Search for signatures of large extra dimensions in high-mass diphoton events using proton-proton collisions at $\sqrt{s} = 13$ TeV with CMS

Andrew Buccilli (University of Alabama)
on behalf of the CMS Collaboration

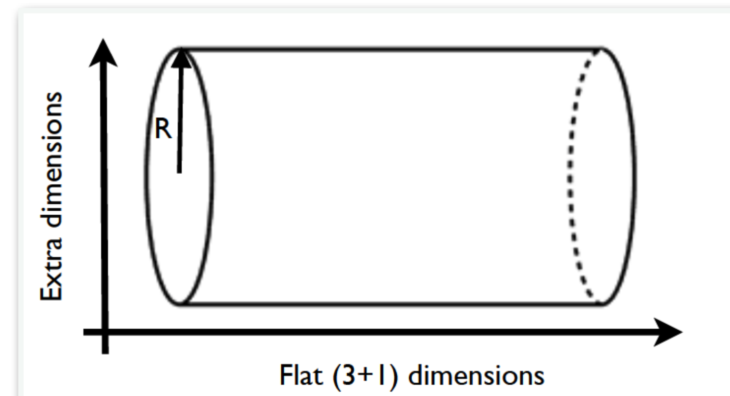
University of Alabama
Experimental Particle Physics Journal Club
26 April 2018

Motivation

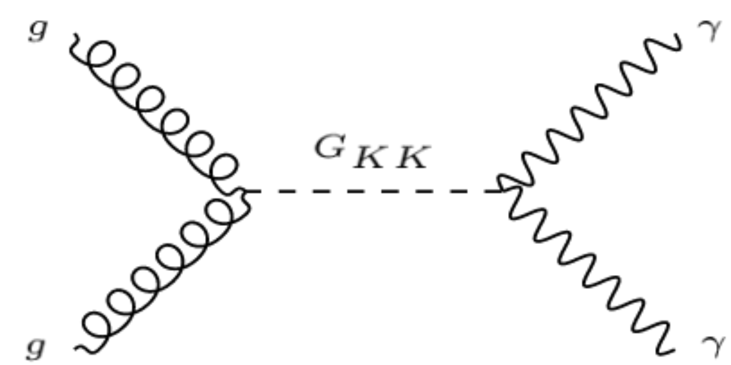
- Hierarchy problem in SM arises from ratio of:
  - Planck scale $M_{Pl} \sim 10^{19}$ **GeV**
  - Electroweak symmetry breaking scale $M_{EWSB} \sim 10^{2}$ **GeV**

$$M_{\mathrm{Pl}}{}^2 \sim M_{\mathrm{Pl}(4+n)}^{2+n} R^n$$

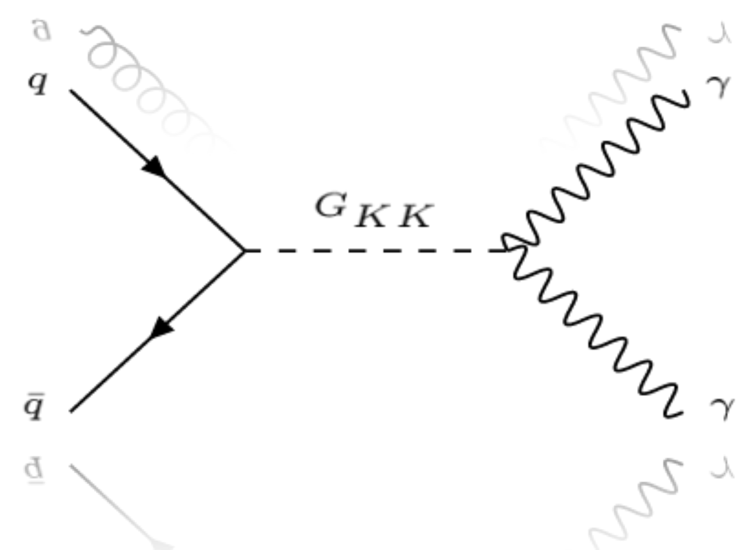
- Arkani-Hamed, Dimopoulos and Dvali (ADD) model of large extra dimensions proposes
  - *n* additional spatial dimensions compactified with radius *R*
  - allows the *fundamental Planck scale* $M_{Pl(4+n)} \sim 1$ **TeV** with observed $M_{Pl} \sim 10^{19}$ **GeV** if $R \sim 10^{30/n-19}$ **m**



- Produces Kaluza–Klein (KK) modes of the graviton *G*, which can decay to 2 photons
  - ADD model gives non-resonant signal enhancement over expected SM background to $m_{\gamma\gamma}$



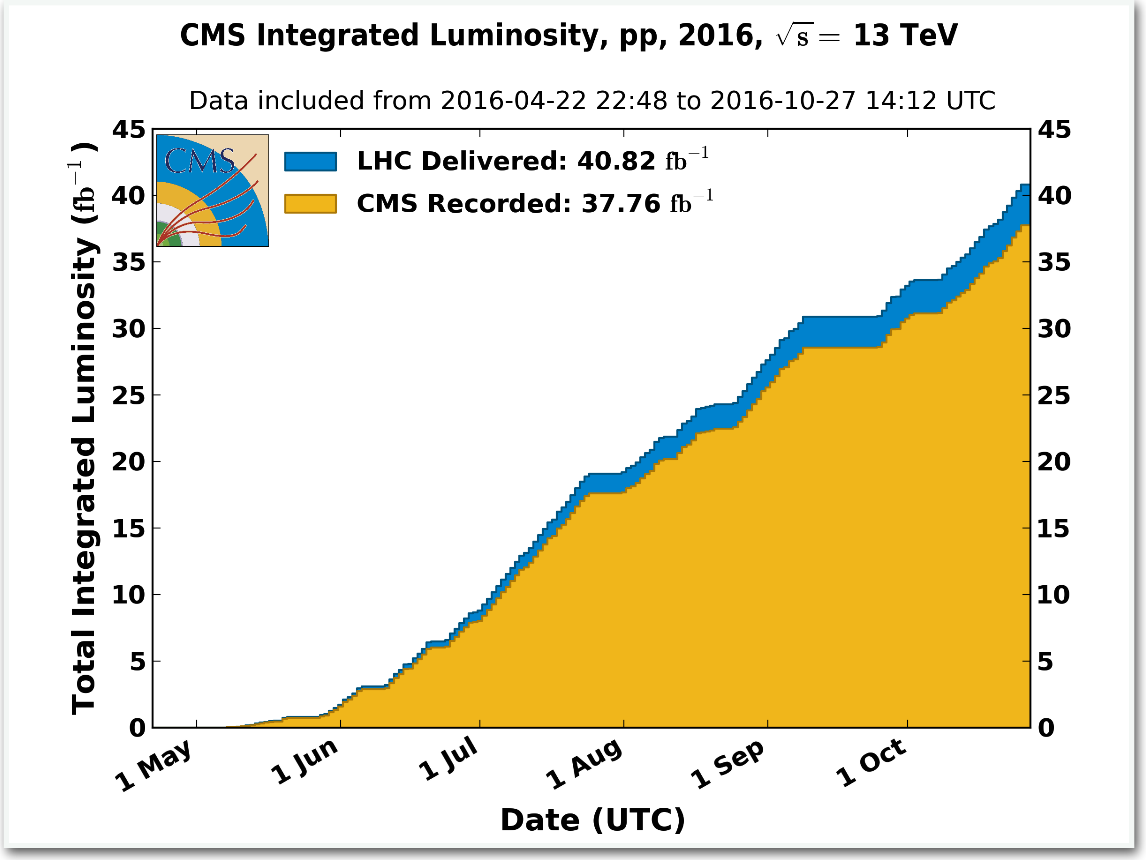
- $G_{KK} \rightarrow \gamma\gamma$
  - CMS detector provides excellent energy resolution for electromagnetic particles
  - Clean, isolated photon signature



- Previous CMS search used 2011 LHC data at $\sqrt{s}$ = 7 TeV

Analysis strategy

- Search for a non-resonant excess in the high-mass diphoton spectrum at $\sqrt{s}$ = 13 TeV using the full 2016 dataset corresponding to an integrated luminosity of 35.9 fb^{-1}



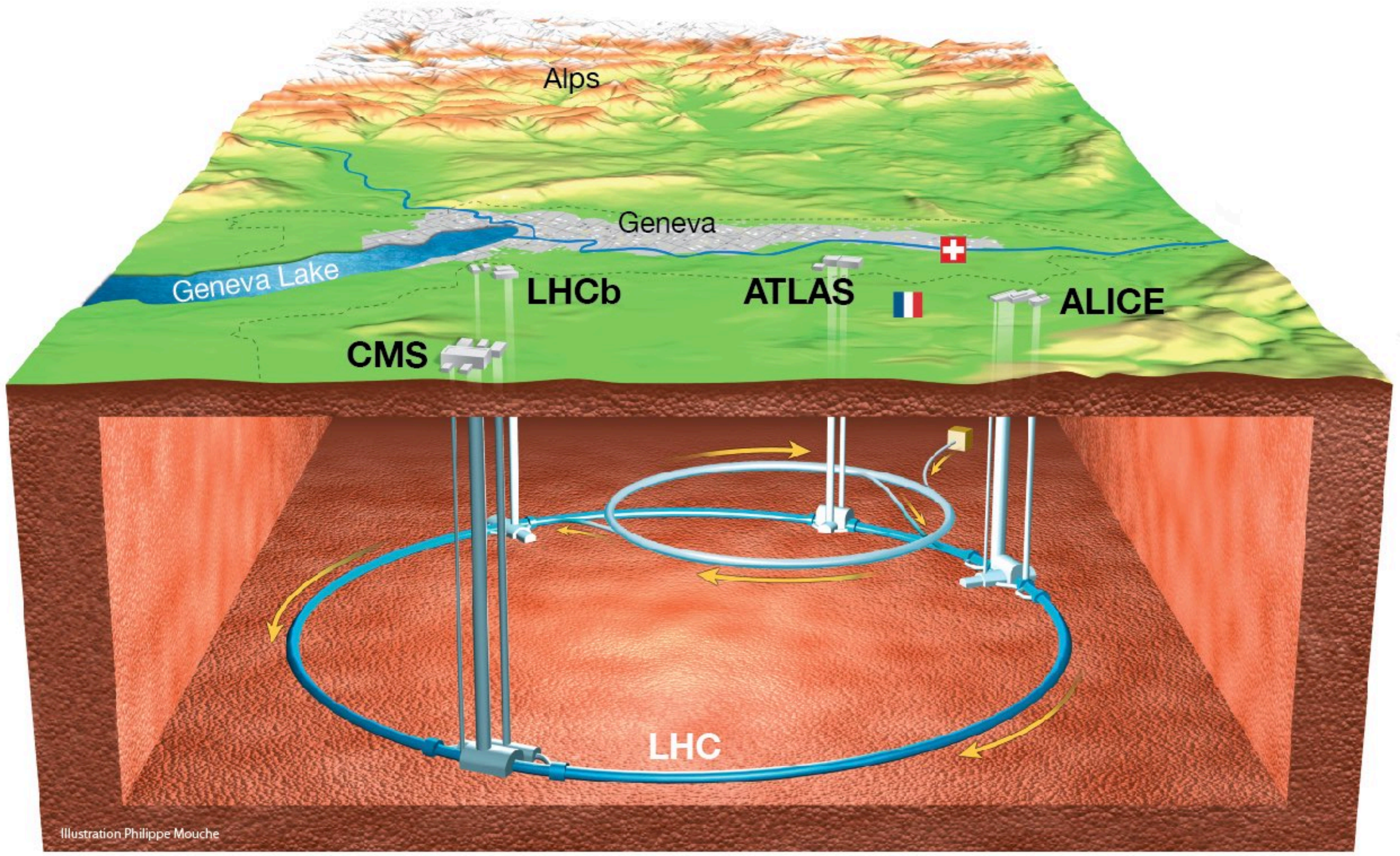
<u>Full background prediction is made:</u>

1. Dominant, irreducible prompt SM diphoton ($\gamma\gamma$) background
   - Next-to-next-to-leading order (NNLO) Monte Carlo (MC) calculation

2. Sub-dominant, reducible jet-faking-photon (γ**+*j*** / ***j*+*j***) background
   - Data-driven estimate

LHC

- Large Hadron Collider
  - 27 km circumference
  - ~100 m underground
- pp, pPb, PbPb, XeXe collisions
- 1232 superconducting magnets
  - 8.3 T
  - 15 m long
- proton bunches
  - 2808 bunches
  - 25 ns spacing
  - 10^{11} protons/bunch



