorithms, e.g. for pattern recognition of (converted) brem photons.

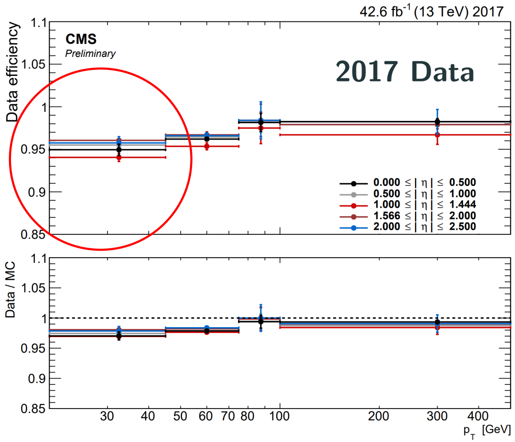
Tracking

- Electron tracking is peculiar due to large radiative losses, which bend track in ϕ.
- Radiative energy loss far from Gaussian.
- Solution: GSF (Gaussian Sum Filter) for electron tracks instead of Kalman Filter.



- Important for photons as well (conversions).

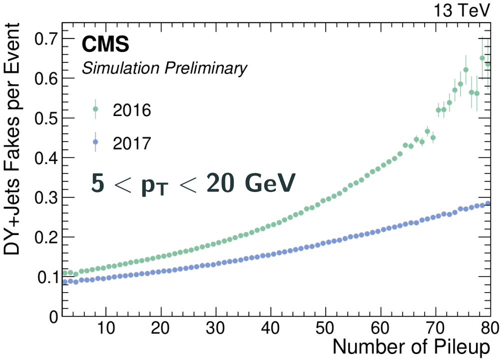




- Tag and probe measurement.
- **Efficiency** $\approx$ **96**% consistently.
  - Slight increase in the endcaps for 2017.
  - Increased efficiency in most recent reconstruction of 2016 data.

Unmatched[1] reconstructed **electron candidates** in DY+Jets MC **2016 vs. 2017**.

- In 2017, the fakes rate is lowered due to the **new pixel detector**, by 30%.
  - Pixel 4th layer/quadruplet seeding reduces number of fake tracks.
- Note: fake rate after ID step is again more comparable.

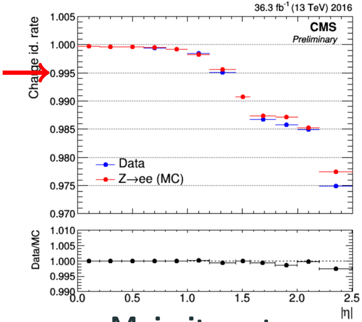


[1] Not matched to a generated electron with $\Delta R = 0.1$.

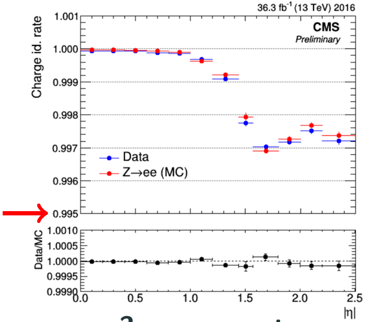
Electron Charge Identification

- Three methods:
  1. GSF track curvature.
  2. KF track curvature.
  3. Supercluster - GSF track $\Delta\phi$ at vertex.

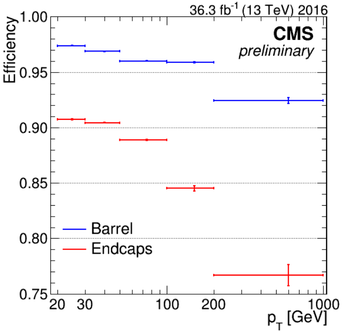
- Probes from $Z \rightarrow ee$ passing **loose ID cuts**.
- Charge misidentification growing with p_T.
- Note: charge mis-ID rate decreased after electron ID due to suppression of e.g. converted brems.



