Search for Higgs decay into a photon and a massless dark photon (ggF) in Run 3

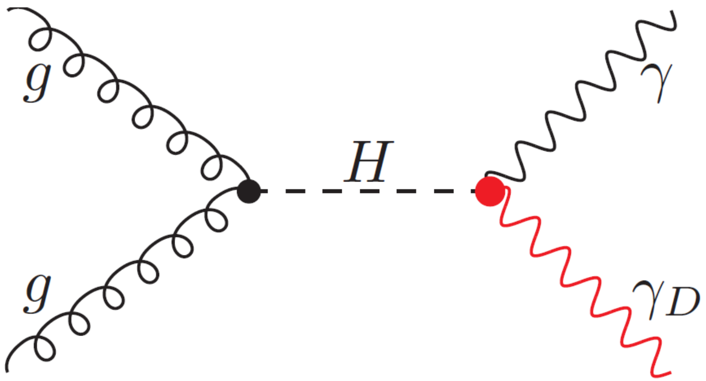
Longsheng Tan[1,2], Andrea Malara[1], Laurent Thomas[1], Pascal Vanlaer[1], Li Yuan[2], Si Hyun Jeon[3]

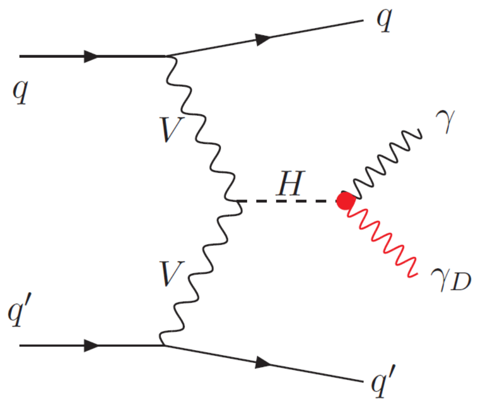
[1] Université libre de Bruxelles / IIHE

[2] Beihang University

[3] Boston University

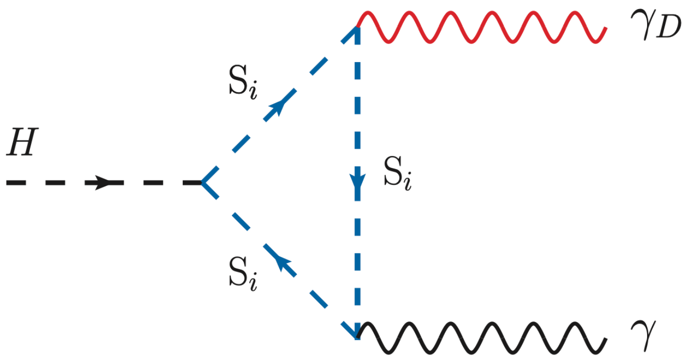
4 August 2026





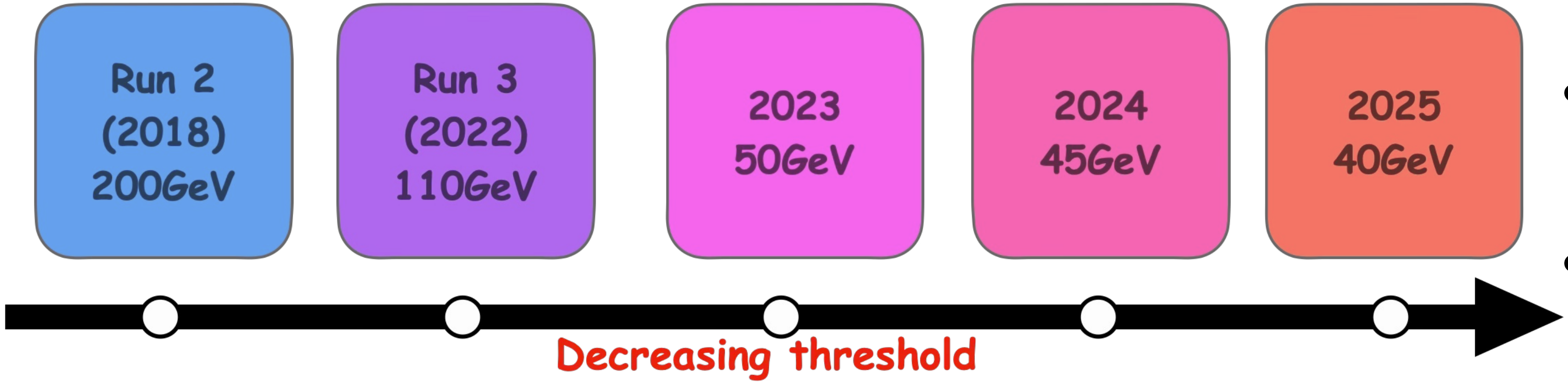
Motivation for $H \rightarrow \gamma\gamma_D$

- $H \rightarrow \gamma\gamma_D$, where the massless dark photon is invisible
- Experimental signature: one isolated photon recoiling against MET
- Exotic Higgs decays provide sensitivity to weakly coupled dark sectors
- Targeting the dominant ggF production provides access to a phase space complementary to previous CMS ZH and VBF searches
- Latest ATLAS results show improvement from ggF
- Previous talk at the CMS NPS Workshop 2026
- A first draft of the analysis note (AN-26-115) is now in place.