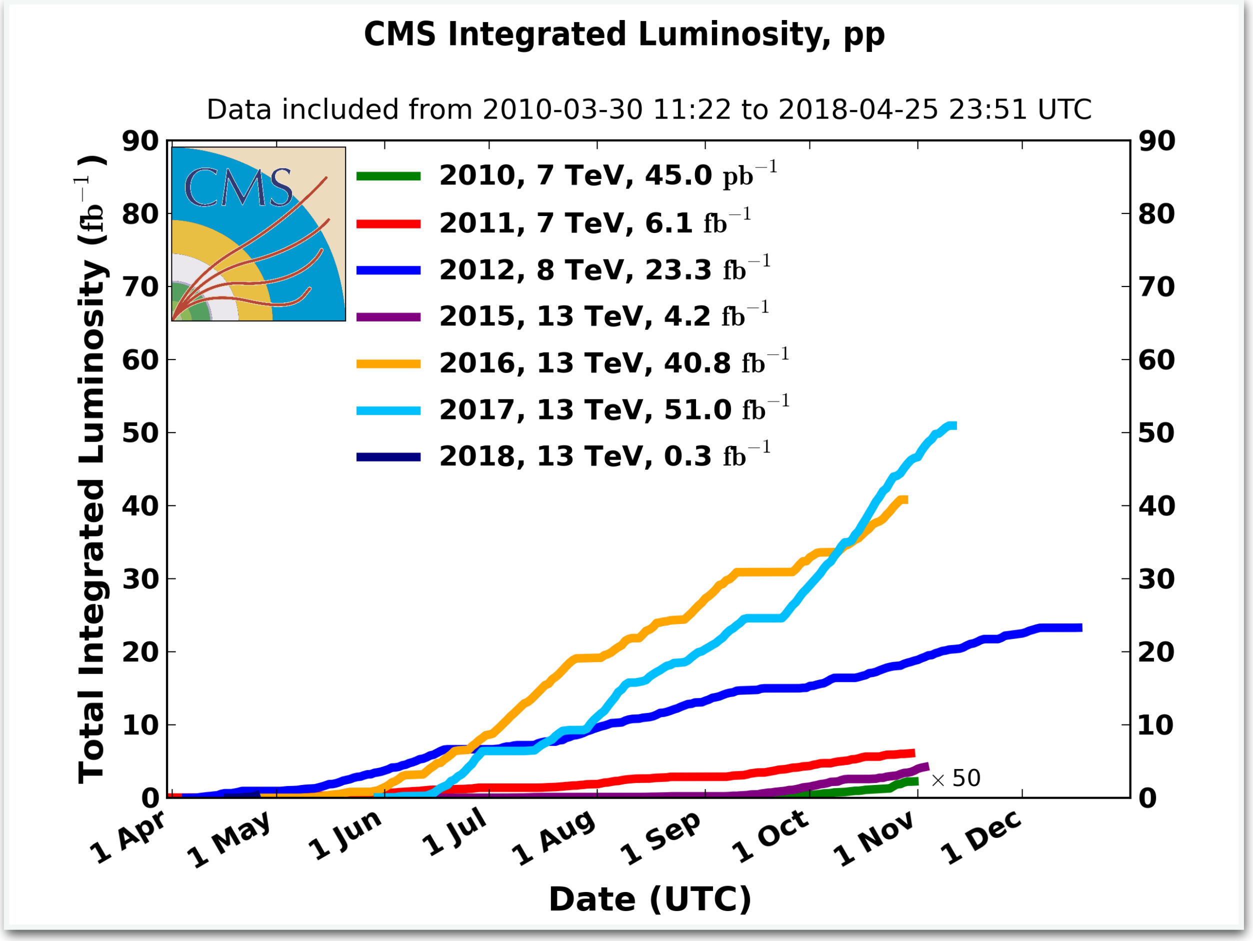
LHC schedule

LHC physics production schedule for pp collisions with integrated luminosity recorded by CMS

physics production run

long shutdown

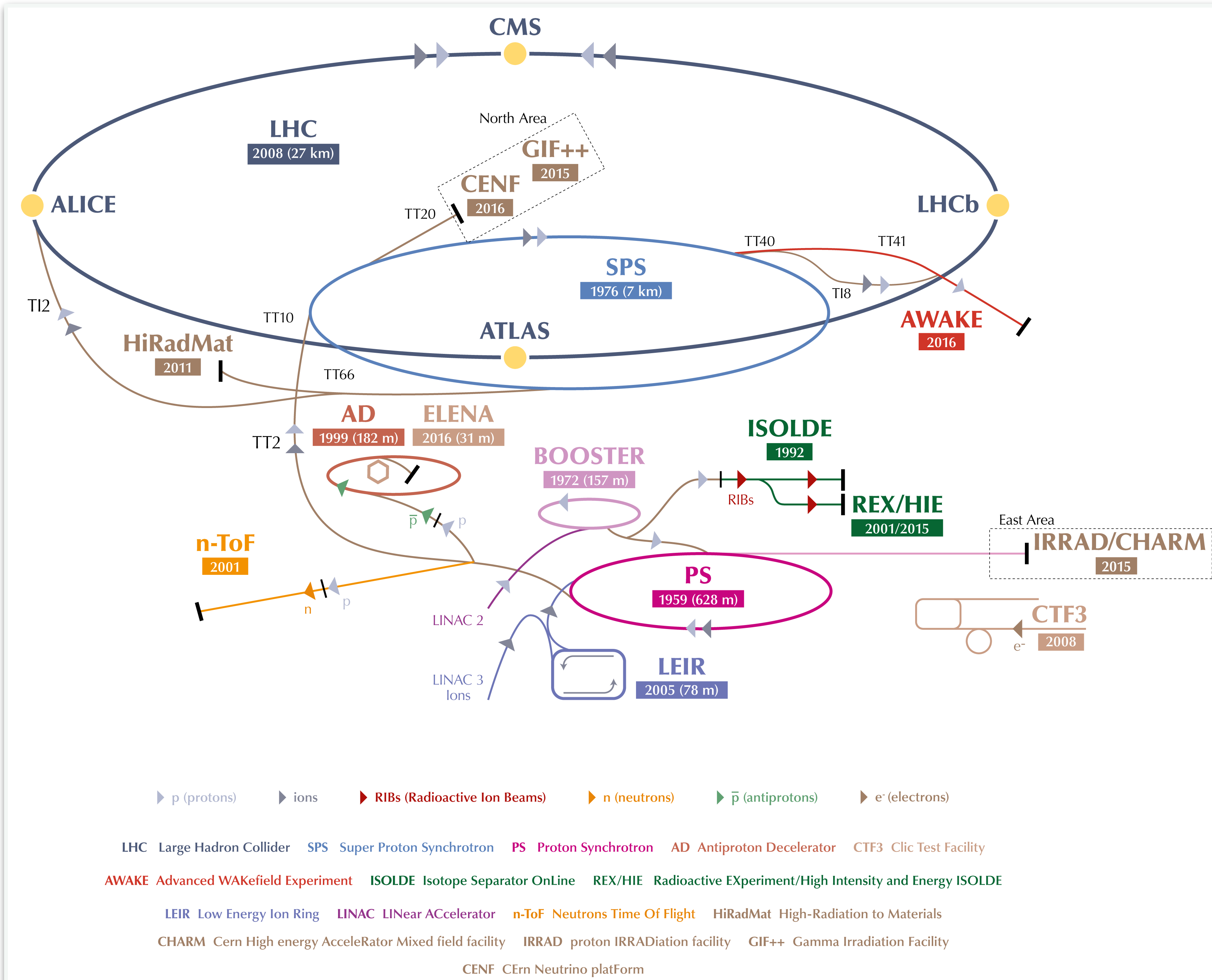
- **Run 1:** 2010-2012
  - 2010: $\sqrt{s}$ = 7 TeV, 41.5 pb^{-1}, 50 ns
  - 2011: $\sqrt{s}$ = 7 TeV, 5.6 fb^{-1}, 50 ns
  - 2012: $\sqrt{s}$ = 8 TeV, 21.8 fb^{-1}, 50 ns
- **LS1:** 2013-2014
- **Run 2:** 2015-2018
  - 2015: $\sqrt{s}$ = 13 TeV, 3.8 fb^{-1}, 50 ns -> 25 ns
  - 2016: $\sqrt{s}$ = 13 TeV, 37.8 fb^{-1}, 25 ns
  - 2017: $\sqrt{s}$ = 13 TeV, 46 fb^{-1}, 25 ns
  - 2018: $\sqrt{s}$ = 13 TeV, 60 fb^{-1} (TBD), 25 ns
- **LS2:** 2019-2020
- **Run 3:** 2021-2023
  - $\sqrt{s}$ = 14 TeV
- **LS3:** 2024-?
  - High-Luminosity Large Hadron Collider



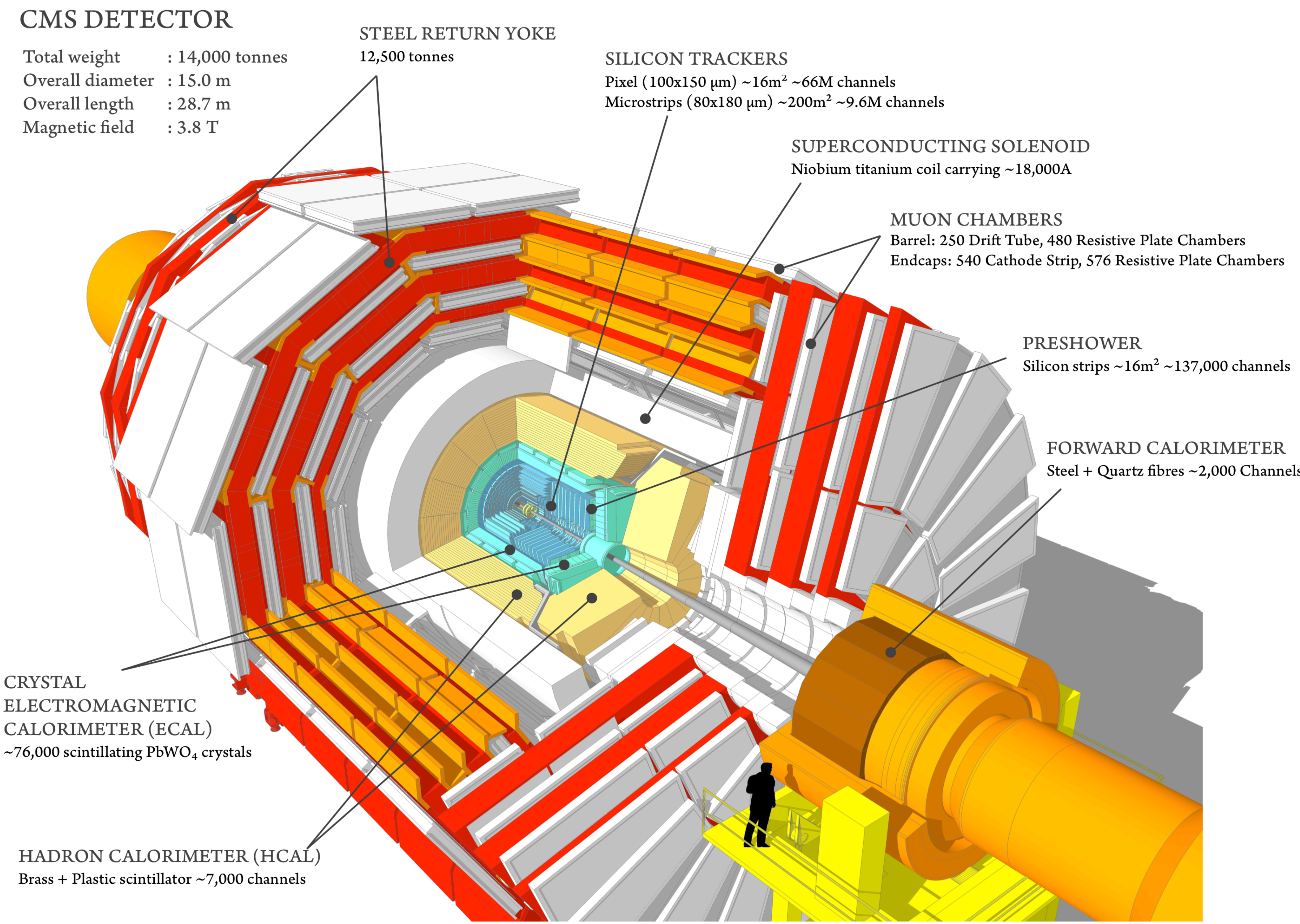
CERN accelerator complex

Proton acceleration chain:

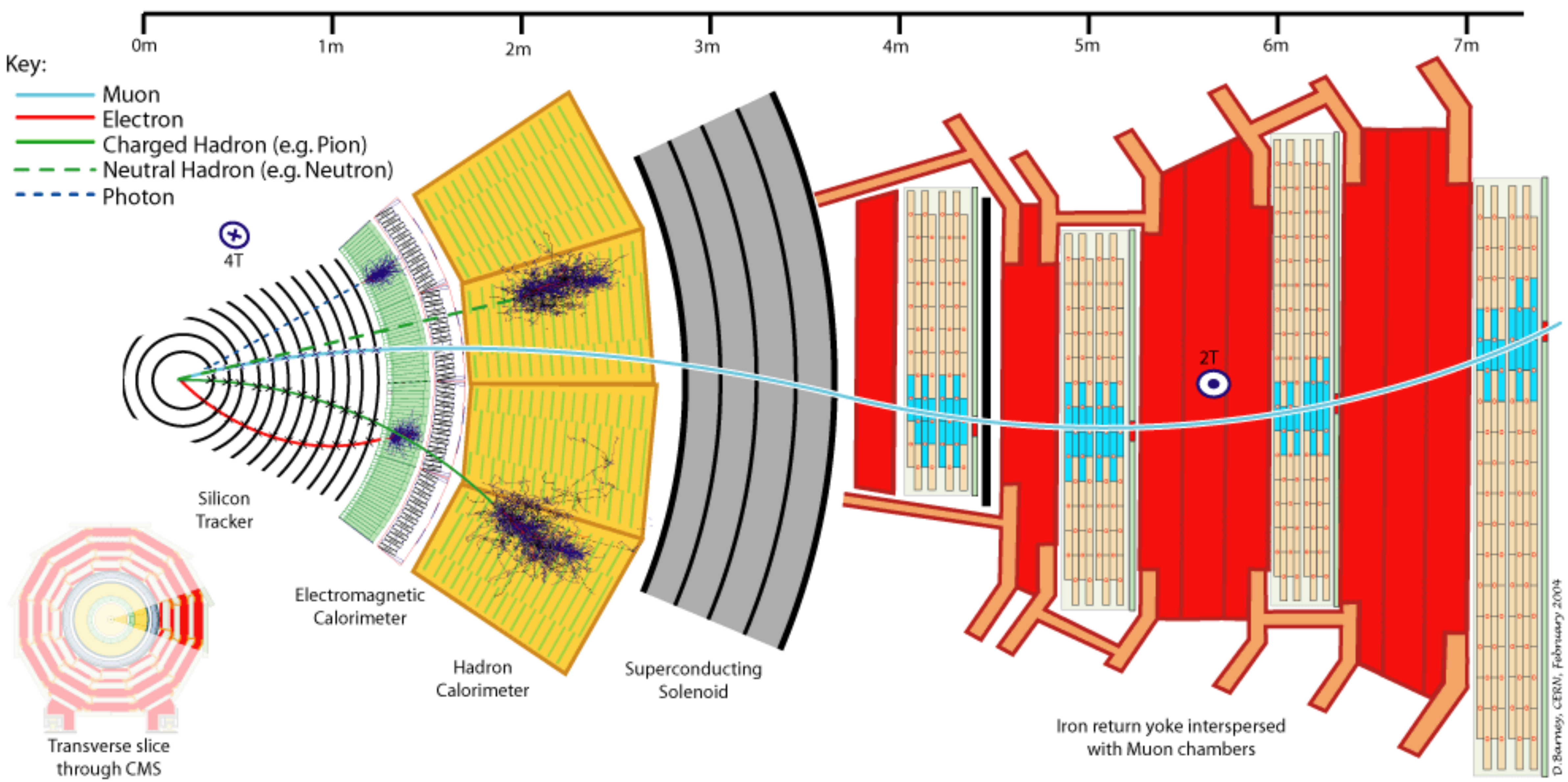
1. Linac 2 (50 MeV)
2. PSB (1.4 GeV)
3. PS (25 GeV)
4. SPS (450 GeV)
5. LHC (14 TeV)