Majority vote

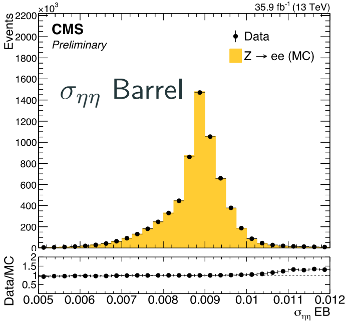


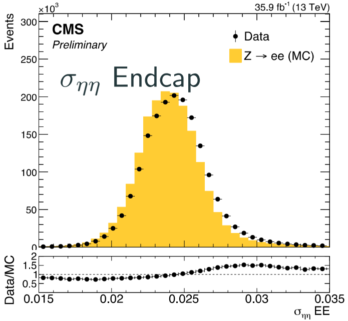
3-agreement



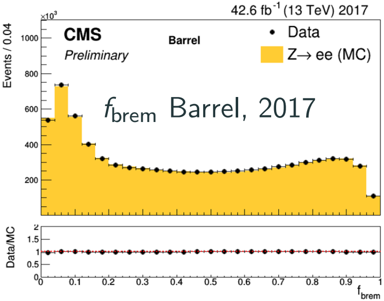
3-agreement efficiency

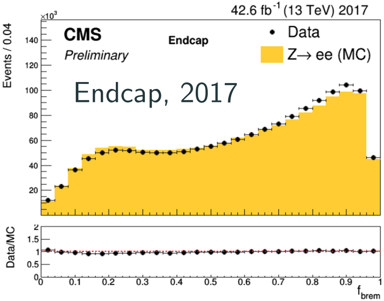
- Standard e/γ identification recipes in CMS (including data/MC scale factors):
  - **Cut based** IDs with several working points for general use.
  - **MVA based (Boosted Decision Trees)**, particularly for maximum separation at low p_T.
- Variables grouped in:
  1. **Shower shape** (e.g. $\sigma_{\eta\eta}$ on the right),
  2. **Track variables** (i.e. quality and brem fraction),
  3. **Track-cluster matching**,
  4. **Conversion ID**,
  5. **Isolation**.



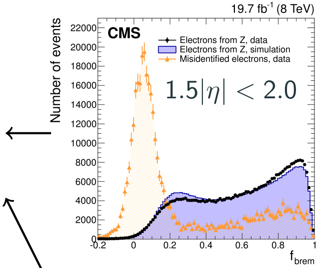


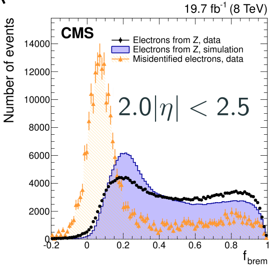
[2] See slide 25 for more details on variables





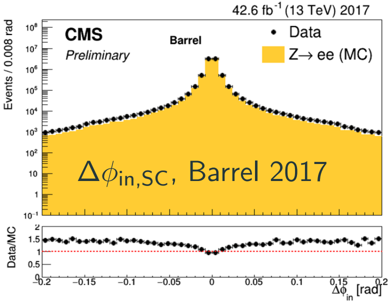
Run I Endcaps

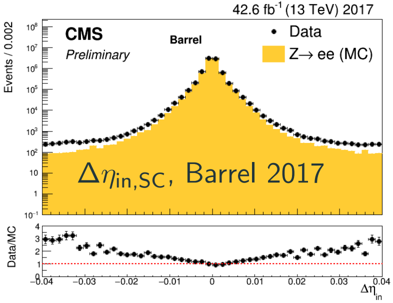




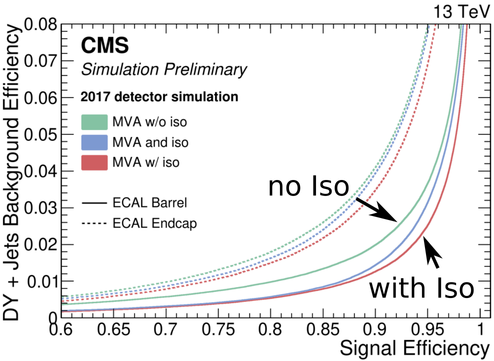
- $f_{brem} = (p_{in} - p_{out})/p_{in}$, fraction of the **momentum lost to bremsstrahlung in tracker**.
- **Excellent tool to access the material budget** and compare it over data and MC.
- Discrepancies in Run I data/MC ⇒ mismodelling of the material budget ⇒ corrected for in Run II.