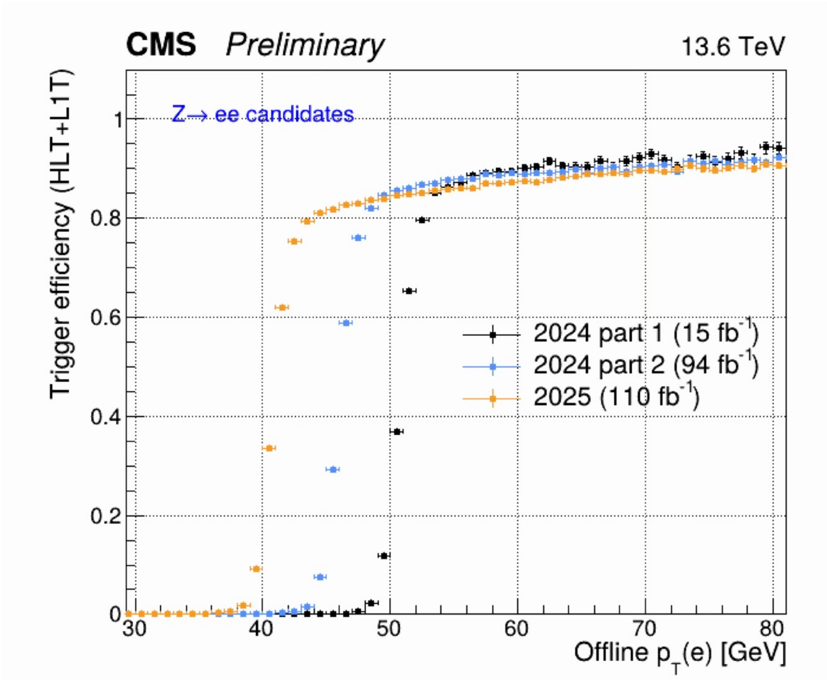


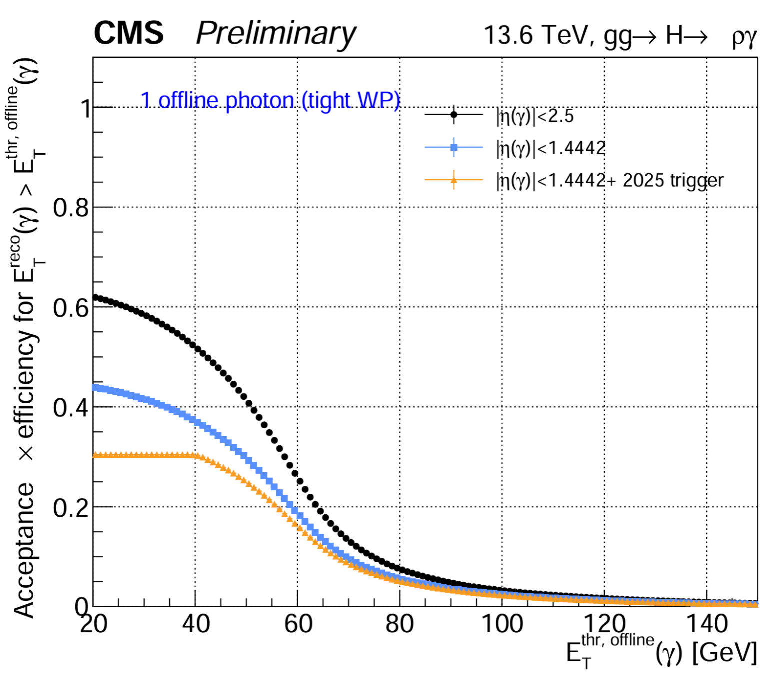
arXiv:2206.05297

A New ggF Opportunity in Run 3



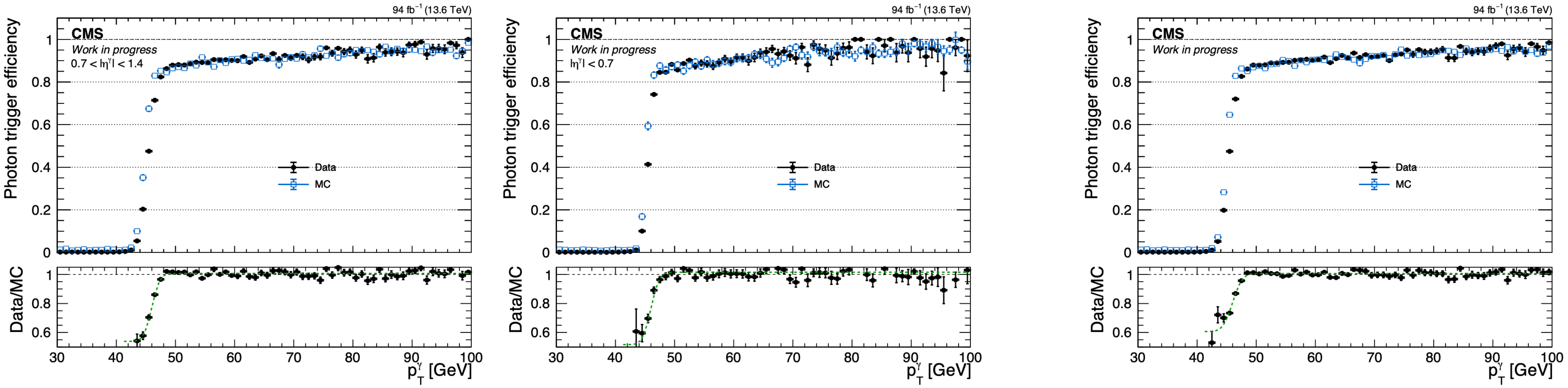
- CMS-DP-2025/086
- Low p_T single photon triggers deployed with high efficiency
- Signal acceptance falls rapidly at high photon E_T
- This analysis currently focuses on the 45 GeV trigger





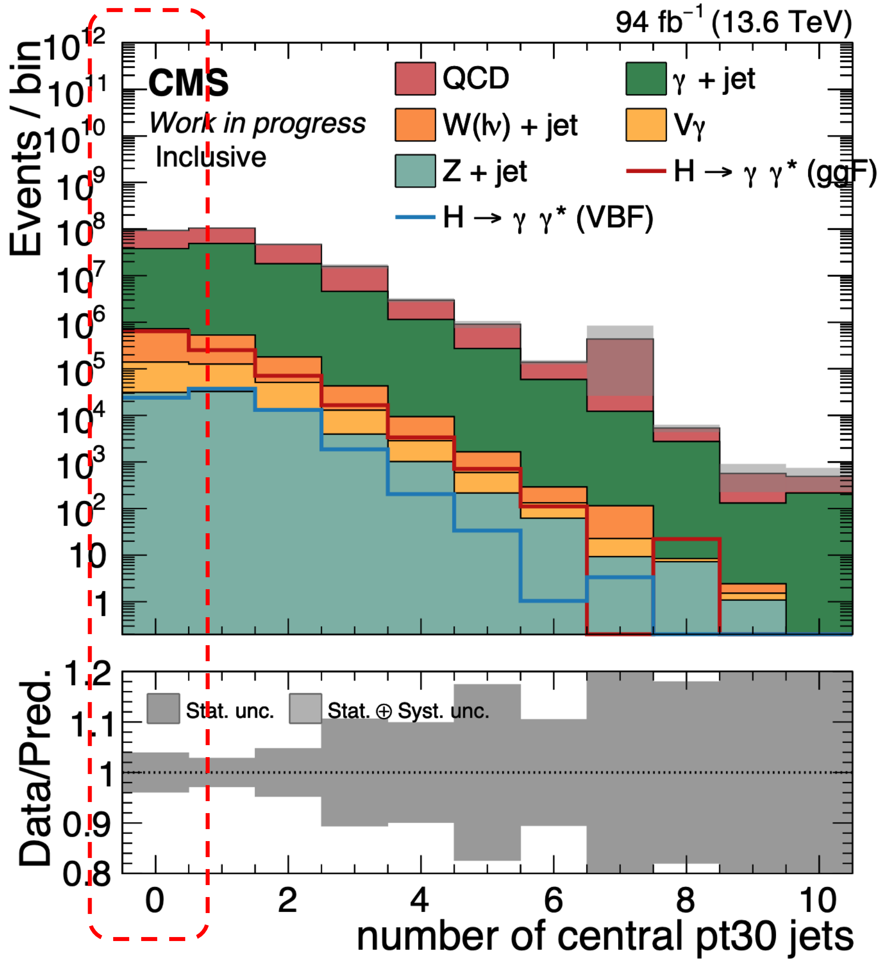
CMS-DP-2025/086

- **Dataset**
  - 2024 EGamma0/EGamma1
  - Eras E–I
  - 93.6 fb^{-1}
  - HLT_Photon45EB_TightID_TightIso
  - Barrel photon: $|\eta^{\gamma}|<1.4442$
- **Trigger performance**
  - $Z \rightarrow ee$ tag-and-probe
  - Endcap electron tag
    - $p_T \geq 32\mathrm{GeV}$, $1.566 < |\eta| < 2.5$
  - Barrel photon probe
    - $p_T \geq 20\mathrm{GeV}$, $|\eta| < 1.4442$
  - $50 \leq m(e, \gamma) \leq 130$ GeV



MET tails in inclusive jets category

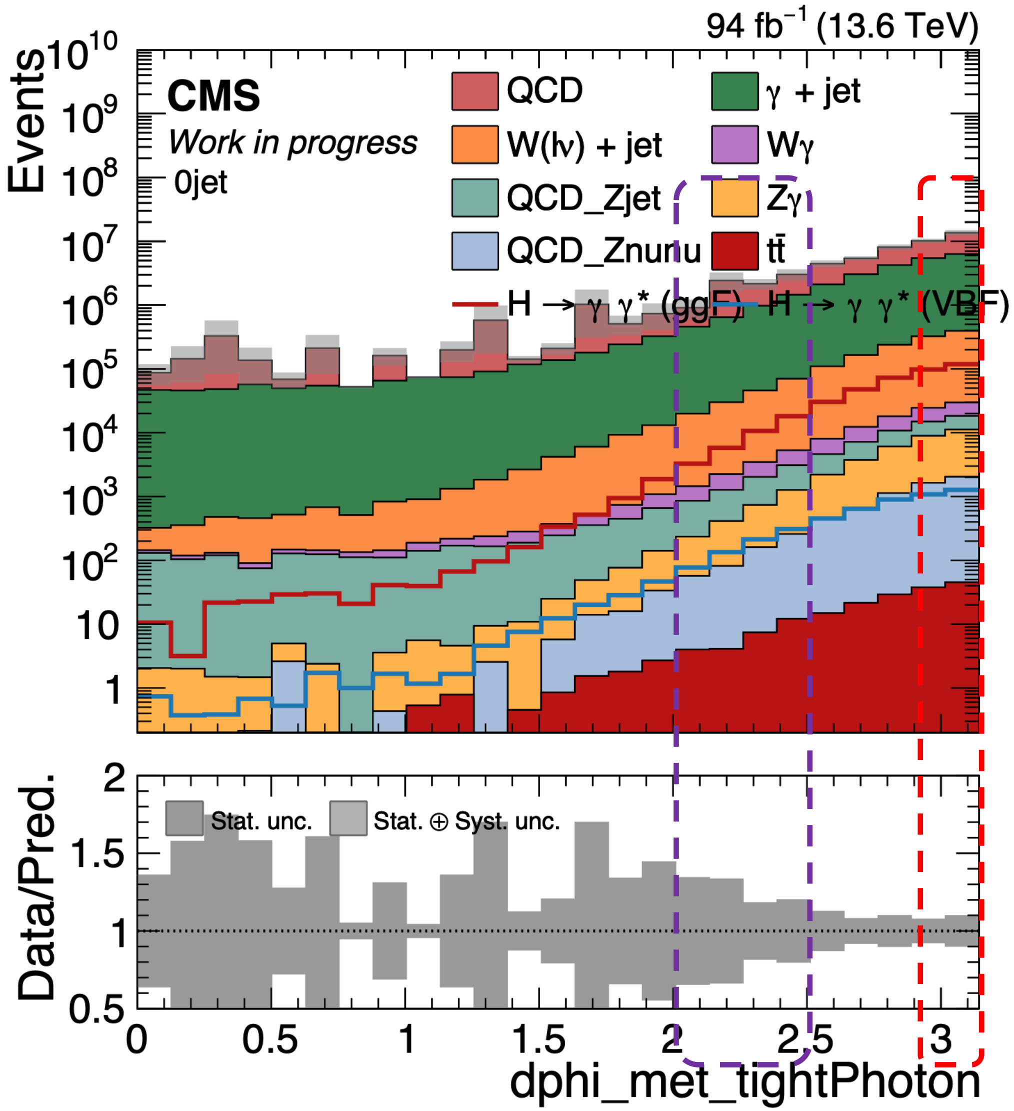
Event Selection and the 0-Jet Category



Process	N(0 jet)/N(inclusive)
ggF signal	65.0%
QCD multijets	35.8%
γ+jets	34.1%

Selection	Signal region	Low-$\Delta\phi$ validation region
Trigger	`HLT_Photon45EB_TightID_TightIso`	
Photon multiplicity	Exactly one tight photon in the ECAL barrel	
Photon kinematics and ID	$p_T^\gamma > 45\,\mathrm{GeV}$, $\lvert\eta^\gamma\rvert < 1.4442$, tight photon ID	
MET	Type-1 corrected PUPPI $40 < p_T^{miss} < 100\,\mathrm{GeV}$	
Lepton veto	No loose electron, muon, or hadronically decaying tau candidate	
b-jet veto	No b-tagged central AK4 jet with $p_T > 15\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.5$	
Fake-jet veto	No AK4 jet with $p_T > 50\,\mathrm{GeV}$ and $\lvert\eta\rvert < 4.7$ failing the “Tight” jet ID	
N_{jet} selection	No inclusive AK4 jet with $p_T > 30\,\mathrm{GeV}$ passing the jet selection	
$\Delta\phi(\gamma, \vec{p}_T^{miss})$	> 2.9	$2.0 < \Delta\phi(\gamma, \vec{p}_T^{miss}) < 2.5$