CMS detector



CMS particle identification



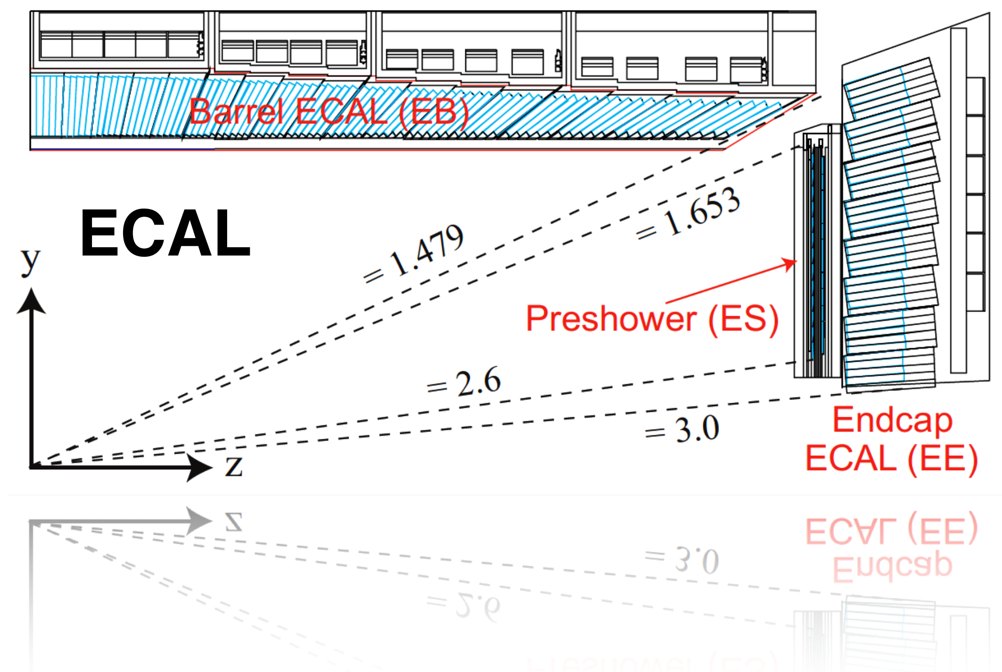
Event selection

- Photons are measured by the CMS electromagnetic calorimeter (ECAL)

- Trigger selection
  - Events with 2 photon candidates each with $p_T > 60$ GeV

- Kinematic selection
  - Photon $p_T > 75$ GeV (fully efficient trigger selection)
  - Diphoton invariant mass $m_{\gamma\gamma} > 500$ GeV
  - EBEB: two photons in EB
  - EBEE: one photon in EB and one in EE

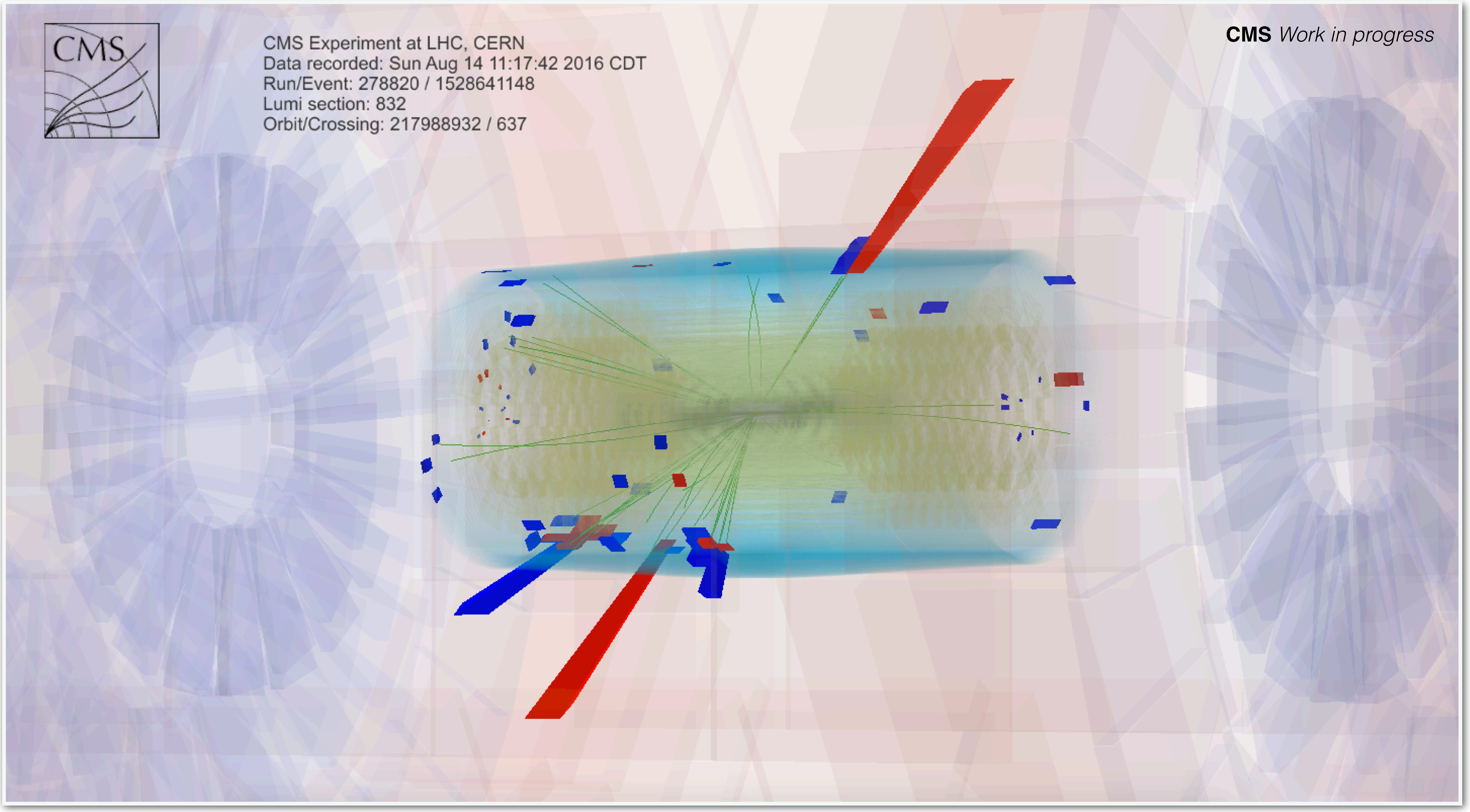
EB: ECAL barrel, $|\eta| < 1.44$
EE: ECAL endcap, $1.57 < |\eta| < 2.5$



- Photon identification tuned for high-p_T photons
  - Relies on
    - Isolation variables
    - Shower shape variables
    - Electron veto
  - 90(85)% efficient in EB(EE)

Event display

- High-mass diphoton event recorded on August 14, 2016



ADD signal model

- ADD model is parameterized by number of extra dimensions ***n_{ED}*** and UV string cutoff scale ***M_S***
- Interference effects
  - Signal generated with LO SM background using Sherpa v.2.1.1
    - Samples generated with $m_{\gamma\gamma} \leq M_S$
- M_S cutoff conventions $\mathcal{F}$
  - Consider **HLZ (n_{ED} =2)**, **GRW**, and **Hewett-**

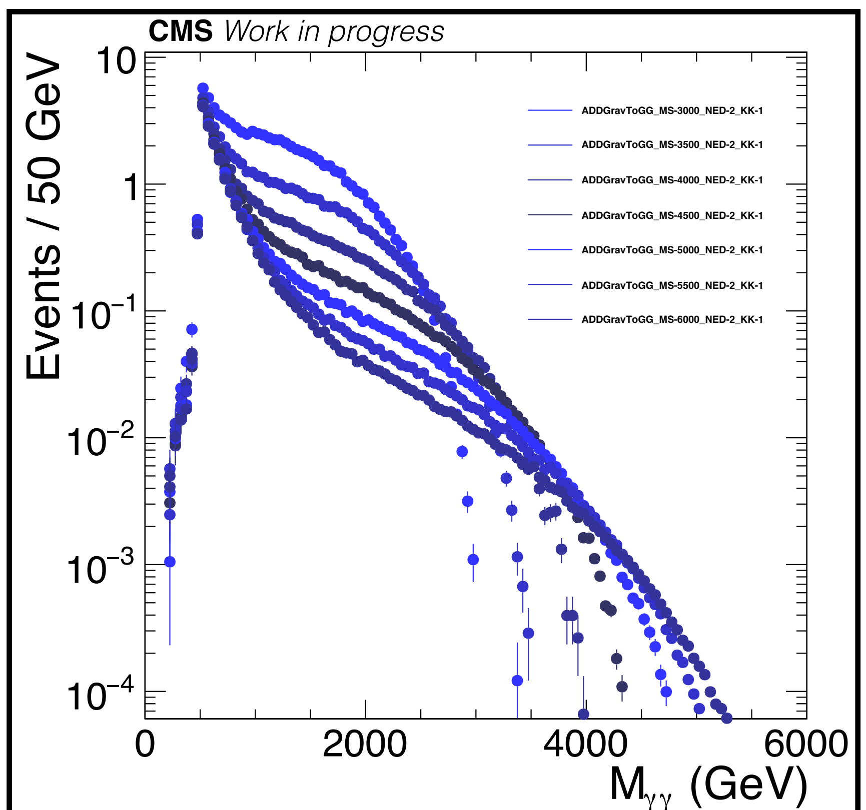
- Total cross section:

$$\sigma_{\text{total}} = \sigma_{\text{SM}} + \frac{\mathcal{F}}{M_S^4}\sigma_{\text{int}} + \frac{\mathcal{F}^2}{M_S^8}\sigma_{\text{ADD}}$$

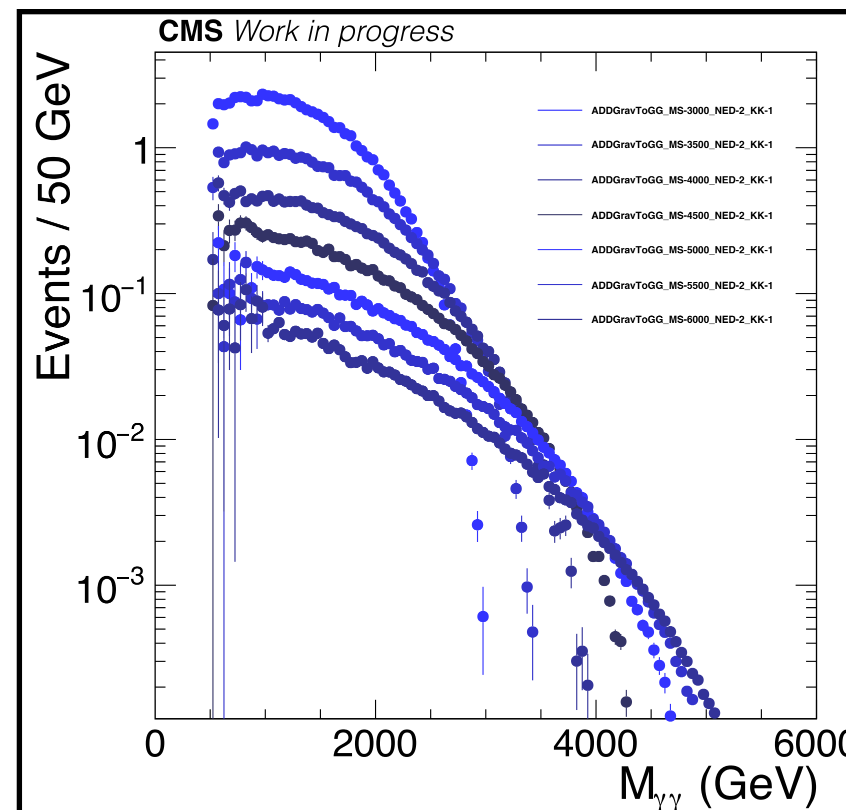
$$\mathcal{F} = \begin{cases} 1 \quad (\text{GRW}), \\ \log\left(\frac{M_S^2}{\hat{s}}\right), \text{ if } n_{ED} = 2 \\ \frac{2}{n_{ED}-2}, \text{ if } n_{ED} > 2 \\ \pm\frac{2}{\pi} \quad (\text{Hewett}), \end{cases} \quad (\text{HLZ}),$$

- Parameter space:
  - M_S = 3-11 TeV
  - n_{ED} = 2,3,…,7

HLZ (n_{ED} = 2)



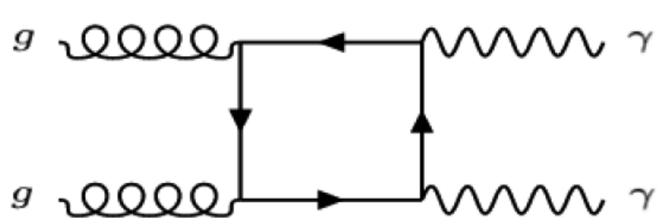
sig.+bkg.+int.



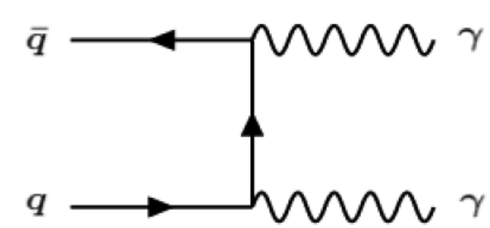
sig.+int.