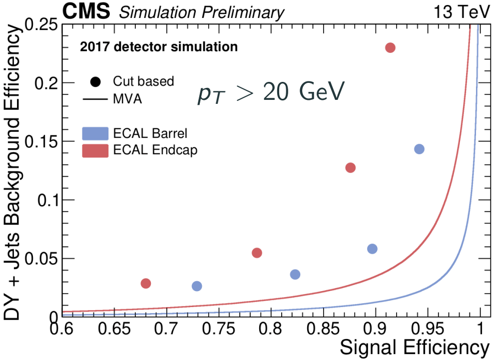
- **GSF Tracking** allows us to extrapolate original electron trajectory (pre-brem) to the calorimeter $\Rightarrow$ precise track matching:
- Very good Supercluster η resolution allows a **tight cut on $\Delta\eta$**
  - Track resolution important.
- Supercluster ϕ resolution worse due to brem, **looser cut on $\Delta\phi$**.





- ROC curves for the electron MVA ID and the cut-based selection working points[3]:



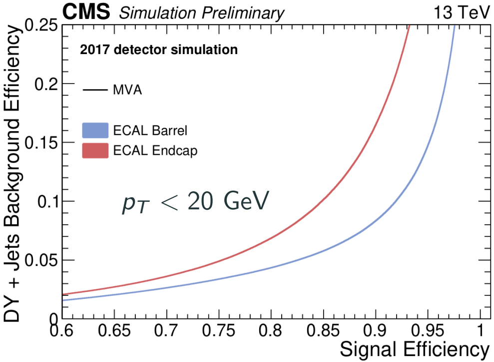


- **New MVA ID** in 2017 **includes PF isolation components**:
  - Better background rejection than the traditional sequential approach at high efficiencies.
- **New pixel** with **less material budget** at $|\eta| > 1.2$:
  - Reduced fake rate in endcaps by 20 % (at 90 % efficiency).

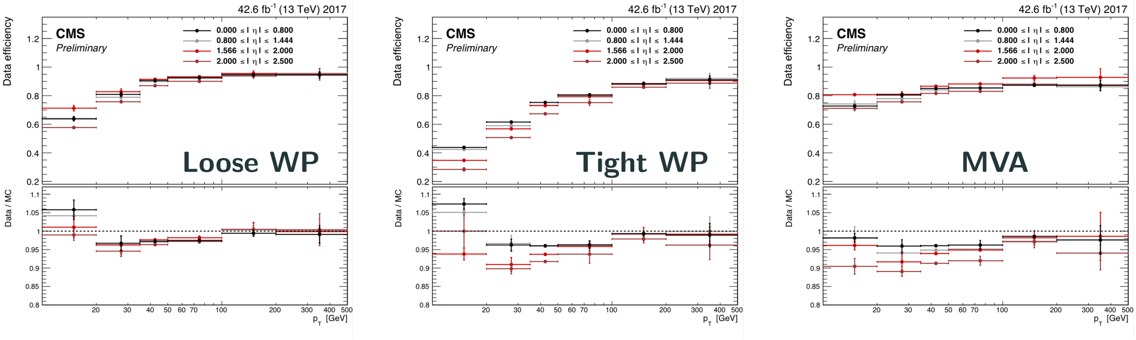
[3]See slide 28 for $t\bar{t}$+jets instead of DY+jets background.

Electron Identification ROC Curves for low p_T

- MVA ID for electrons **performant at low $\mathbf{p_T}$**, allowing for multi-lepton analyses with high efficiency.
- Mainly fakes from pileup in endcaps at low p_T:
  - **New pixel** reduces fakes: > 20 % effect.
  - **Isolation** components **in MVA**: > 10 % effect.