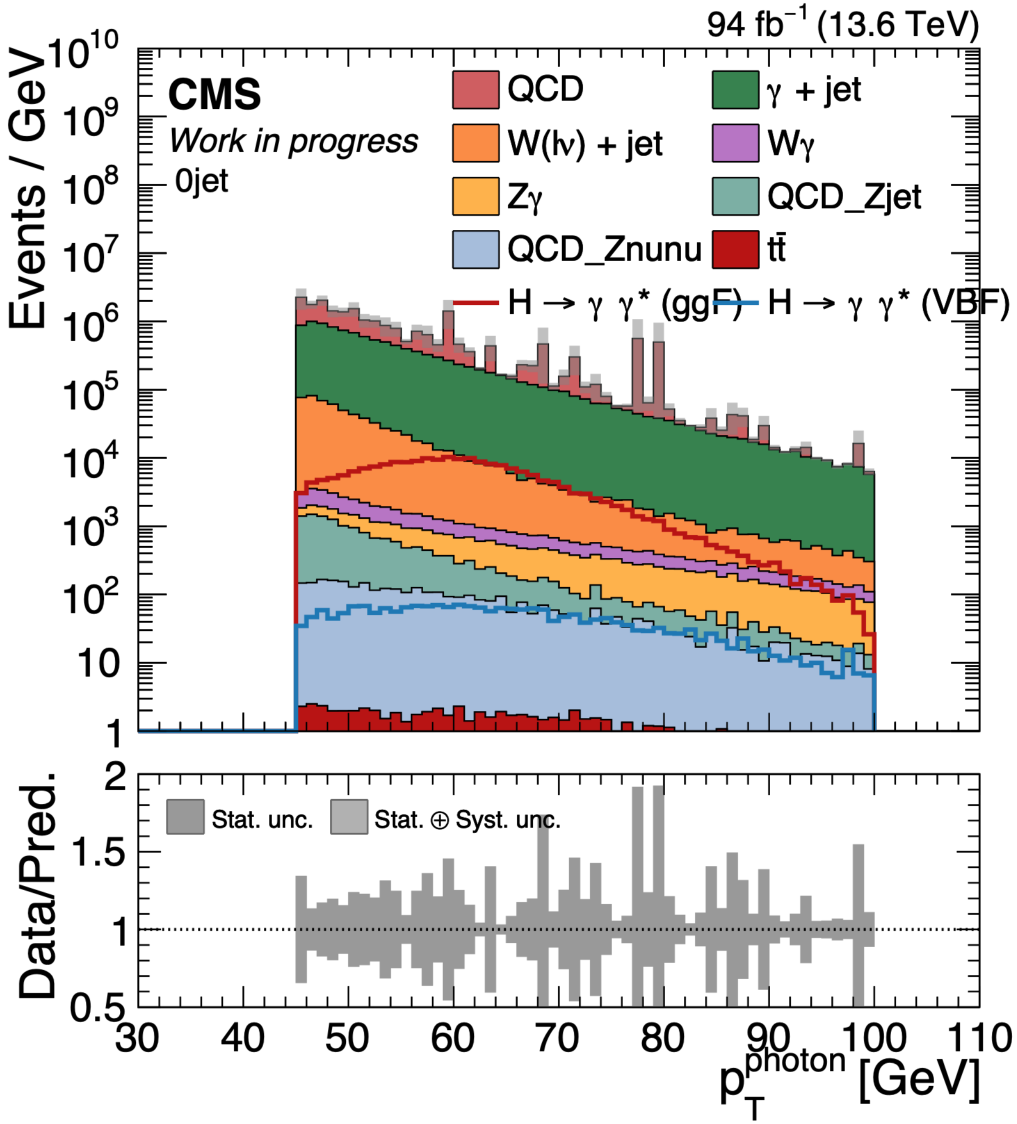
Signal and Validation Regions



<table>
<tr><th rowspan="2">Region</th><th colspan="2">N(region) / N(full selected)</th></tr>
<tr><th>ggF signal</th><th>γ+jets & QCD</th></tr>
<tr><td>VR $2.0 < \Delta\phi < 2.5$</td><td>8.7%</td><td>15.0%</td></tr>
<tr><td>SR $\Delta\phi > 2.9$</td><td>51.2%</td><td>42.2%</td></tr>
</table>

- $\Delta\phi(\gamma,\text{MET})$ is used to define the analysis regions
- Low $\Delta\phi$ region: background model validation region (VR)
- High $\Delta\phi$ region: signal region

Signal Extraction and Statistical Model



- **Observable**
  - Binned fit to the p_T^{γ} spectrum
  - Fit range: $49 < p_T^{\gamma} < 100$ GeV
  - Signal feature near $p_T^{\gamma} \approx m_H/2$
- **Model**
  - ggF: double-Gaussian function
  - VBF: fourth-order Bernstein polynomial
  - Background: three-order exponential
- **Inference**
  - Binned profile likelihood with Combine
  - Common branching fraction for ggF and VBF

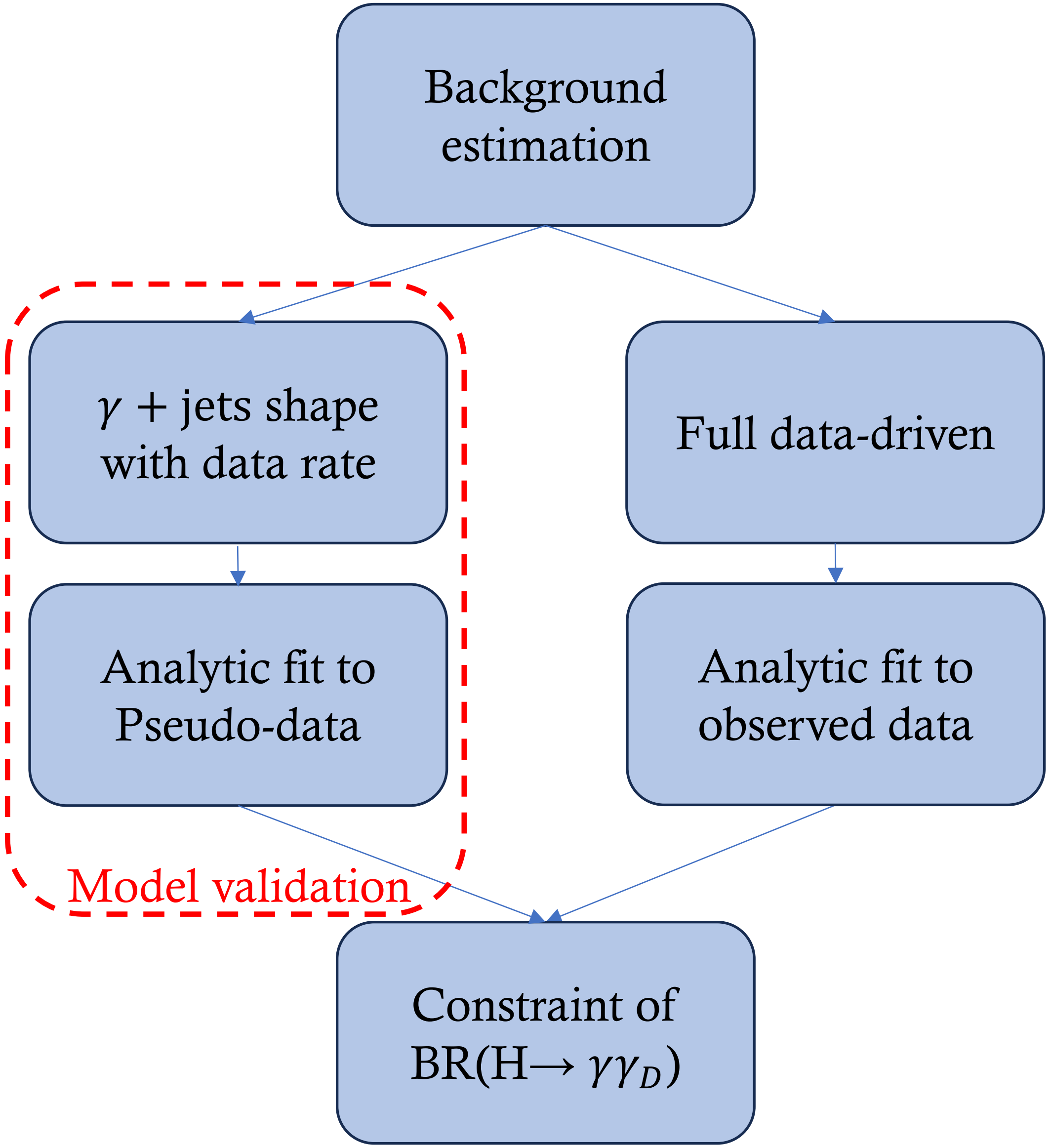
Background Estimation

Process	SR yield	SR fraction
QCD multijet	8,019,464	51.1%
γ + jets	7,256,605	46.3%
$W(\ell\nu)$ + jets	364,789	2.3%
$W\gamma$	14,285	0.1%
$Z\gamma$	13,200	0.1%
$Z(\ell\ell)$ + jets	7,664	$< 0.1\%$
$Z(\nu\nu)$ + jets	2,813	$< 0.1\%$
Top quark	66	$< 0.1\%$
Total simulated background	15,678,886	100%
ggF $H \rightarrow \gamma\gamma_D$	192,019	–
VBF $H \rightarrow \gamma\gamma_D$	2,018	–