Real background

- Irreducible, prompt SM $\gamma\gamma$ background generated using Sherpa v.2.1.1 with:



Box

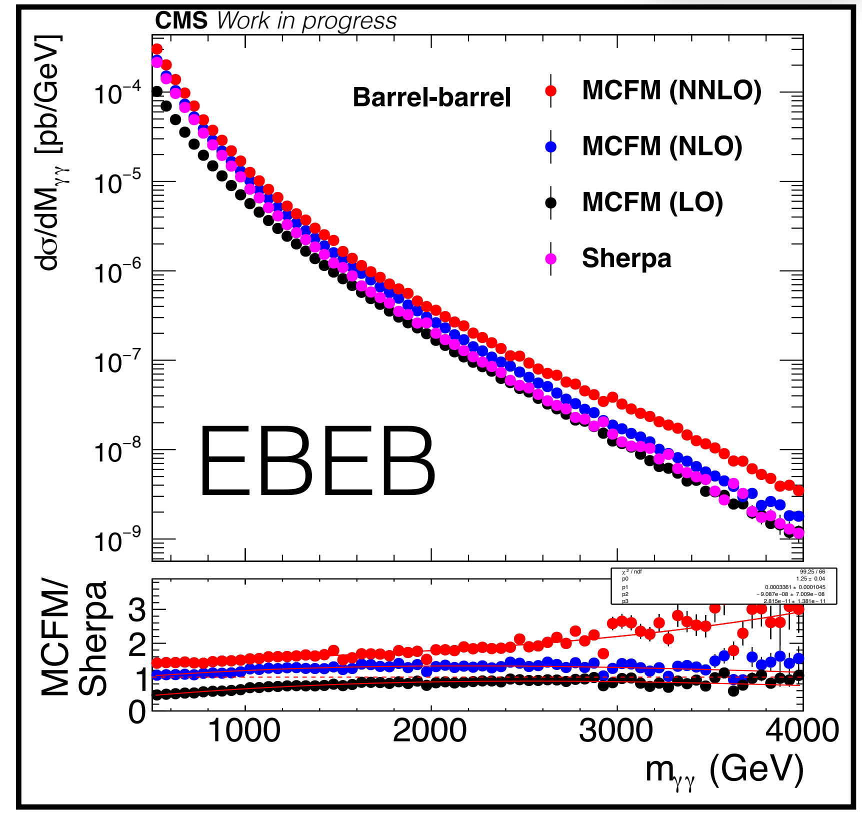


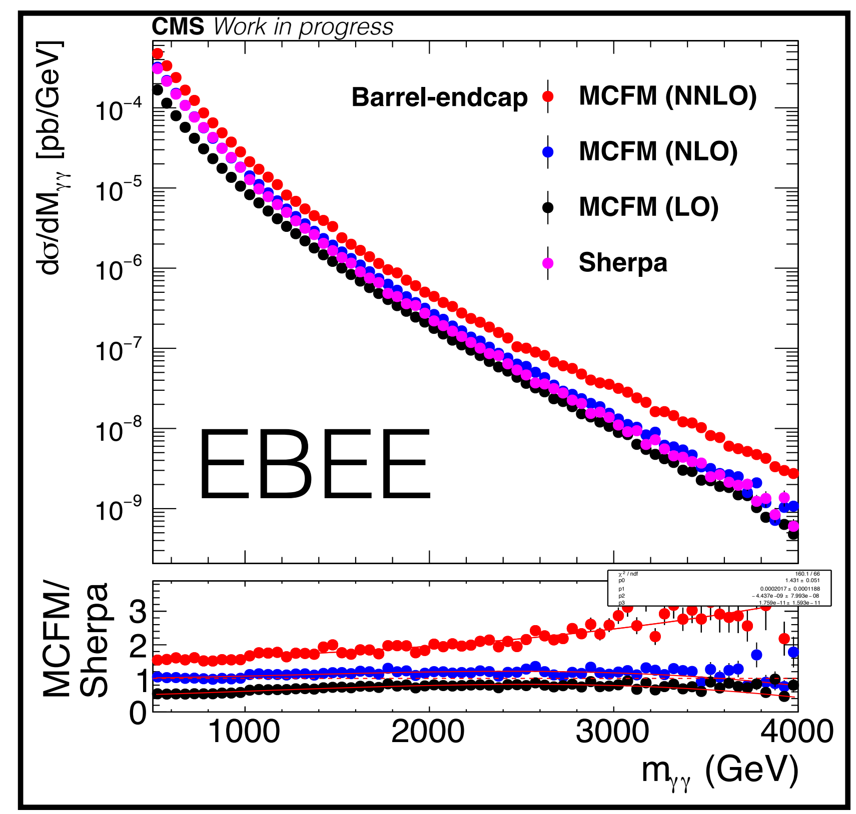
Born + 0, 1, 2, or 3 jets

- NNLO contribution to SM processes calculated using MCFM v.8.0

- ***K* factor = MCFM / GEN-level Sherpa**
  - Function of $m_{\gamma\gamma}$
  - Used to reweight fully simulated prompt diphoton events (from Sherpa) for final prediction

- *K* factor: 1.4-2.2(2.5) in EBEB(EBEE) for $m_{\gamma\gamma}$ < 3 TeV

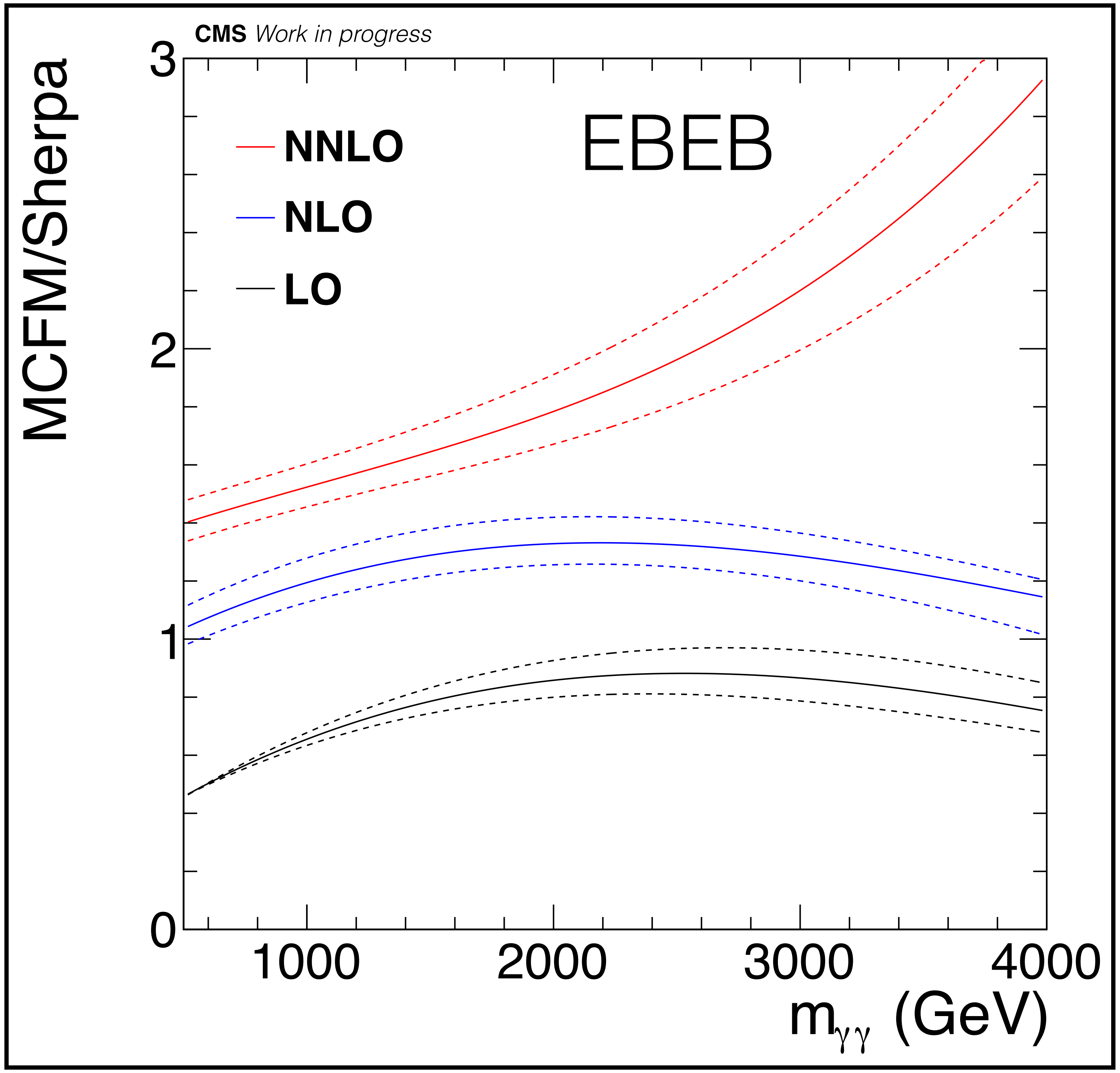




Scale variations on K factor

- Scale variations
  - $\mu_f = \mu_r = 2\ m_{\gamma\gamma}$
  - $\mu_f = \mu_r = m_{\gamma\gamma}$
  - $\mu_f = \mu_r = 1/2\ m_{\gamma\gamma}$

- Similar level of variation observed in EBEE



Fake background

- Reducible background from SM γ***+j*** or ***j+j*** events where one or two jets with large EM activity fake a photon signature

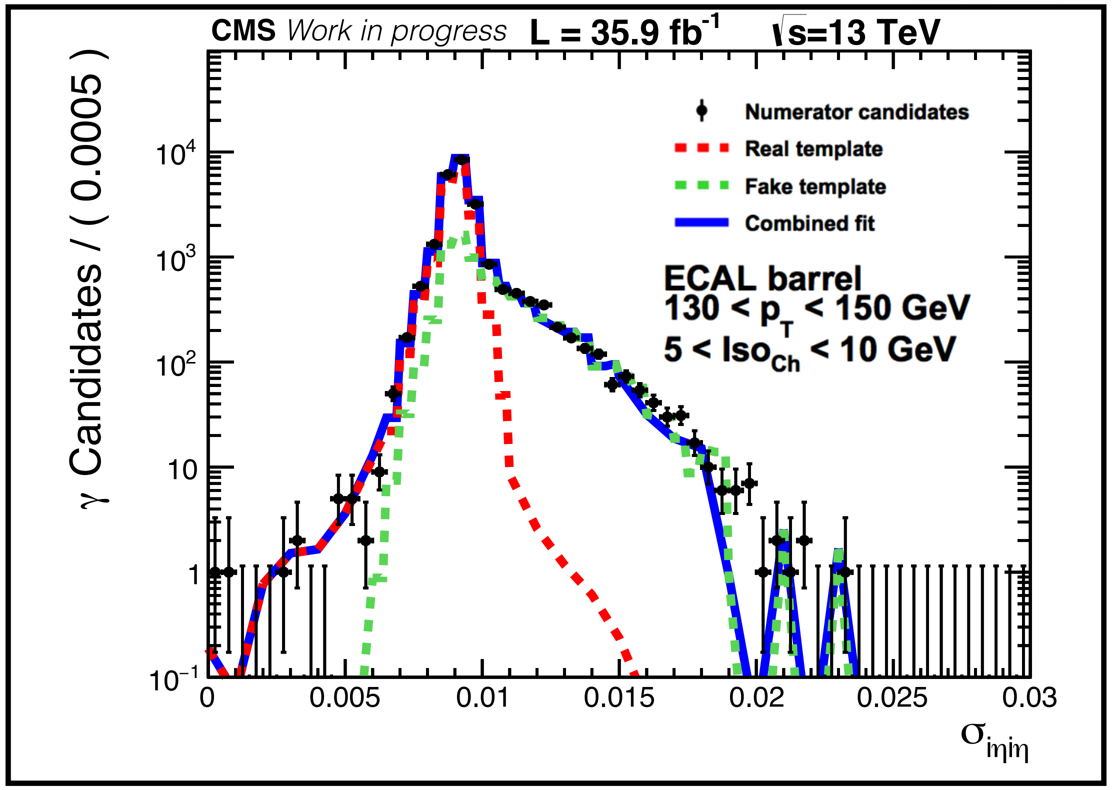
- Fake rate function ***f*(*p*$_T$)** is measured in a jet-enriched data reference sample as a function of photon p_T in EB and EE

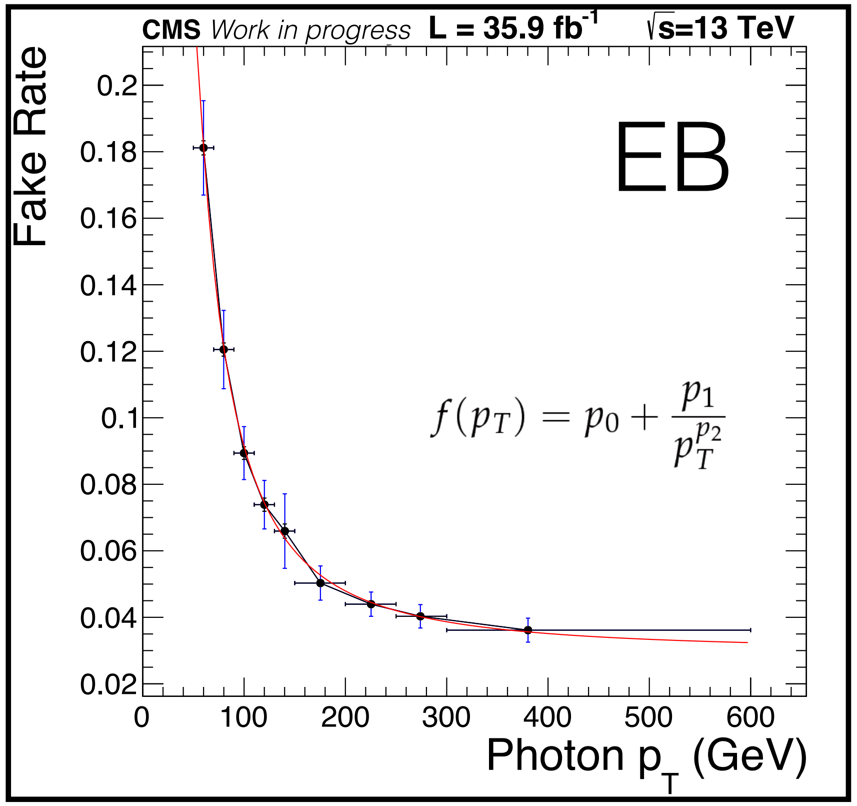
- **fake rate: *f*(*p*$_T$) = numerator / denominator**
  - **numerator:** number of jets passing photon ID
  - **denominator:** number of jets passing a looser photon-like ID