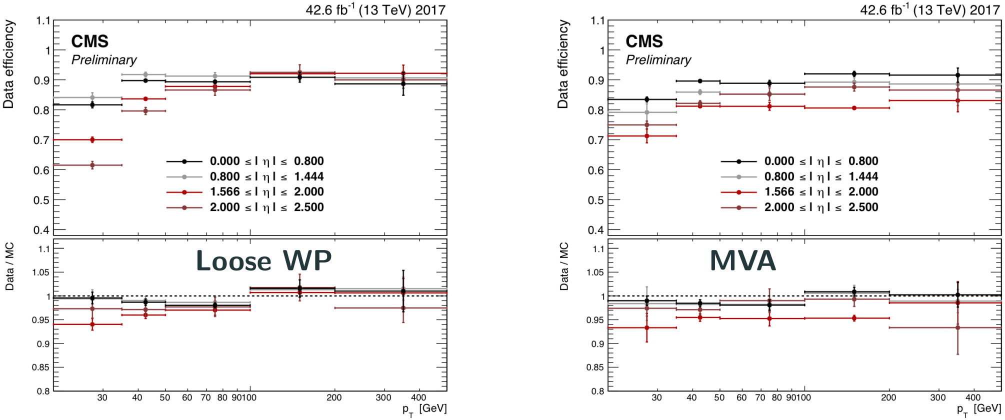


Measuring Selection Efficiencies - Electrons



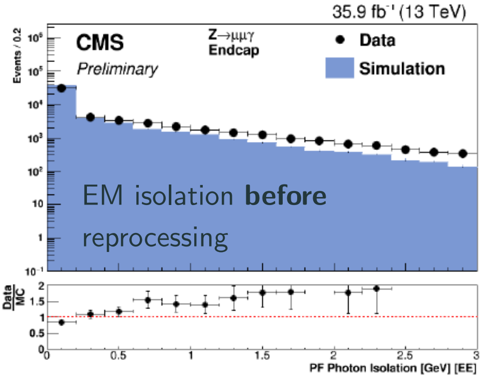
- **Tag n' probe** with $Z \to e^+e^-$, syst. uncertainties:
  - Background shape (expo times erf or expo only),
  - Signal shape (analytic or template from MC),
  - Monte Carlo Generator,
  - Tag Selection.
- Data/MC corrections up to 5 % barrel, 10 % endcap.
  - Minimal effect from end-of-year pixel issues: $\sim$ few percent.

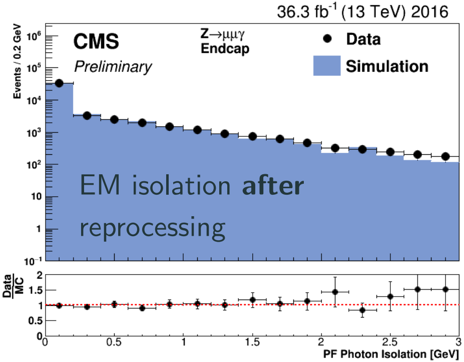
Measuring Selection Efficiencies - Photons



- $Z \to e^+e^-$ also used for photons:
  - Reconstruct electrons as photons.
  - Don't apply electron veto.
- Data/MC corrections up to 5 %.

Reprocessing of 2016 Data

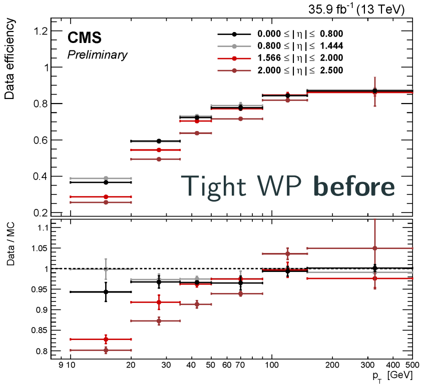


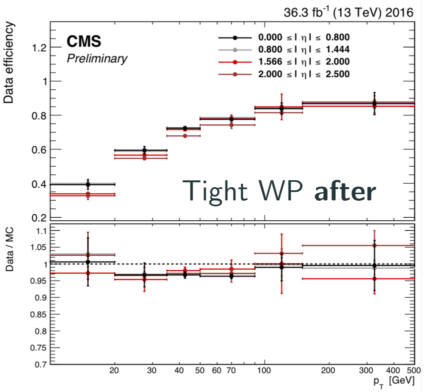


Reprocessed 2016 data with improved calibration and reconstruction:

- Major change: **improved ECAL pedestals and calibrations**.
- Data/MC agreement improved in almost all cases.
  - Example of **electromagnetic particle flow isolation** for photons.