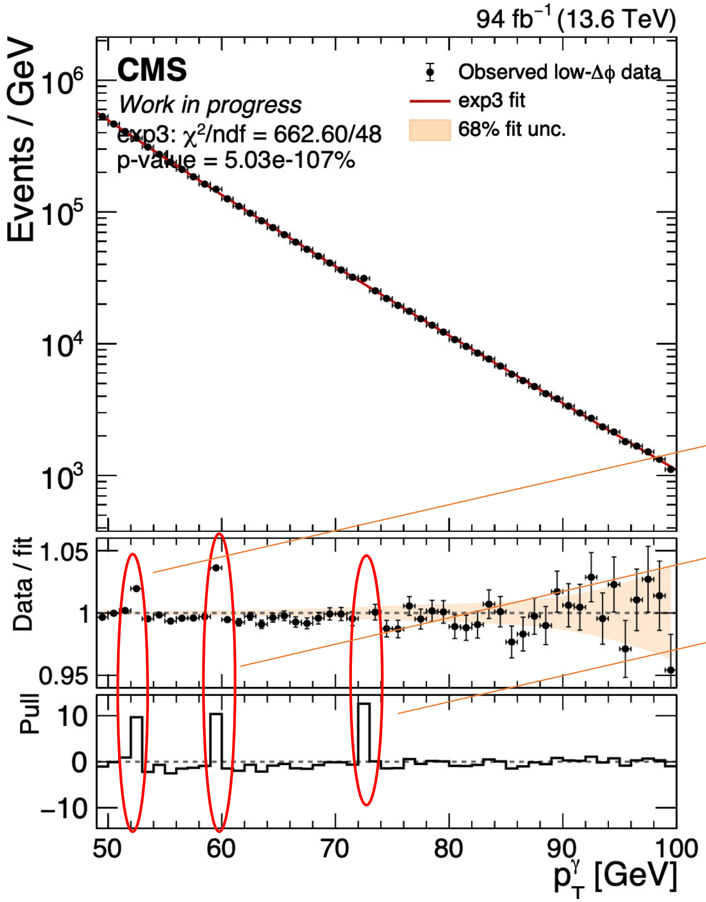
- QCD multijet
  jet → photon misidentification and instrumental MET

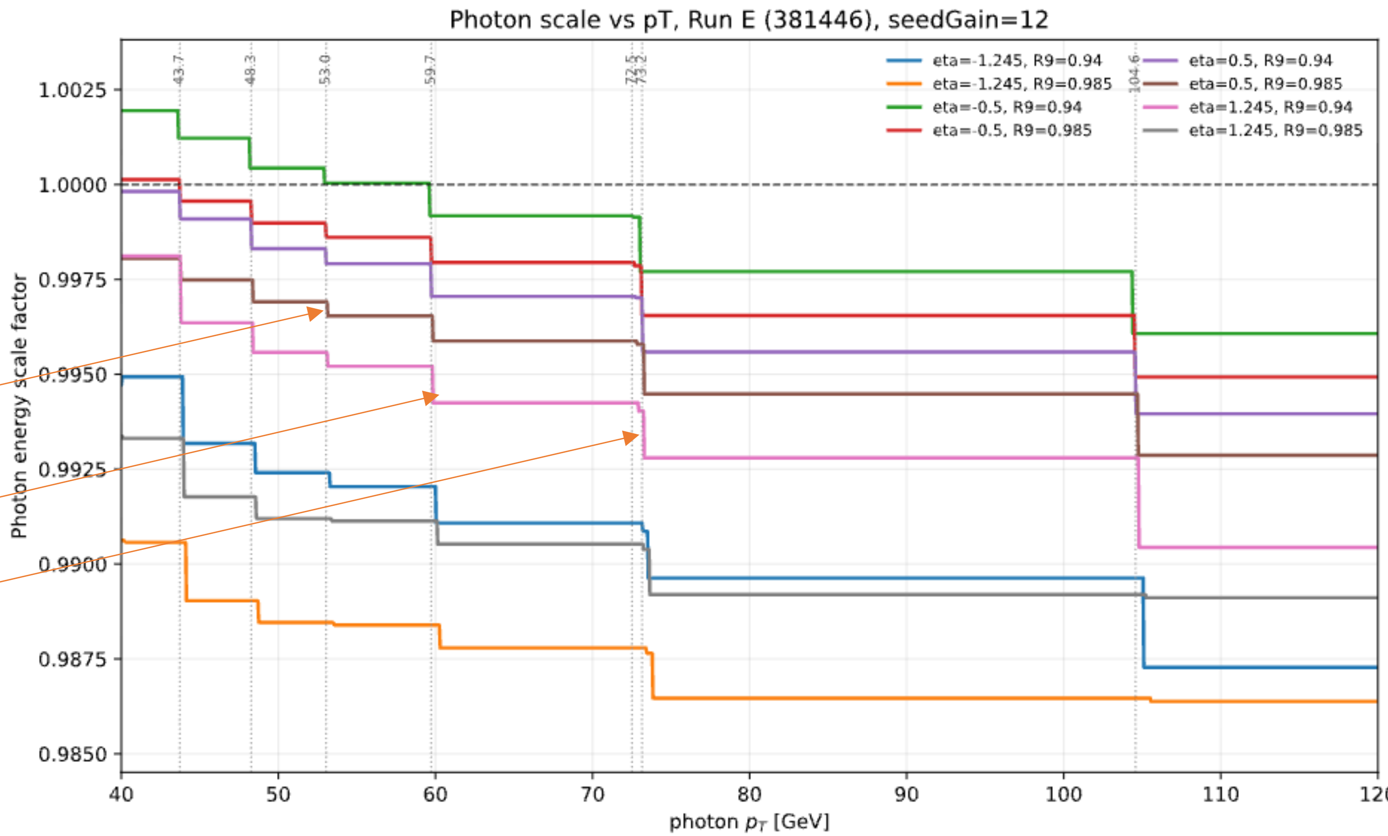
- γ+jets
  prompt photon with instrumental MET