- Contamination from real photons is removed from **numerator** through a template fit

- **Denominator** objects in analysis data sample are reweighed by ***f*(*p*$_T$)** to give a fake prediction

- Fake contribution is less than 4(14)% in EBEB(EBEE) for $m_{\gamma\gamma}$ > 1 TeV

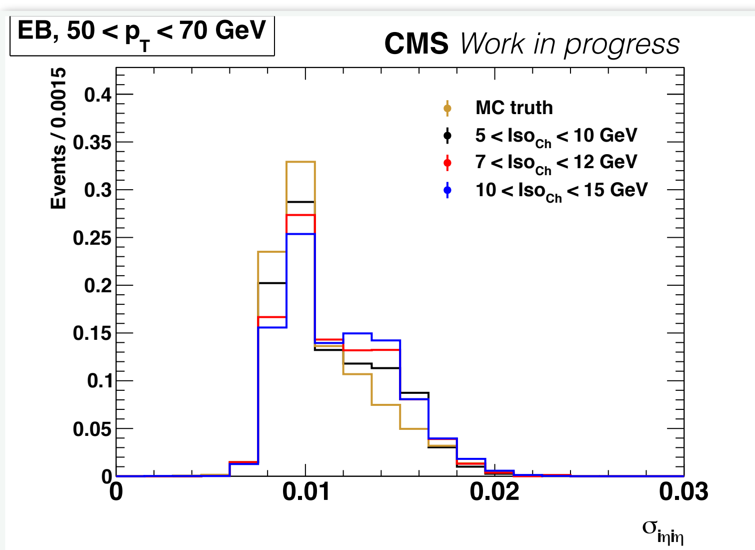


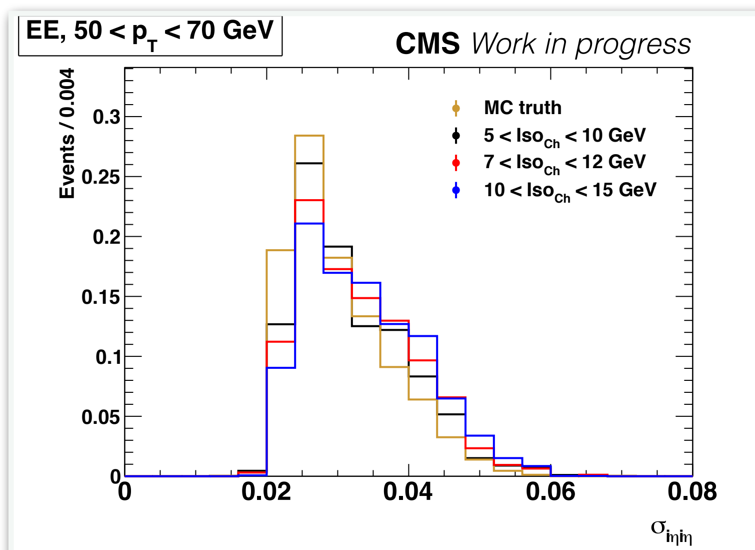


Closure test to photon fake rate method

MC based closure test is performed by comparing

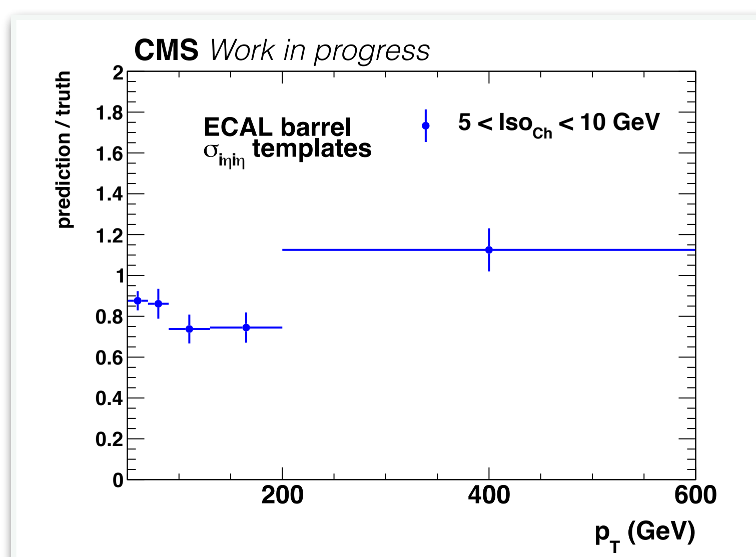
1. Fake rate method applied to jet-enriched MC sample (treat MC as data) to determine a **fake prediction** on MC sample
2. **MC truth** calculation of actual fake photons in MC sample

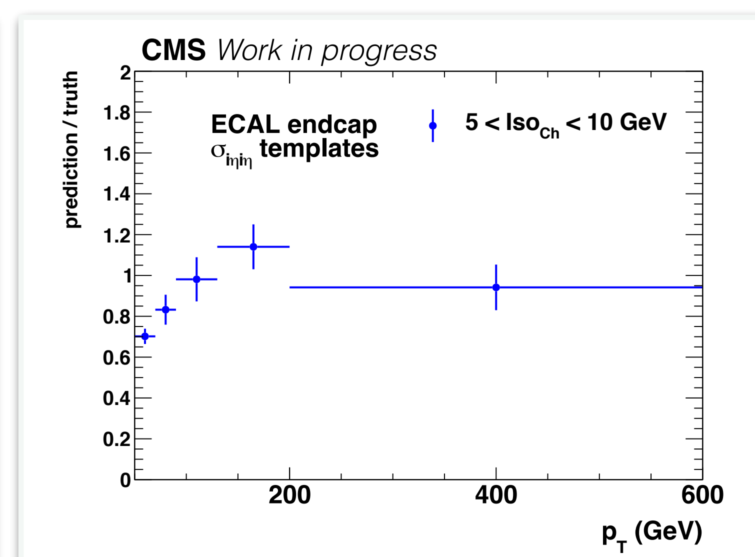




- MC fake templates to MC truth
  - choosing templates from a sideband has some bias

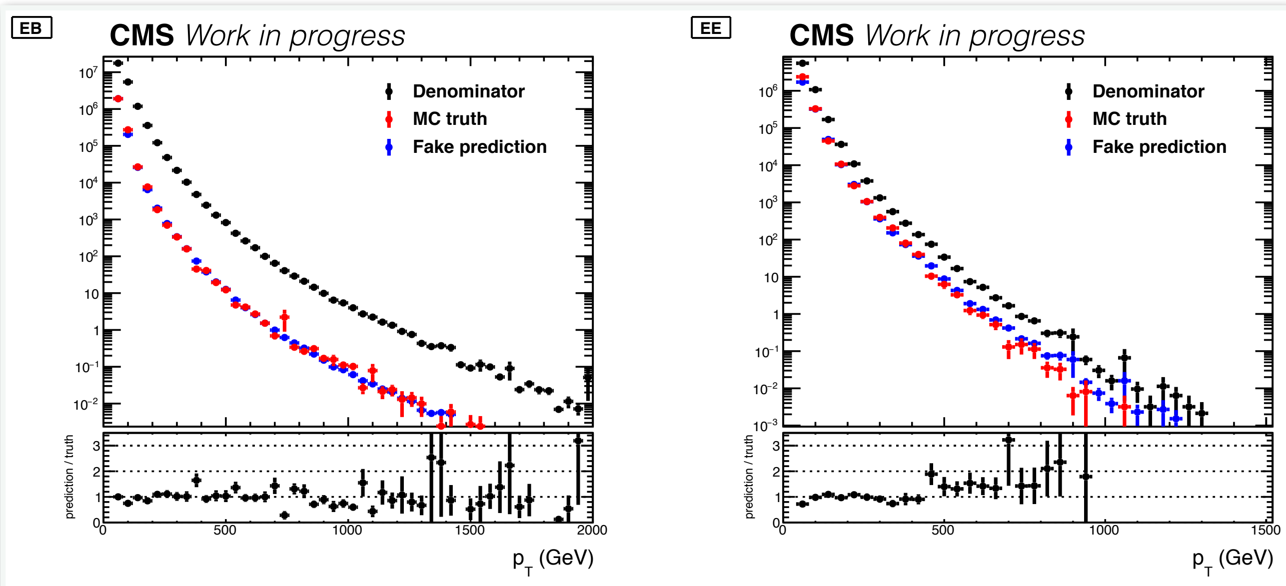
- Closure test predicted fake rate compared to truth fake rate
  - Similar agreement for inverted relationship between template and sideband variables

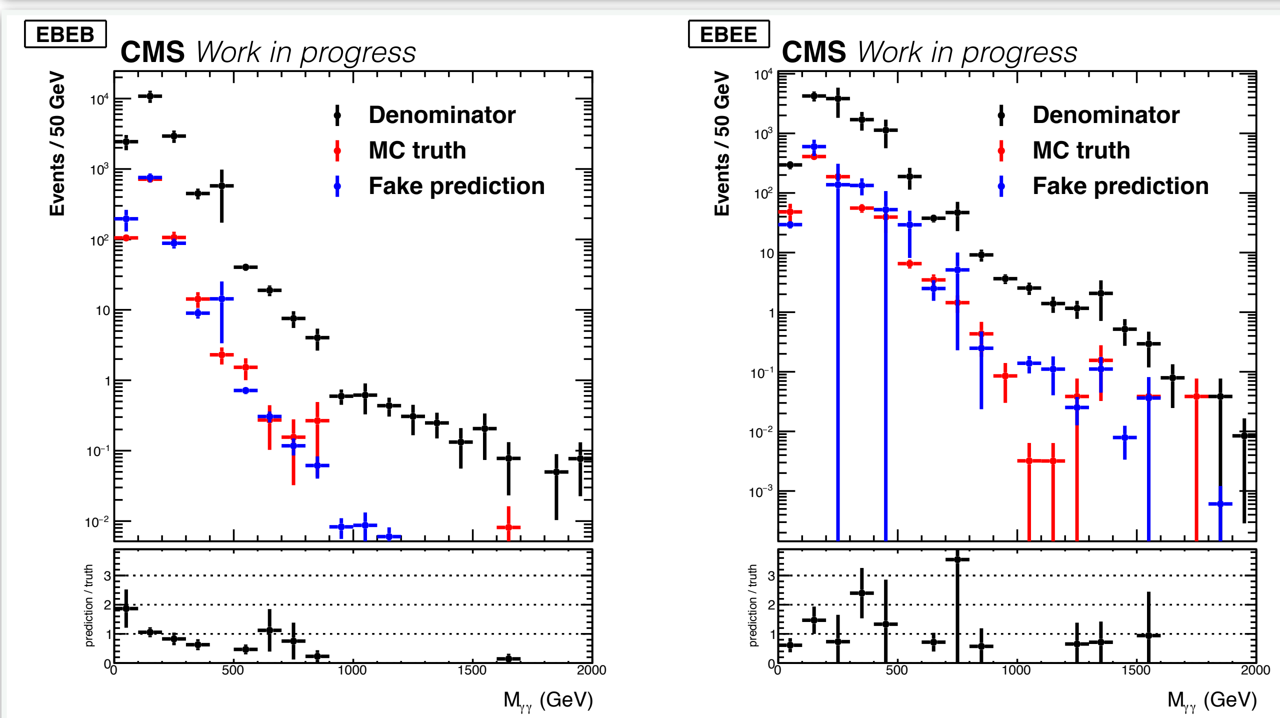




Closure test of kinematics

- The **re-weighted denominator objects** reflect the kinematics of known fake objects from **MC truth**