Electron cut based identification before and after 2016 data reprocessing:

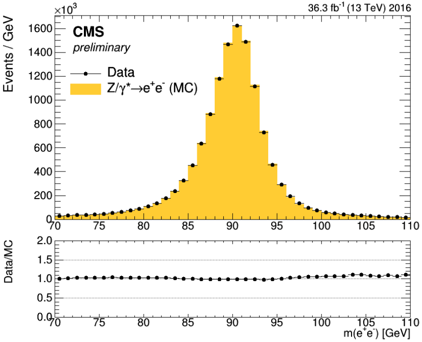


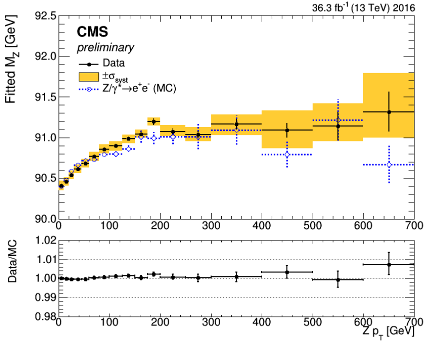


- Data/MC agreement substantially improved due to final calibrations applied.

$Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-\gamma$ Invariant Mass with Full Energy Corrections

$Z \rightarrow e^+e^-$ mass distribution, electrons satisfy $p_T > 25(20)$ GeV and medium cut based ID.



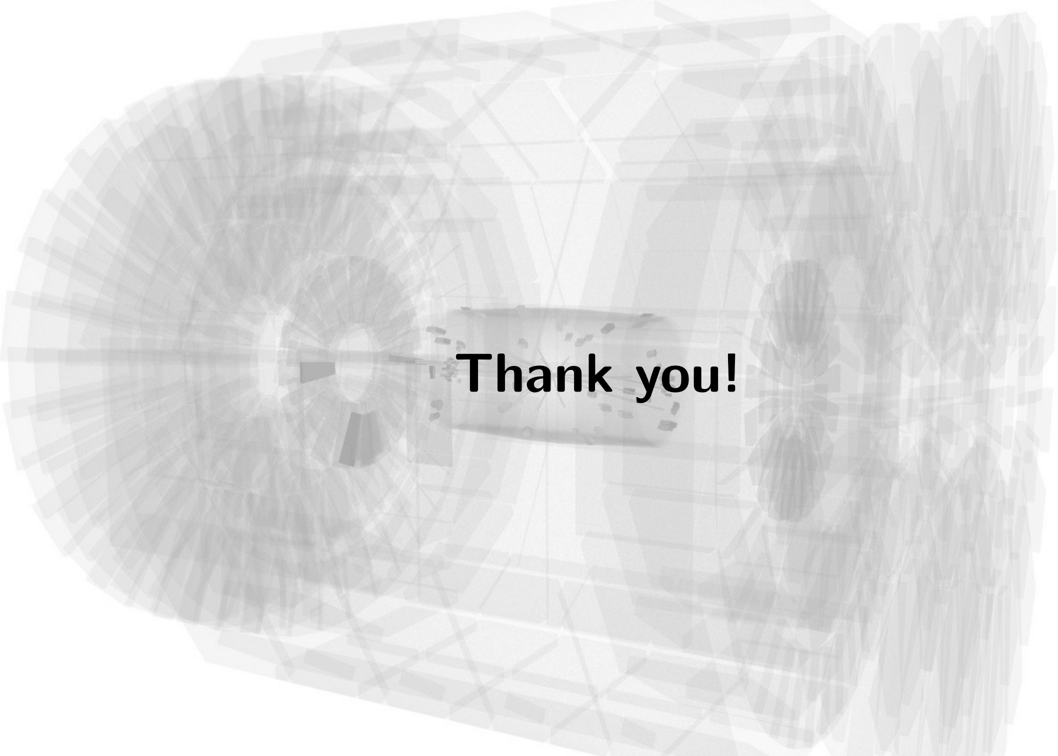


- Fit with Breit-Wigner ⊛ Crystal Ball (81 to 101 GeV) to obtain mass scale shift and resolution ($\approx$ 1.8 GeV).
- Fitted Z mass as a function of Z boson transverse momentum.
  - Excellent stability vs Z p_T.

Resolution very comparable with Run I.

Conclusion

- Presented **first CMS e/γ performance results with 2017 data** (initial caibration) and the 2016 recalibration.
- 2016 **recalibration significantly improves data/MC** agreeement.
  - Demonstrates highlevel benfits of **good understanding of low detector calibrations**.
- Data/MC agreement in **2017** very **comparable** to **2016**, despite challenging data taking conditions.
- Good electron and photon performance in Run II, benefiting from:
  - **Complex e/γ reconstruction algos** integrated with particle flow.
  - Monte Carlo simulations where the material budget of the inner detector is well modeled,
  - **Sophisticated MVA techniques** to ensure good background separation down to low p_T.