- Minor backgrounds
  W/Z+jets, Vγ, and top-quark processes



Current status of the low-Δϕ validation

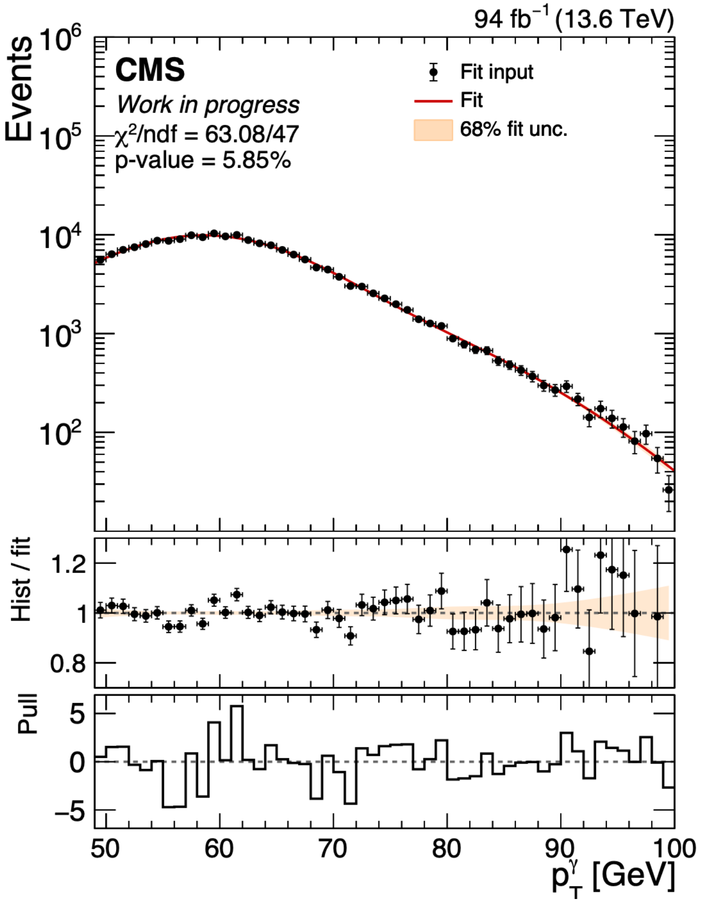




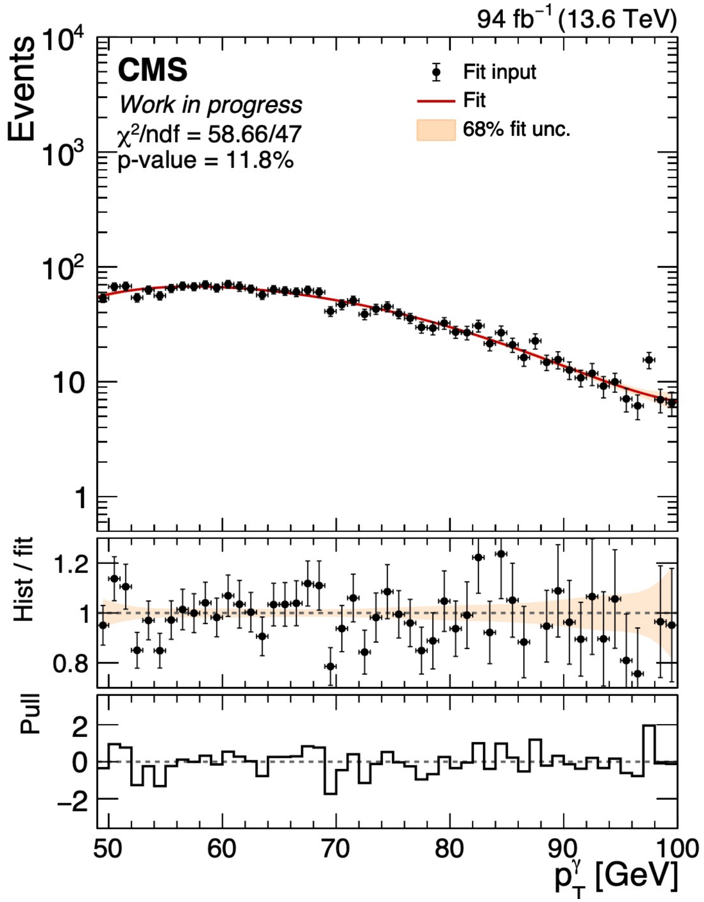
- The nominal exp3 model describe the current low-Δϕ data well except several points
- Localized residuals are aligned with discontinuities in the piecewise photon energy-scale correction
- A smooth EGM correction parameterization is being implemented before repeating the final background-model validation

Signal and background model

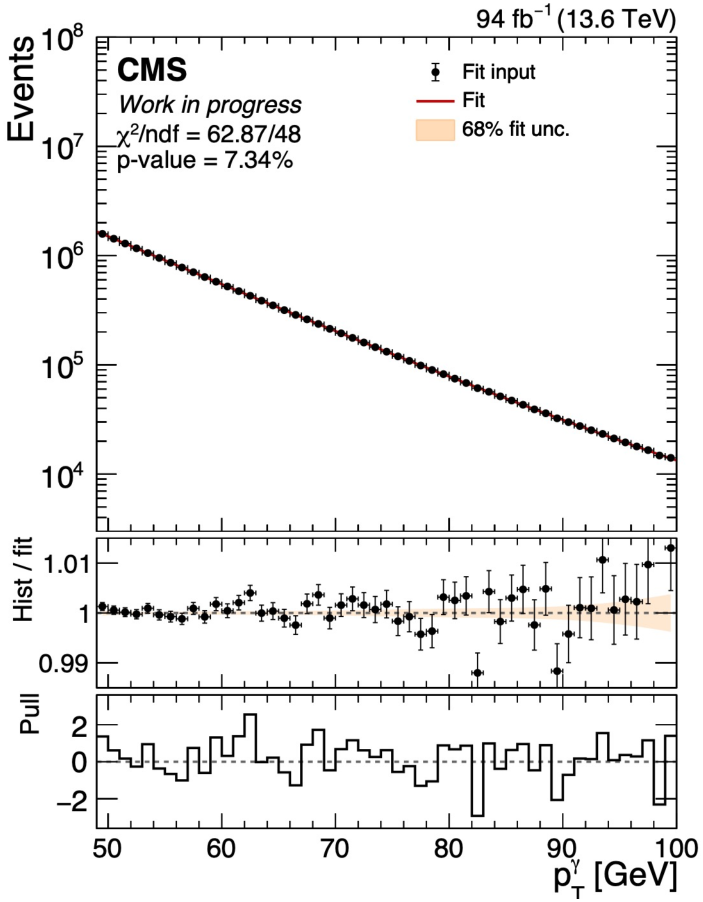
ggF signal — double Gaussian



VBF signal — Bernstein polynomial



Background — exp3



Fits to the signal templates and the background validate the chosen parameterizations and provide initial values for the likelihood fit.

Background model workflow

γ + jets MC

↓

Fit smooth RooGenericPdf

↓

Generate Poisson pseudo-data

↓

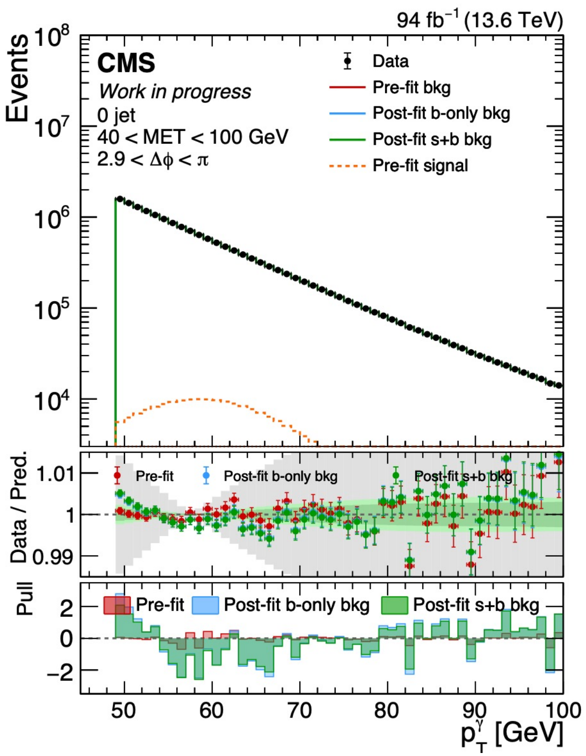
Refit the pseudo-data

↓

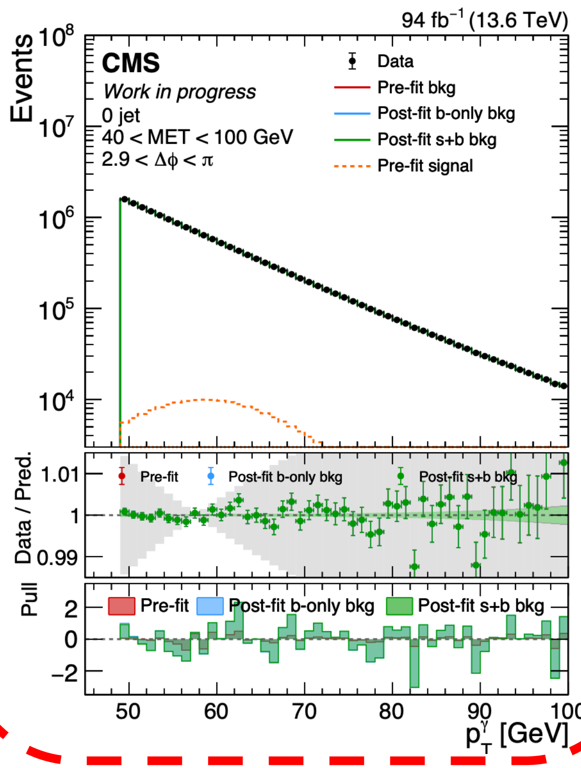
Normal card	ShapeBin card
↓	↓
Direct RooGenericPdf	Wrap same pdf function with RooParametricShapeBinPdf
↓	↓
Bin-central approximation $f(x_i)\Delta x_i$	Bin-integrated implement $\int f(x)\,dx$