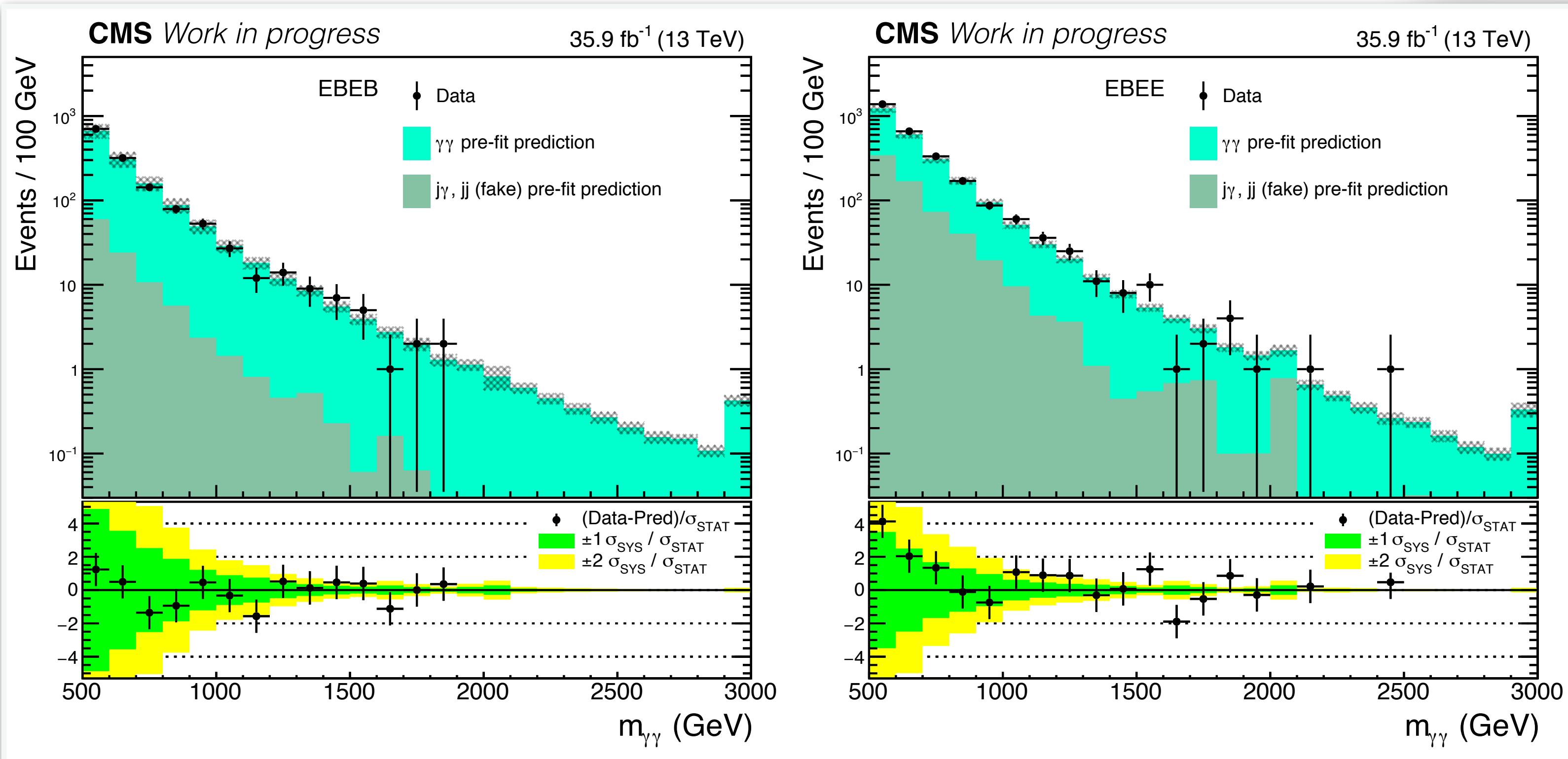


- Photon p_T in EB and EE

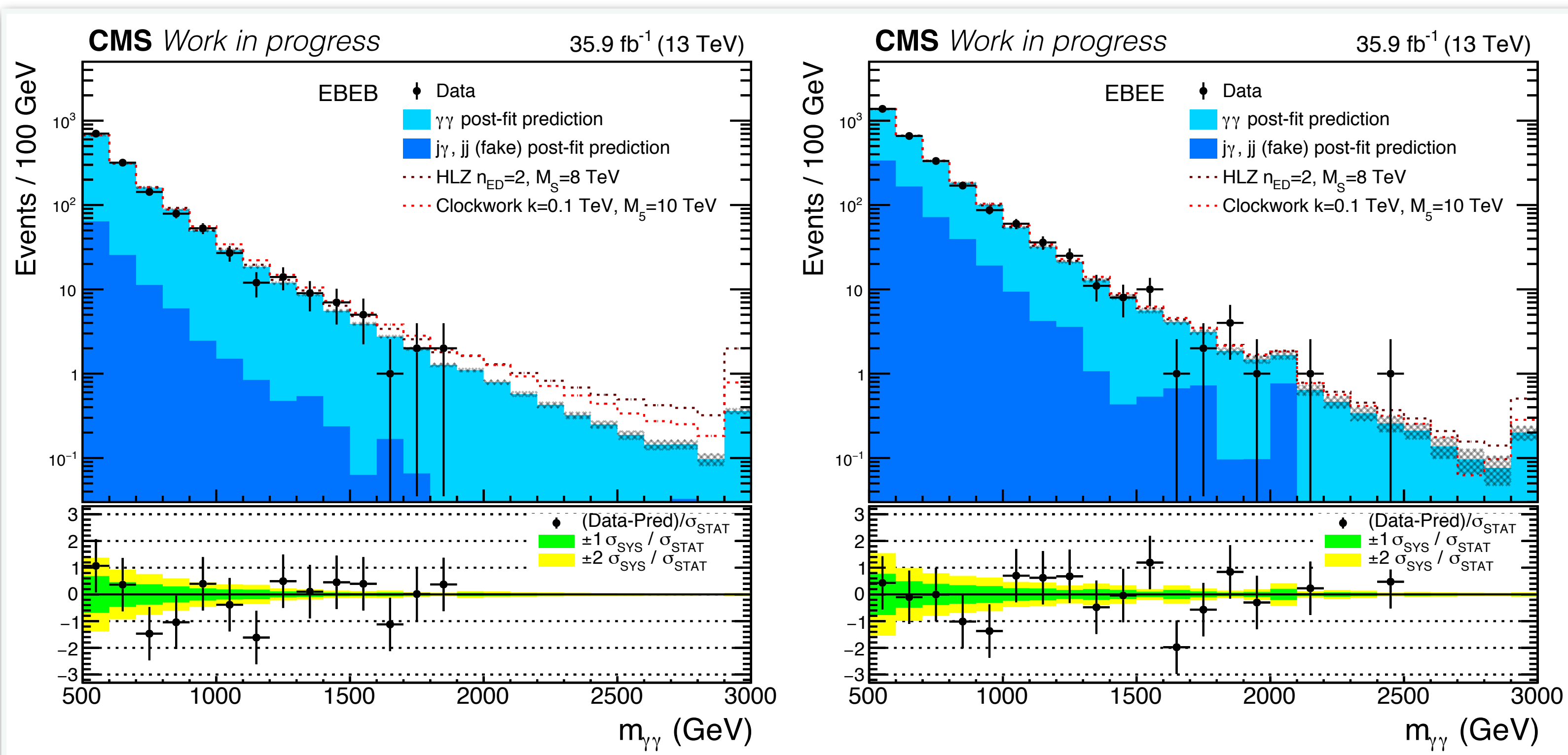
- Diphoton $m_{\gamma\gamma}$ γ+j objects in EBEB and EBEE

Results

- **Pre-fit** prediction of $m_{\gamma\gamma}$ spectra in good agreement with data
  - EBEB and EBEE

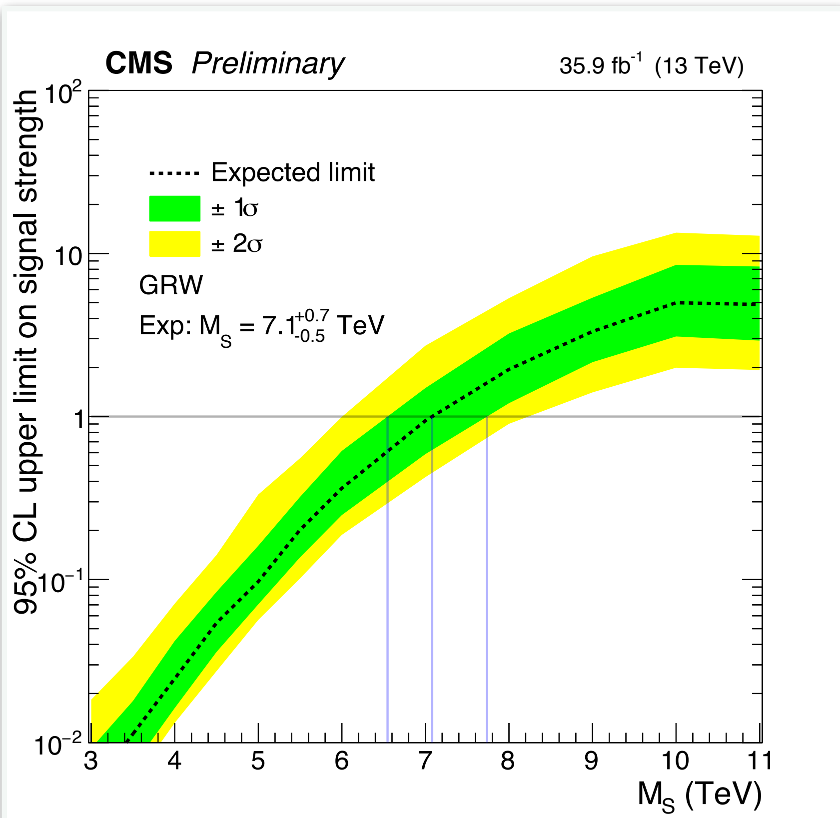


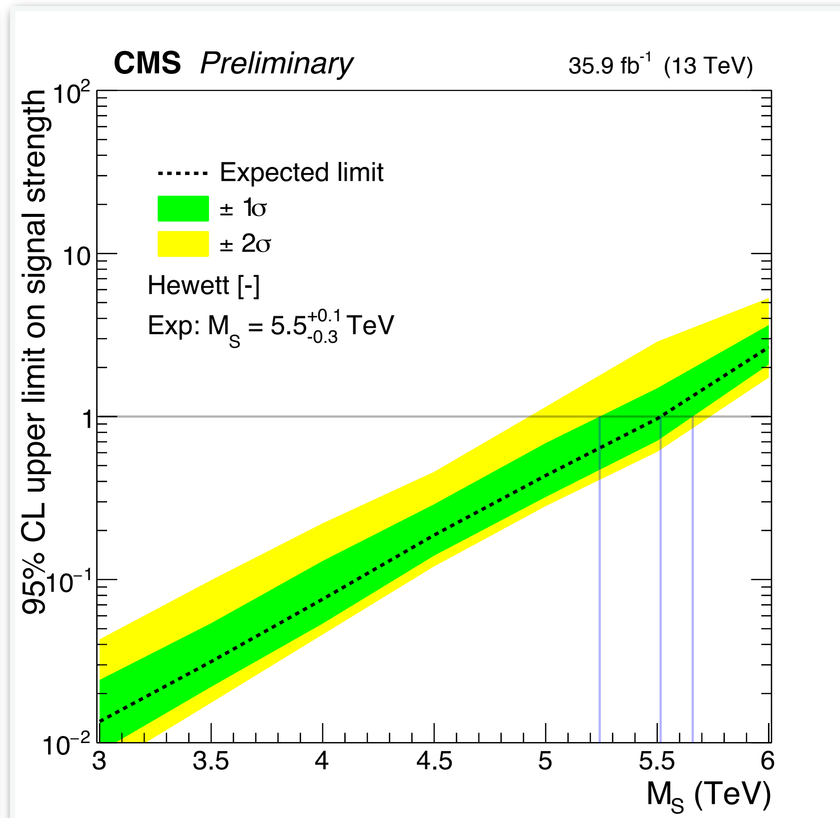
- **Post-fit** in $m_{\gamma\gamma}$ gives final prediction
  - Assign nuisance parameters to each source of systematic uncertainty
  - Perform a MLE fit
  - Allow real background normalization to float

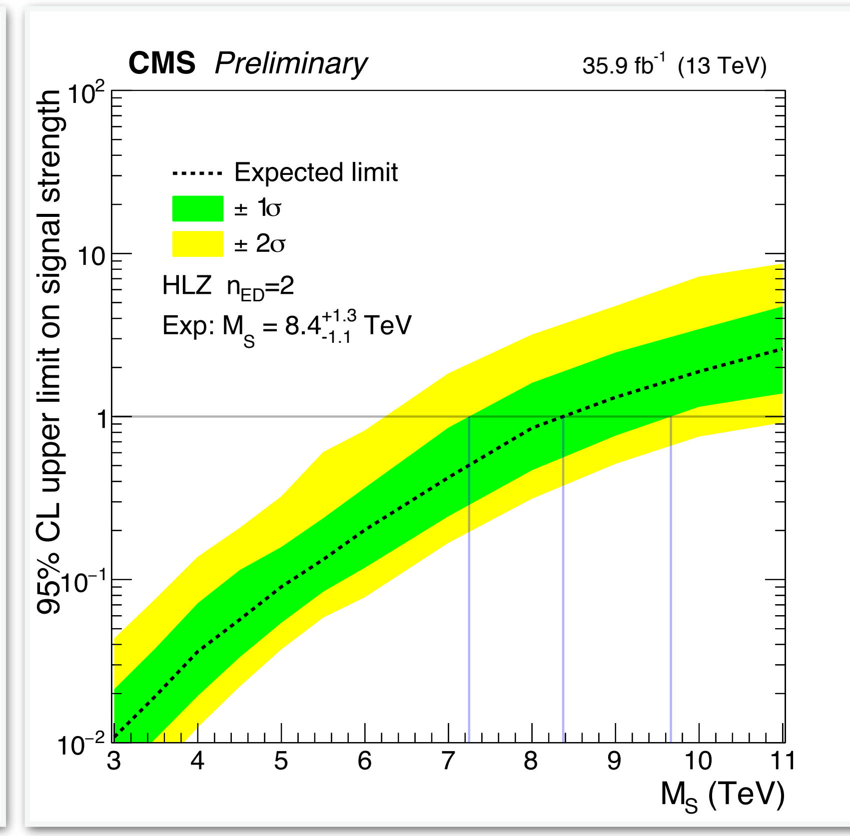


Limits on ADD model

- Upper limits on the ADD signal strength are translated into lower limits of M_S







- Expected limits on M_S range from 5.6-8.3 TeV

Signal:	GRW	Hewett		HLZ					
		negative	positive	$n_{ED}=2$	$n_{ED}=3$	$n_{ED}=4$	$n_{ED}=5$	$n_{ED}=6$	$n_{ED}=7$
Expected:	$7.0^{+0.5}_{-0.5}$	$5.6^{+0.1}_{-0.2}$	$6.3^{+0.4}_{-0.4}$	$8.3^{+1.2}_{-1.0}$	$8.3^{+0.6}_{-0.6}$	$7.0^{+0.5}_{-0.5}$	$6.3^{+0.5}_{-0.5}$	$5.9^{+0.4}_{-0.4}$	$5.6^{+0.4}_{-0.4}$

CMS *Work in progress* **CMS** *Work in progress*
CMS *Work in progress*

Conclusion and outlook

- A non-resonant, high-mass diphoton search using 35.9 fb^{-1}, corresponding to the full 2016 dataset, was performed with CMS
  - NNLO prediction of the prompt SM diphoton background
  - Data-driven estimate of the fake background
- Observed data is consistent with SM background
- Limits set on the ADD model of large extra dimensions
  - Expected limits on M_S = 5.6-8.3 TeV depending on the cutoff convention
    - Improves current best limits by ATLAS (*arXiv:1707.04147*)
- First CMS non-resonant, high-mass diphoton search done using LHC Run 2 data
  - Previous CMS result used 2011 data with $\sqrt{s}$ = 7 TeV pp collisions (*arXiv:1112.0688*)
- Review of paper draft under internal review
  - Journal submission expected to follow