Thank you!

[1] "Search for high mass resonances in dielectron final state," Tech. Rep. CMS-PAS-EXO-18-006, CERN, Geneva, 2018.

[2] *The CMS electromagnetic calorimeter project: Technical Design Report*.
Technical Design Report CMS, Geneva: CERN, 1997.

[3] V. Karimäki, M. Mannelli, P. Siegrist, H. Breuker, A. Caner, R. Castaldi, K. Freudenreich, G. Hall, R. Horisberger, M. Huhtinen, and A. Cattai, *The CMS tracker system project: Technical Design Report*.
Technical Design Report CMS, Geneva: CERN, 1997.

[4] A. Dominguez, D. Abbaneo, K. Arndt, N. Bacchetta, A. Ball, E. Bartz, W. Bertl, G. M. Bilei, G. Bolla, H. W. K. Cheung, M. Chertok, S. Costa, N. Demaria, D. D. Vazquez, K. Ecklund, W. Erdmann, K. Gill, G. Hall, K. Harder, F. Hartmann, R. Horisberger, W. Johns, H. C. Kaestli, K. Klein, D. Kotlinski, S. Kwan, M. Pesaresi, H. Postema, T. Rohe, C. Schäfer, A. Starodumov, S. Streuli, A. Tricomi, P. Tropea, J. Troska, F. Vasey, and W. Zeuner, "CMS Technical Design Report for the Pixel Detector Upgrade," Tech. Rep. CERN-LHCC-2012-016. CMS-TDR-11, Sep 2012.
Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.

[5] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, "RESEARCH NOTE FROM COLLABORATION: Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC," *Journal of Physics G Nuclear Physics*, vol. 31, pp. N9–N20, Sept. 2005.

[6] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, I. Puljak, C. Rovelli, R. Salerno, and Y. Sirois, "Electron reconstruction in CMS," Tech. Rep. CMS-NOTE-2006-040, CERN, Geneva, Feb 2006.

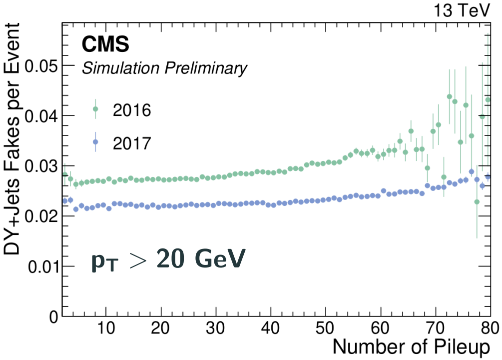
[7] "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV," *Journal of Instrumentation*, vol. 10, p. P06005, June 2015.

[8] "Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV," *Journal of Instrumentation*, vol. 10, p. P08010, Aug. 2015.

[9] CMS Collaboration, "Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV," *ArXiv e-prints*, Apr. 2018.

Unmatched[4] reconstructed **electron candidates** in DY+Jets MC **2016 vs. 2017**.

- In 2017, the fakes rate is lowered due to the **new pixel detector**, by approximately 15% at $p_T > 20$ GeV.
  - Pixel 4th layer/quadruplet seeding reduces number of fake tracks.
- Note: fake rate after ID step is again more comparable.



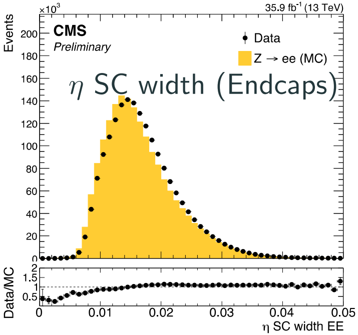
[4]Not matched to a generated electron with $\Delta R = 0.1$.

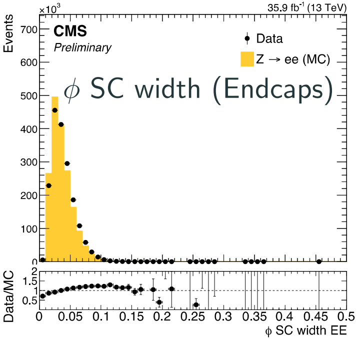
We use different types of shower shape variables:

- **Supercluster variables**: we use the SC widths in η and ϕ.