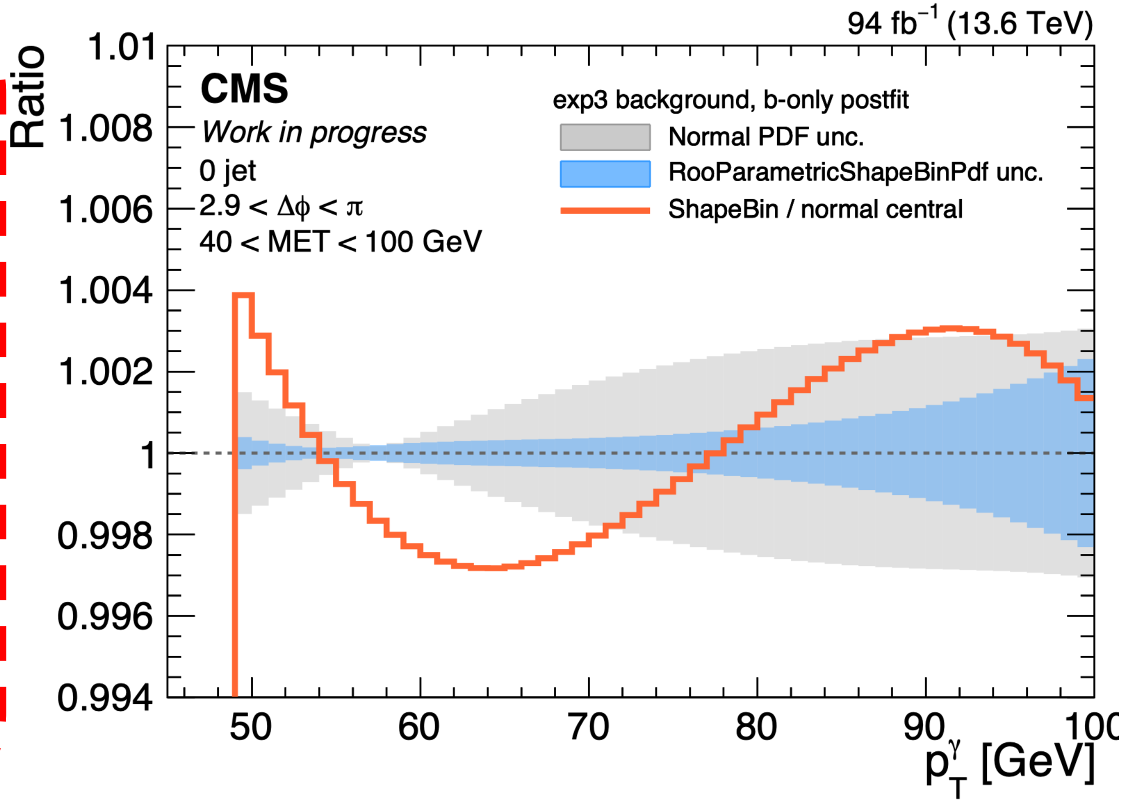
Prefit and Postfit

RooGenericPdf



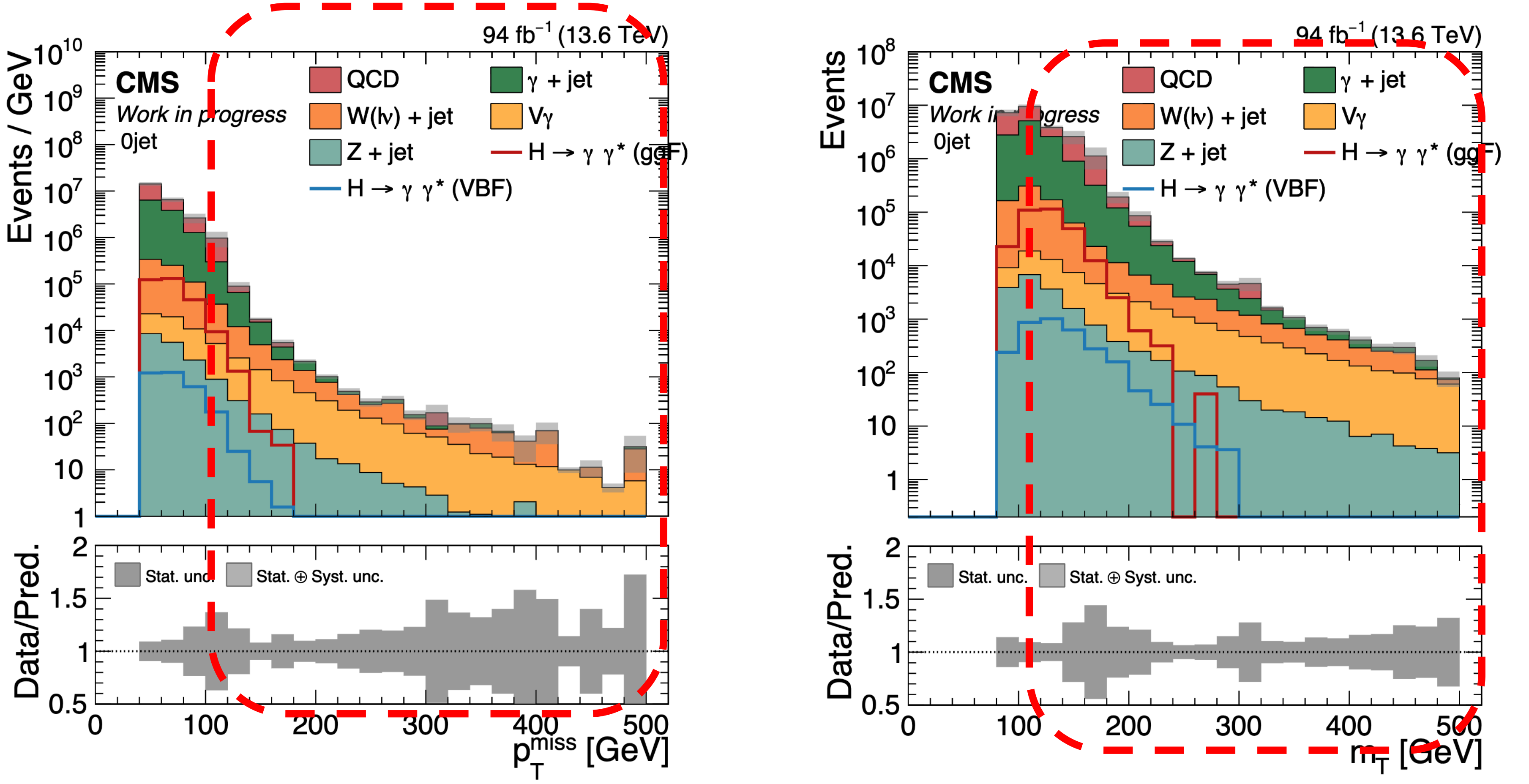
RooParametricShapeBinPdf





- Both models use the same exp3 function and the same dataset; only the bin prediction differs.
- This changes the profiled background parameters and the propagated bin uncertainty. The central predictions differ by at most ±0.4%.

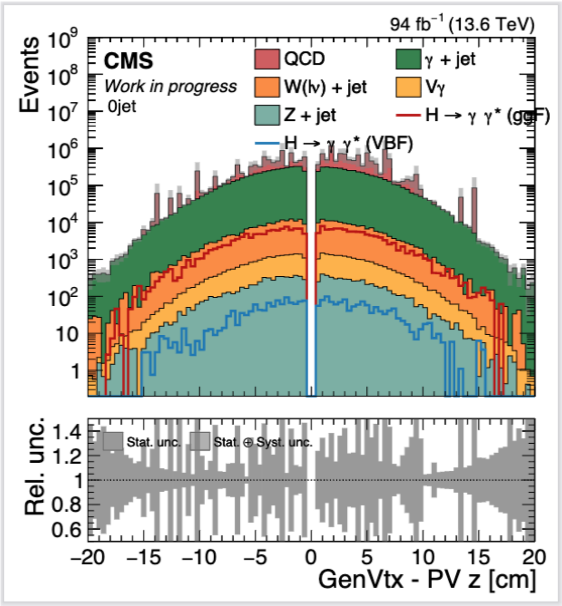
MET Tails in current 0-jet category



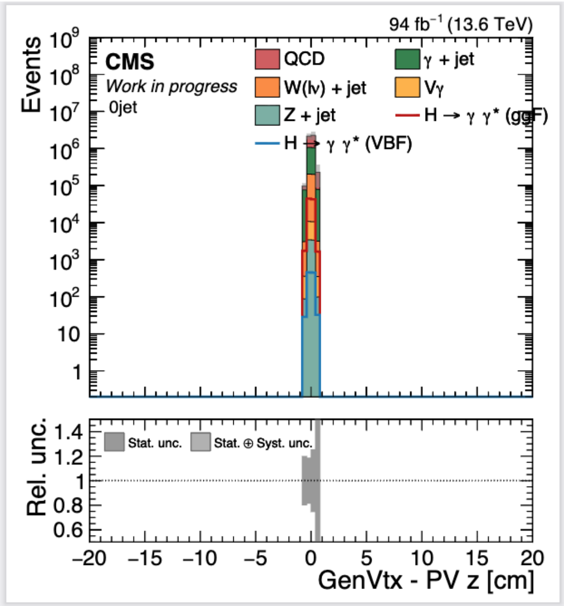
- MET>40
- The current 0-jet MC prediction shows harder high-MET and high-m_T tails than the ATLAS, driven mainly by γ+jets.