Backup

Large extra dimensions

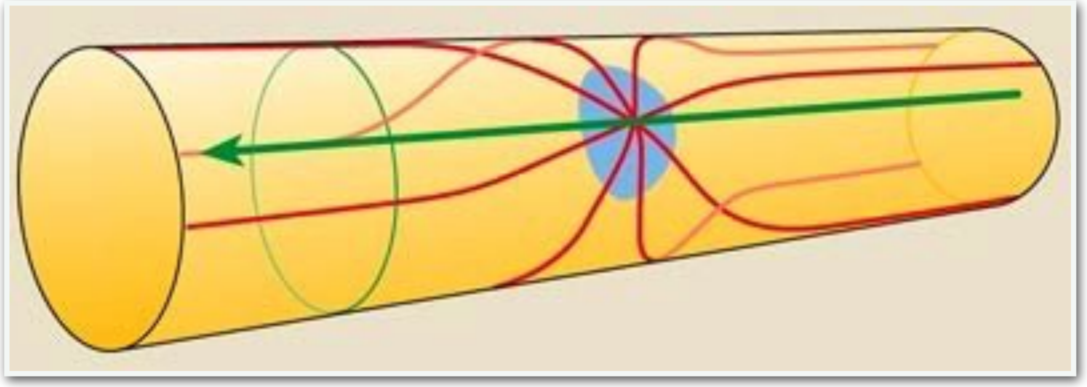
- Gravitational potential in (n+3+1) dimensions

$$V(r) \sim \frac{m_1 m_2}{M_D^{n+2}} \frac{1}{r^{n+1}}, \ (r \ll R)$$

$$V(r) \sim \frac{m_1 m_2}{M_D^{n+2} R^n} \frac{1}{r}, \ (r \gg R)$$

- For $r \gg R$, we get our usual $1/r$ potential and an observed Planck scale

$$M_{pl}^2 \sim M_D^{2+n} R^n$$



- If $M_D \sim 1$ TeV, we find

$$R \sim 10^{\frac{30}{n} - 19} \ \mathrm{m}$$

- We get large extra dimensions

Luminosity

- Instantaneous luminosity

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y}$$

- Integrated luminosity

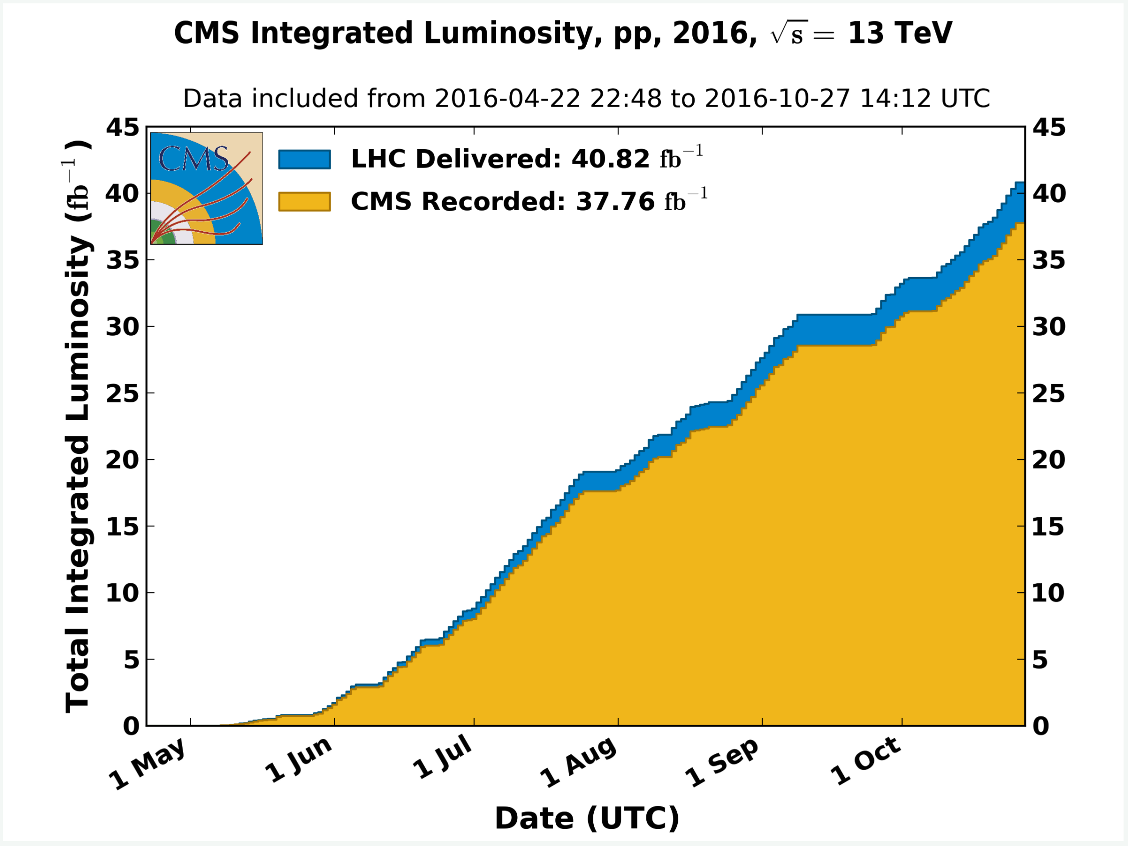
$$L = \int \mathcal{L}(t)dt$$

- Number of events

$$N = \sigma L$$

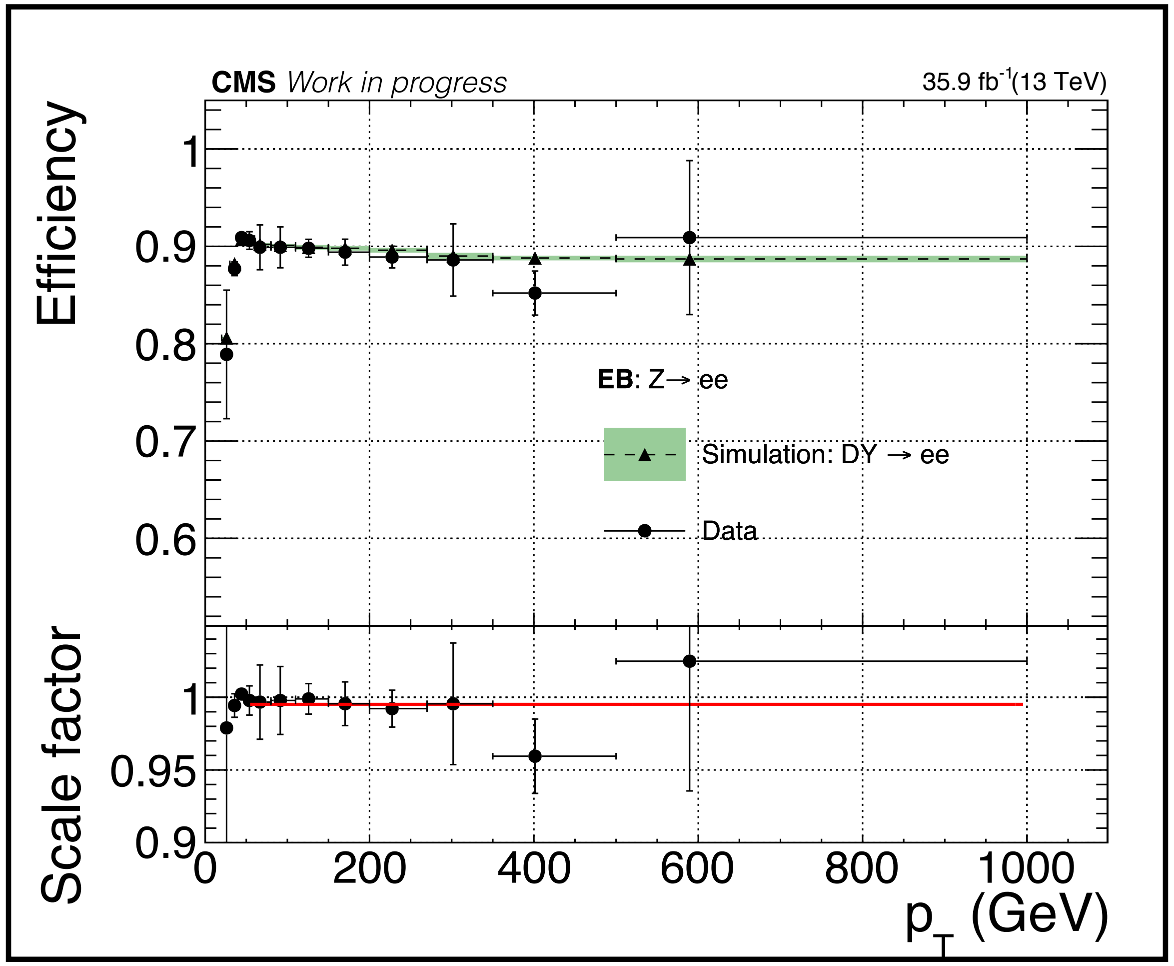
1 $\text{fb}^{-1} = 10^{39}\ \text{cm}^{-2}$

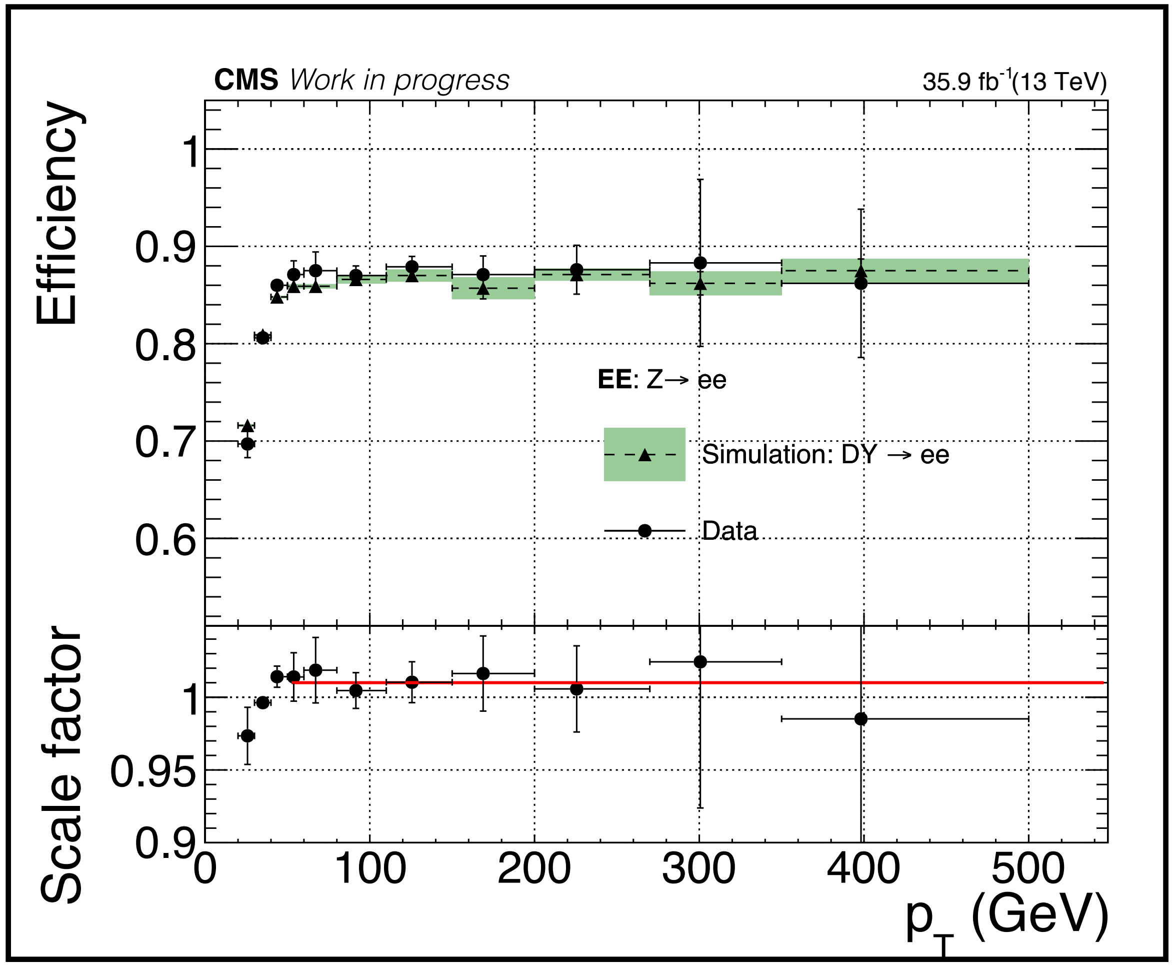
- f: collision frequency
- n_i: number of particles in bunch i
- $\sigma_{x,y}$: rms beam size in x and y
- σ: cross section



Photon selection efficiency

- Measured using the Tag and Probe technique





ADD signal model

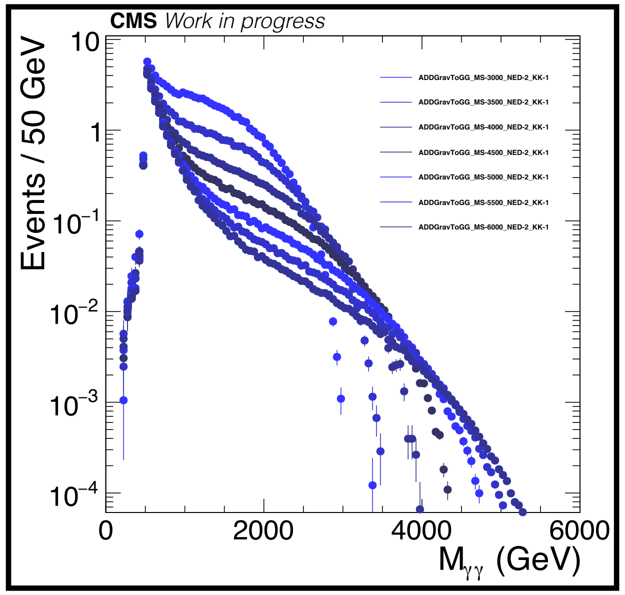
- ADD model is parameterized by number of extra dimensions $\boldsymbol{n_{ED}}$ and UV string cutoff scale $\boldsymbol{M_S}$
- Interference effects:
  - Signal generated with LO SM background using Sherpa v.2.1.1 (**signal+background** samples)
    - Samples generated with $m_{\gamma\gamma} \leq M_S$
  - Separate LO background samples are generated to extract the signal (background only samples)
- Cutoff conventions $\mathcal{F}$
  - Translate between different conventions with same $\boldsymbol{\eta_G}$
  - Consider **HLZ (n_{ED} =2), GRW, and Hewett-**
- Parameter space:
  - M_S = 3-11 TeV
  - n_{ED} = 2,3,…,7