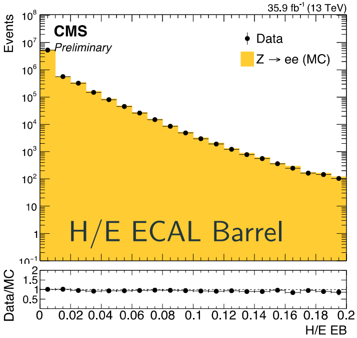
Drawback: superclusters are "forced" to look like e/γ energy deposits, as initital clusters look like single-particle clusters. Better derive variables from assumption-inependent clusters:

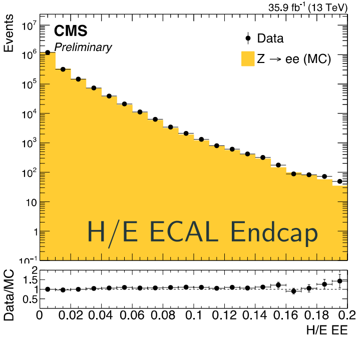
- **5x5 shower shapes**: natural choice as the energy of an e/γ is almost fully contained in 5x5 crystal arrays.





- **5x5 shower shapes**: other 5x5 shower shape variables we use are:
  - Energy ratios like $R_9 = \frac{E_{3x3}}{E_{5x5}}$, a good indicator of photons converting before the ECAL,
  - covariances involving ϕ, like $\sigma_{\phi\phi}$ and $\sigma_{\eta\phi}$, although there are less powerful due to the spread in ϕ of the EM objects.
- **Longitudinal shower shapes** which we don't have much in CMS:
  - Less significant: the preshower energy over the ECAL energy.
  - H/E, the ratio of the hadronic energy to EM energy:



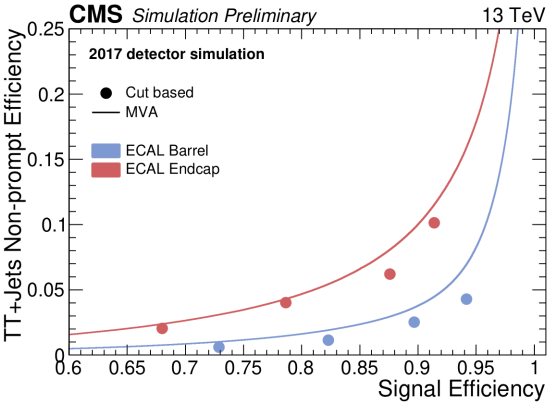


H/E is affected by pileup and by energy dependent leakage through ECAL gaps.

$\Rightarrow$ Scaling cut $H/E < c + a \cdot \rho + c \cdot E$, pileup variable ρ, energy E.

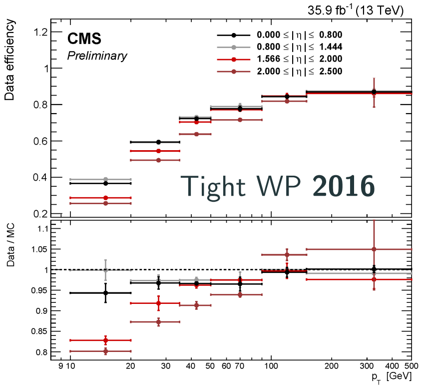
Other Variables

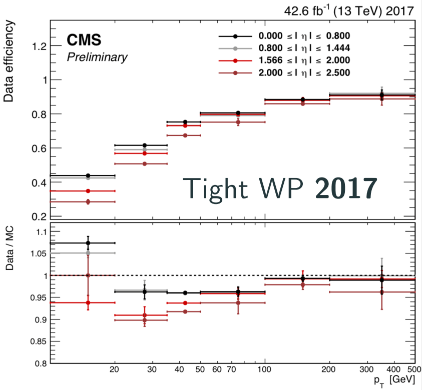
- **Track variables**:
  - Number of hits and χ^2 of GSF and KF track
  - f_{brem}
- **Track-cluster matching**:
  - E/p_{in}, E/p_{out} and similar
  - Geometrical distances
- **Conversion ID**:
  - Conversion vertex fit probability
  - Number of missing expected hits
- **Isolation**:
  - PF isolation components (neutral hardron, charged hadron and EM)
  - Combined PF isolation with pileup correction



- Signal from Drell-Yan+Jets Monte Carlo and non-prompt electrons from tt+Jets Monte Carlo as background. The cut-based selection was trained for background in tt+Jets events, while the MVA selection (which has different use cases) was trained for background in DY+Jets events.