MET Tails in current 0-jet category

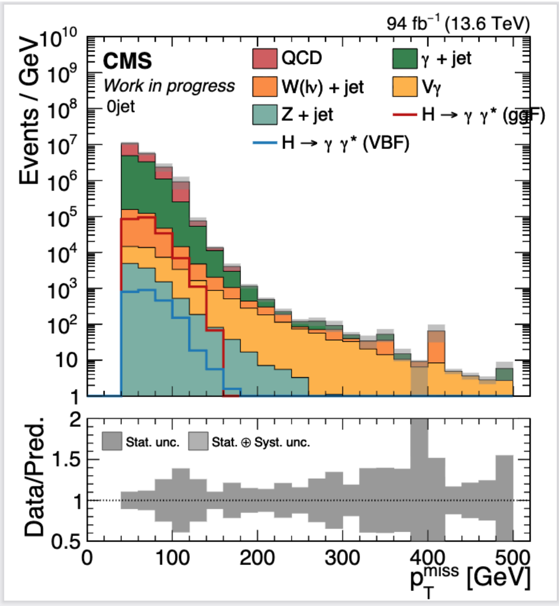
|dz| > 0.5 (PV mismatched)



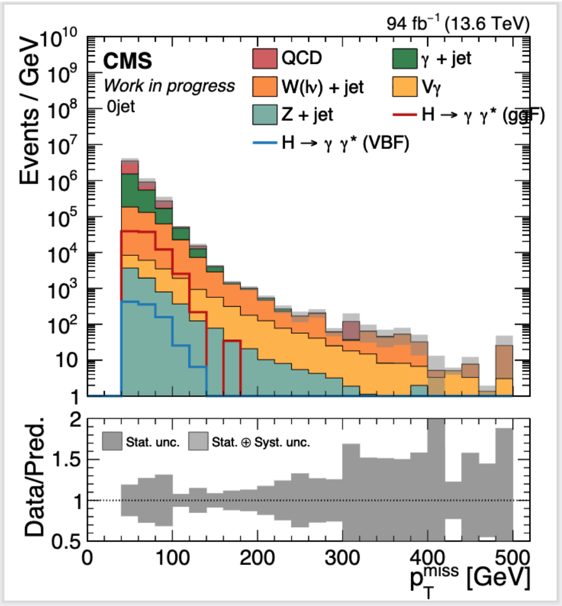
|dz| <= 0.5 (PV matched)



|dz| > 0.5 (PV mismatched)



|dz| <= 0.5 (PV matched)



- Split MC by $|Z_{gen} - Z_{PV}| <= 0.5$ cm and > 0.5 cm.
- The selected MC sample shows a large contribution from events with $|Z_{gen} - Z_{PV}| > 0.5$.
- PV matching improves the high-tail to some extent, but does not remove it.

Summary and Outlook

- Run 3 low-threshold single photon triggers enable a new CMS search for $H \rightarrow \gamma\gamma_{\mathrm{D}}$ targeting the ggF topology.

- The current baseline uses 93.6 fb^{-1} of 2024 data and focuses on the 0-jet category with $40 < p_{\mathrm{T}}^{\mathrm{miss}} < 100\mathrm{GeV}$.

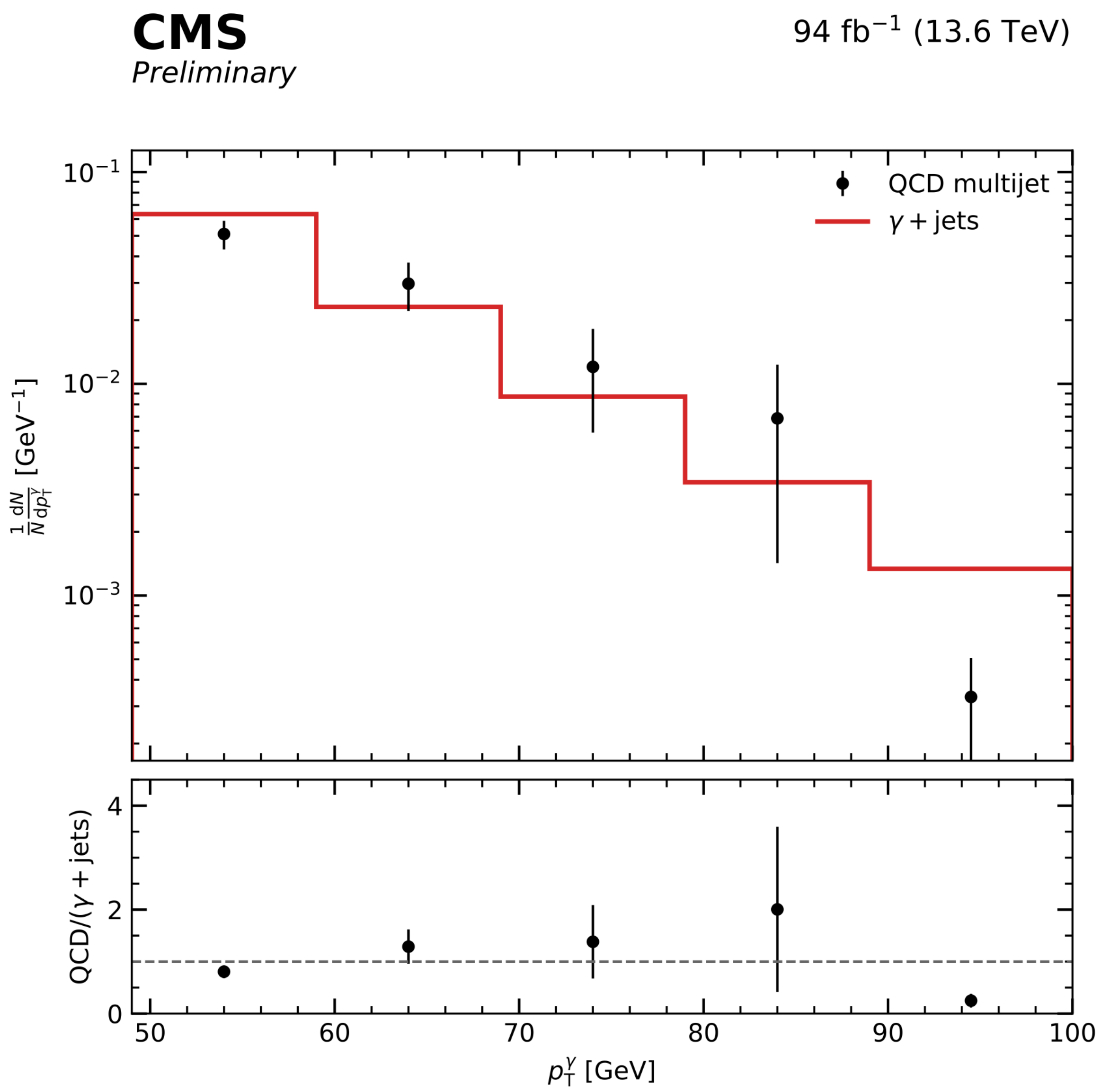
- The current blinded implementation uses pseudo-data for model development; the final total background will be determined directly from data.

- Next steps
  - Background-model bias tests;
  - instrumental MET tail studies;
  - Enable smooth EGM correction;
  - integration of the 2025 dataset.

- With the analysis note ready, we kindly request the opening of a CADI line to start the review after the summer.

Back up

Similarity between Gjets and QCD multijet



CMS iihe BRUXELLES BRUSSEL

Background model families

ULB BOSTON UNIVERSITY

1. Exponential polynomials of orders $n = 2, 3, 4$, and 5:

$$f_{\exp,n}(x) = \exp\left(\sum_{k=1}^{n} a_k \tilde{x}^k\right). \quad (3)$$

2. Bernstein polynomials of orders $n = 5, 6$, and 7:

$$f_{\mathrm{Bern},n}(x) = \sum_{k=0}^{n} c_k \binom{n}{k} t^k (1-t)^{n-k}, \quad (4)$$

   where c_0 is fixed to unity and the remaining coefficients are constrained to be positive.

3. Power-law functions multiplied by exponential polynomials, with $n = 2, 3, 4, 5$, and 6 free shape parameters:

$$f_{\mathrm{powexp},n}(x) = \left(\frac{x}{100\,\mathrm{GeV}}\right)^q \exp\left(\sum_{k=1}^{n-1} a_k \tilde{x}^k\right). \quad (5)$$

4. A sum of two exponential functions:

$$f_{2\exp}(x) = f \exp(a\tilde{x}) + (1-f)\exp(b\tilde{x}), \quad (6)$$

   where $0 \le f \le 1$.

Alternative background models

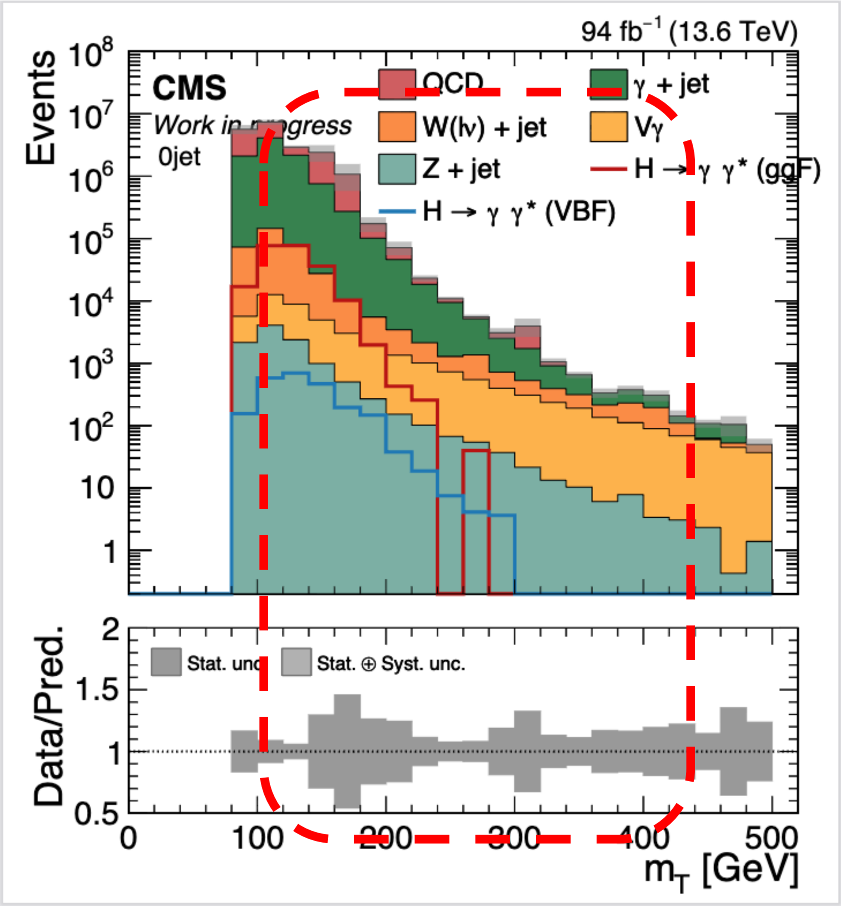
Signal and background model

- **Nominal analytic (RooGenericPdf):**
  data_obs
  γ + jets → fitted analytic PDF → Poisson pseudo-data
  background
  direct RooGenericPdf
  → bin-central approximation in the binned Combine model