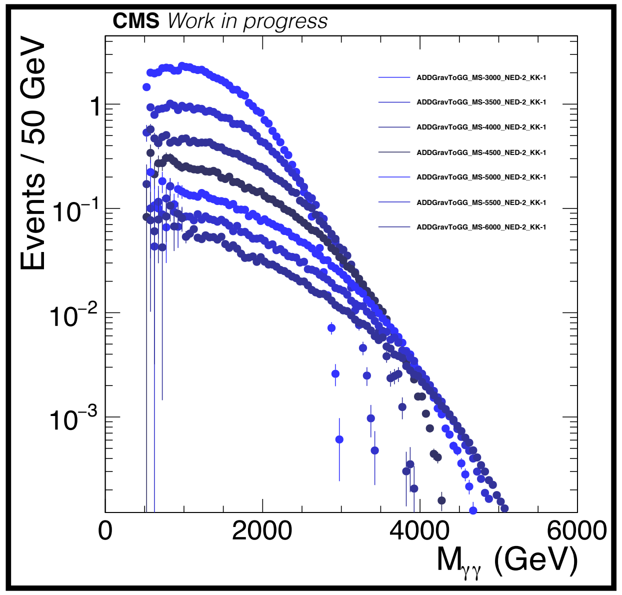
- Total cross section:

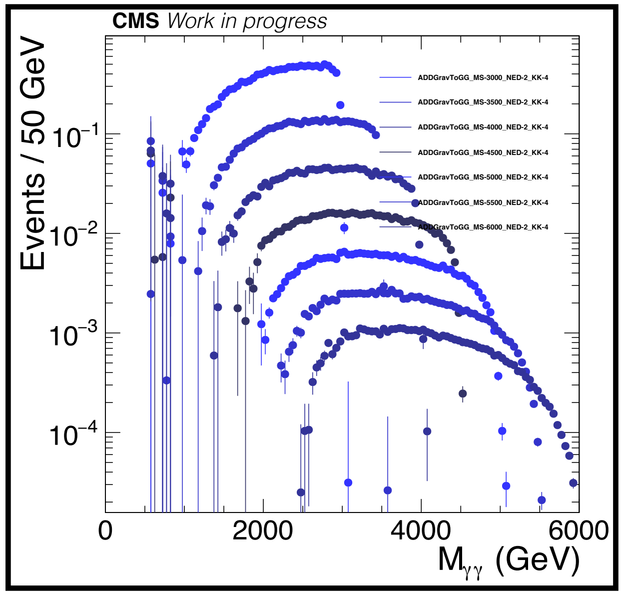
$$\sigma_{\text{total}} = \sigma_{\text{SM}} + \frac{\mathcal{F}}{M_S^4}\sigma_{\text{int}} + \frac{\mathcal{F}^2}{M_S^8}\sigma_{\text{ADD}}$$

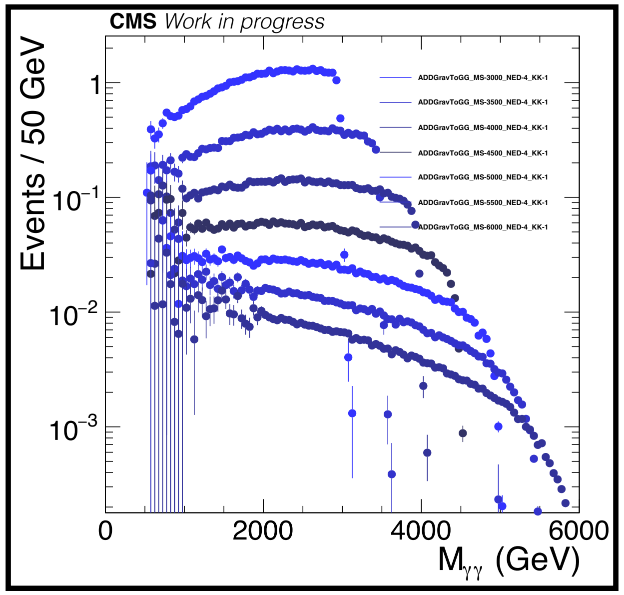
$$\mathcal{F} = \begin{cases} 1 \quad \text{(GRW)}, & \\ \log\left(\frac{M_S^2}{\hat{s}}\right), \text{ if } n_{ED} = 2 & \\ \frac{2}{n_{ED}-2}, \text{ if } n_{ED} > 2 & \text{(HLZ)}, \\ \pm\frac{2}{\pi} \quad \text{(Hewett)}, & \end{cases}$$

$$\eta_G = \mathcal{F} / M_S^4$$







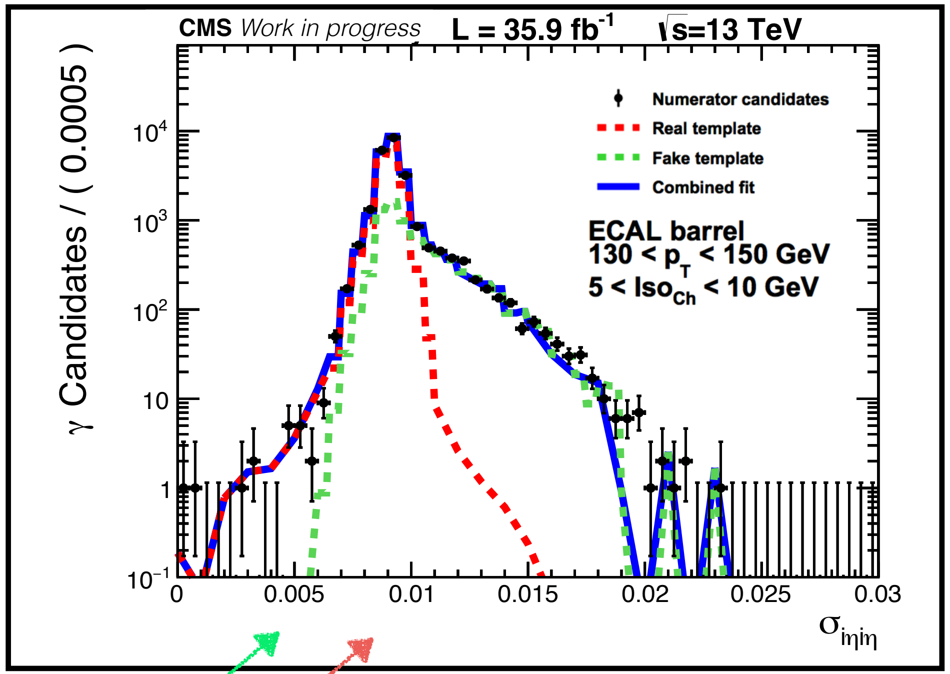


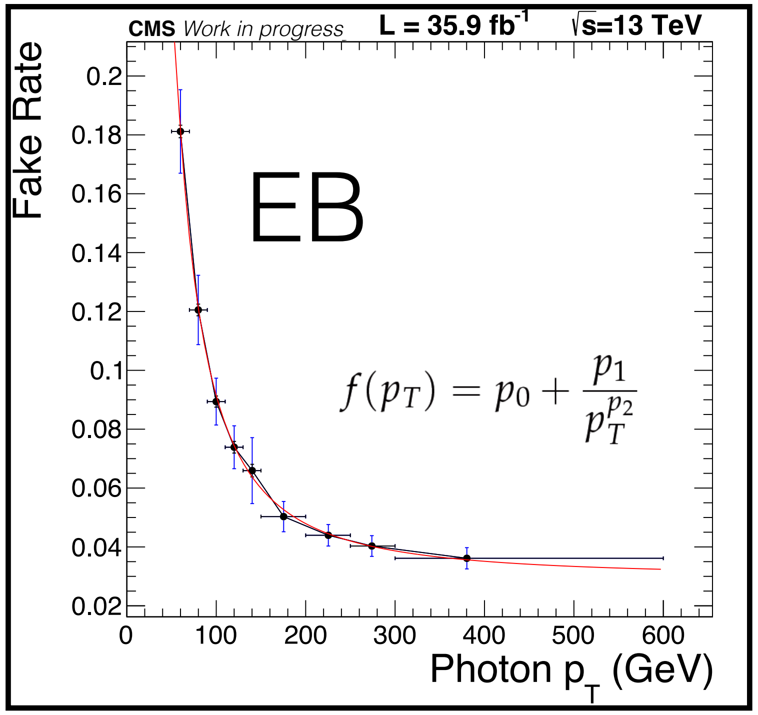
signal+background | **signal** | **signal** | **signal**

HLZ (n_{ED} = 2) | **Hewett-** | **GRW**

Fake background

- Reducible background from SM γ**+j** or **j+j** events where one or two jets with large EM activity fake a photon signature, typically a π^0 or η
- Fake rate function f is measured in a jet-enriched data reference sample as a function of photon p_T in EB and EE, then is applied to analysis data sample to give a fake prediction
- **fake rate = $f(p_T)$ = numerator / denominator**
  - **numerator** = number of jets passing photon ID
  - **denominator** = number of jets passing a looser photon-like ID
- Real, signal photons must be removed from **numerator**
  - Real and fake photon templates are fit to data in the template variable $\sigma_{i\eta i\eta}$, a shower-shape variable sensitive to real and fake photons
    - **Fake templates** - from reference data in sideband of the isolation variable Iso_{Ch}
    - **Real templates** - from MC
    - **Template fit** - performed using MLE fit to reference data
- **Denominator** objects in analysis data are reweighted by the fake rate function f
  - Fake contribution is less than 4(14)% in EBEB(EBEE) for $m_{\gamma\gamma}$ > 1 TeV





Systematics

Prompt diphoton background

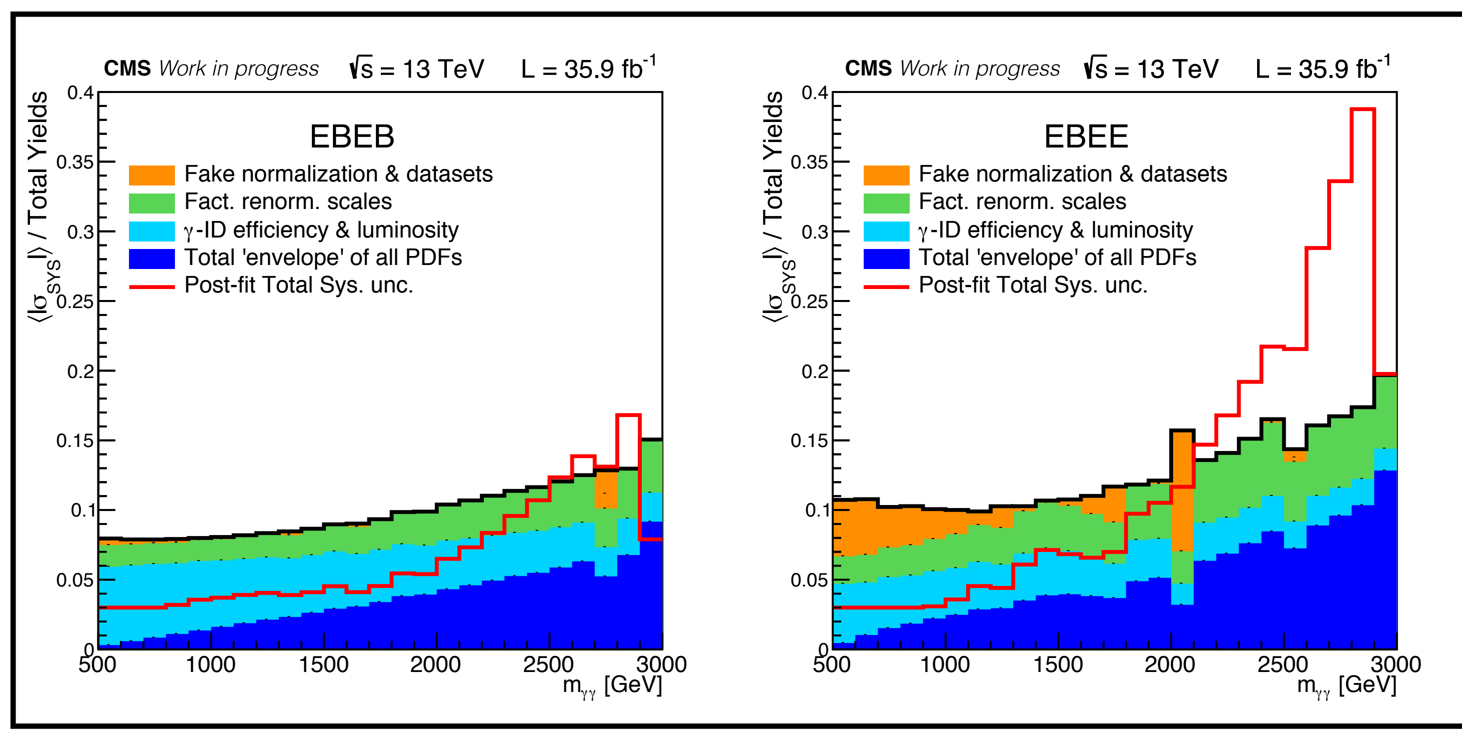
- The normalization is allowed to float arbitrarily, which absorbs uncertainties in
  - higher order terms not included in our NNLO *K* factor
  - luminosity (2.5%) and photon ID efficiency (6%)
- Scale variations on μ_r and μ_f for the MCFM *K* factor calculation
- 26 NLO CT10 PDF eigenvector pair variations using DIPHOX at NLO
- Shape differences in NNLO MCFM and NLO MCFM and NLO DIPHOX

Fake background

- ~30% uncertainty on fake rate normalization
  - pileup, photon η, and non-closure
- Shape differences in fake rate from JetHT and DoubleMuon datasets

MC signal

- 2.5% for luminosity
- 6% for photon ID efficiency
- 26 NLO CT10 PDF eigenvector pair variations using DIPHOX at NLO on acceptance



EBEB **EBEE**