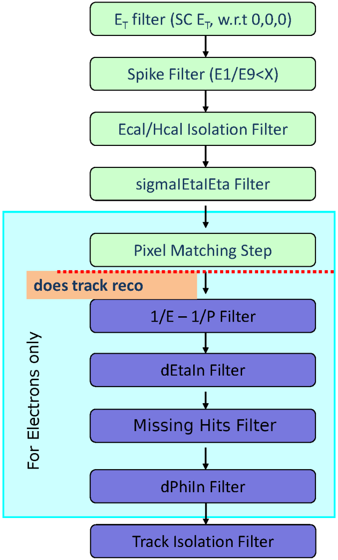
Electron cut based identification in **2016 after reprocessing** (left) and 2017 (right):

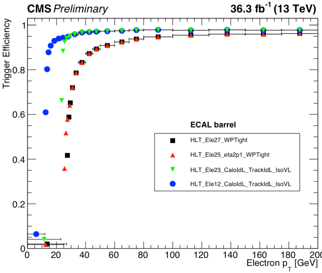


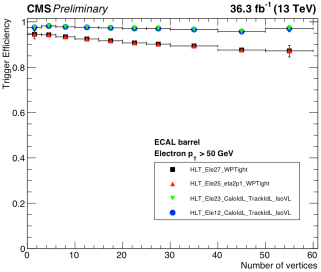


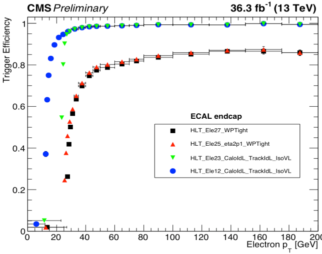
- Similar data /MC agreement to 2016 initial reconstruction

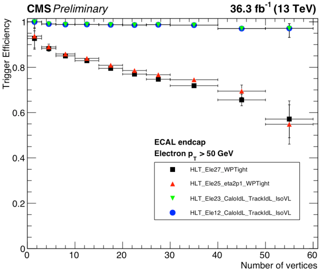
- Reducs rate from 100 kHz from Level 1 trigger to about 800 Hz.
- Fully implemented in software
- In general a tradeoff between efficiency rate and computing resources.
- For electrons and photons, there exist several HLT single (double) electon/photon paths which follow the reconstruction/identificaion up to different steps.

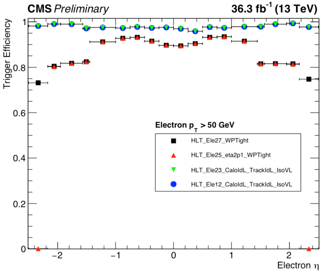












HLT trigger efficiencies vs. p_T, η and the number of vertices for the full 2016 dataset.

Pixel Detector: DCDC converter issue

- DCDC converters providing LV (VA/VD) for pixel detector modules started failing on October 5, 2017
  - ⇒ about 50 broken DCDC converters in total (5 % of the detector, BPIX+FPIX)
  - Affected channels during data-taking larger (up to 10 %) single power-cycling for module SEU recovery was only done rarely as precaution
- For E/gamma reconstruction and identification, the overall effect on efficiencies is small
- This effect is not in the simulation and loss due to this issue would show up in the data/MC scale factors presented in this talk
  - the scale factor deviation from 1 indicates the maximum possible effect
  - however the scale factors also result from other effects apart from DCDC converter issue
    - given that the electron scale factors are fully affected by this and the photon scale factors are negligibly affected, the reader can see from the plots in this talk the scale of the issue
  - from comparing periods with and without this DCDC issue, **this effect is estimated to be a few percent** depending on the tightness of the selection for ID and very little effect for reconstruction
- the main effect in the ID is worsening of variables which depend on precise tracker resolution, such as the track matching plots shown earlier
  - not all of the disagreement is due to this but a significant component is