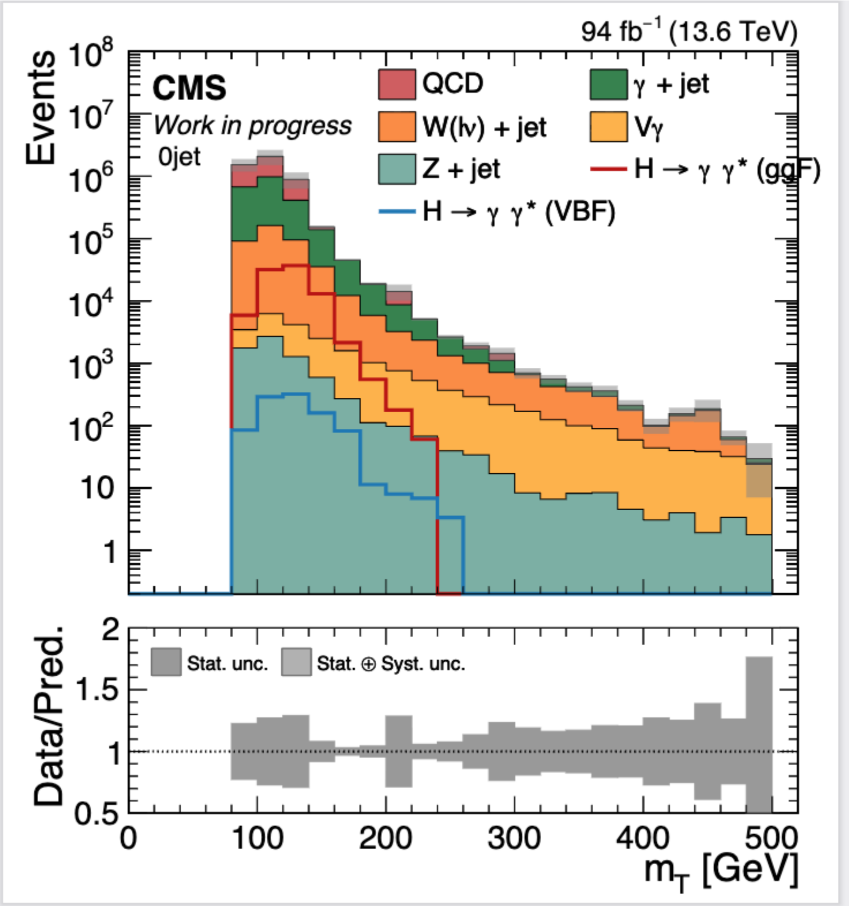
- **Shapebin analytic (RooParametricShapeBinPdf):**
  data_obs
  exactly the same pseudo-data as the corresponding normal card
  background
  same analytic function and the same fitted parameters
  → wrapped by RooParametricShapeBinPdf after the pseudo-data fit
  → exact bin-integrated prediction in the binned Combine model

MET Tails in current 0-jet category

|dz| > 0.5 (PV mismatched)

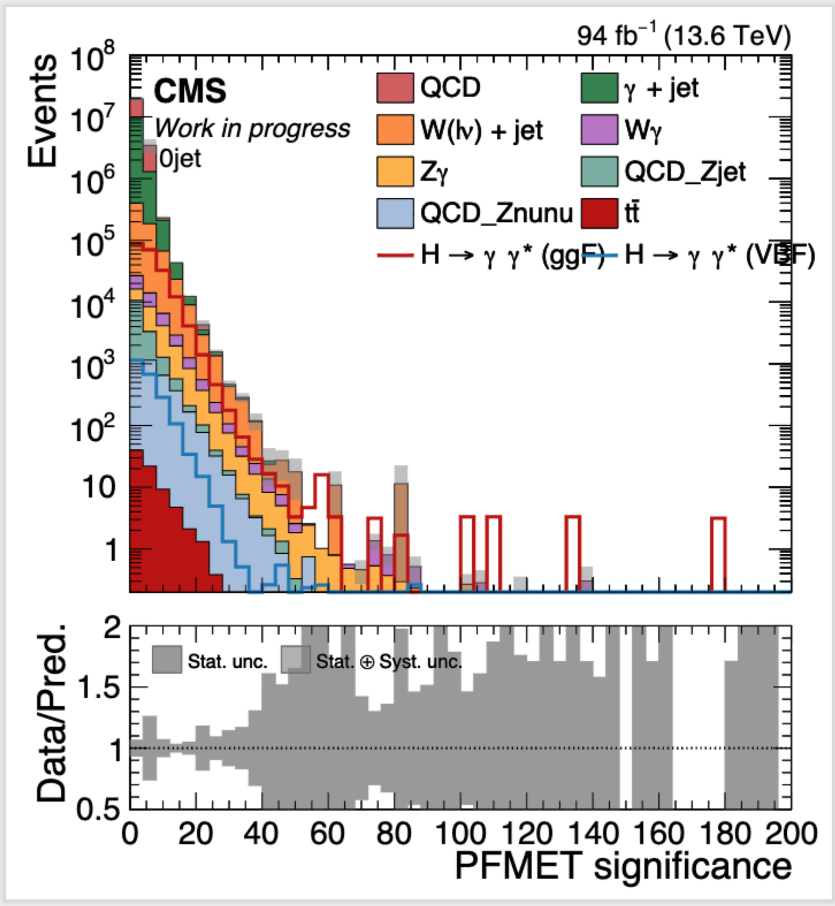


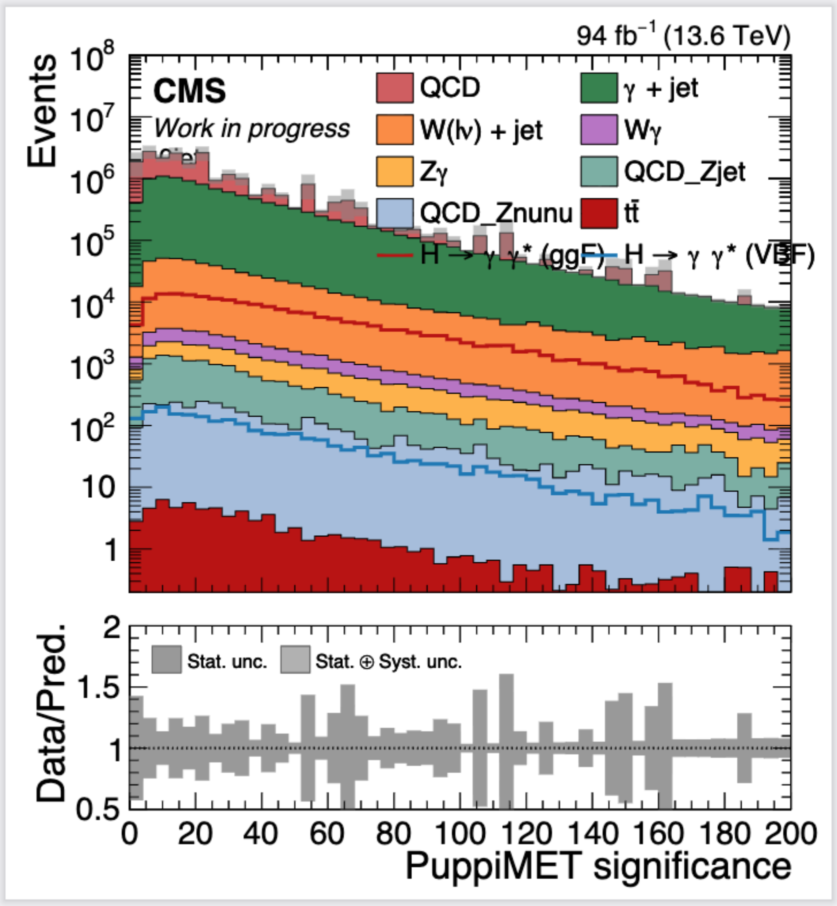
|dz| <= 0.5 (PV matched)



PV matching improves the high-tail to some extent, but does not remove it.

MET Tails in current 0-jet category





PF and PUPPI MET significance exhibit substantially different numerical ranges and shapes.