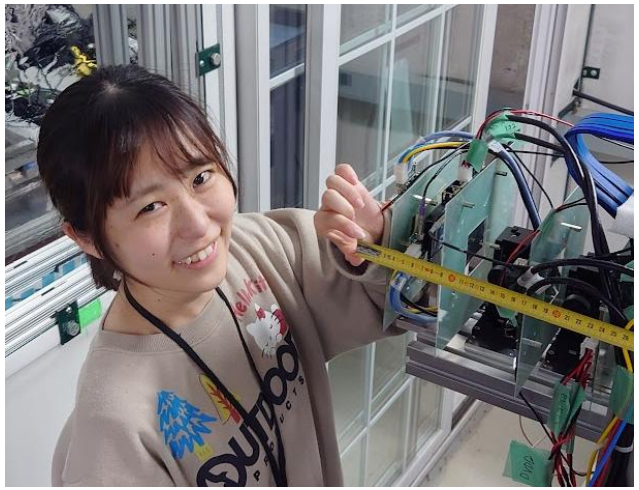
My activities for $HH \to b\bar{b}\gamma\gamma$

Sayuka Kita
(University of Tsukuba)

筑波大学
University of Tsukuba

Self-introduction

Sayuka Kita



- Ph.D. student in University of Tsukuba, Japan (will graduate in March 2026)
- Researching in HEP lab under Shigeki Hirose and Fumi Ukegav
- Have been contributing in ATLAS experiment since 2023
  - ✓ SCT operation team
  - ✓ Member of Run 2 + partial Run 3 $b\bar{b}\gamma\gamma$ analysis
- Collaborating with KEK IPNS (Kazuki Kojima)

My contributions for $b\bar{b}\gamma\gamma$ analysis

- Validation study on Run 2 vs Run 3, data and MC
- BDT input variable optimization
- Evaluation of photon conversion systematics (CP)

- Postdoc in KEK IPNS
- BDT hyperparameter optimization
- Evaluation difference between FS vs AF3

Run 2 + partial Run 3 $b\bar{b}\gamma\gamma$ analysis

- Overview

Glance, Int note

  - Target: HiggsPairs2025 (May 2025)
  - Data: Run 2(data15-18) + Run 3(data22, 23) (data24 is under discussion)
  - Status: 2nd EB request on 25th Feb. → Unblined on 14th Mar.

Difference setup from Legacy analysis

	Legacy analysis	Run 2 + p Run 3
DAOD derivation	HIGG1D1	PHYSLITE
b-jet selection	Exact 2b-jets @ DL1r 77%	At least 2b-jets @ GN2 85%
Kinematic fit	-	Yes

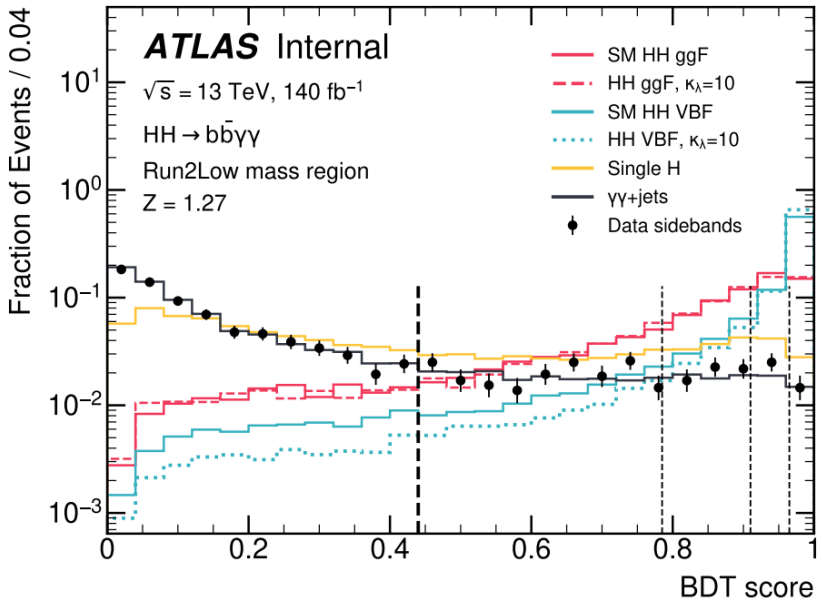
Sensitivity

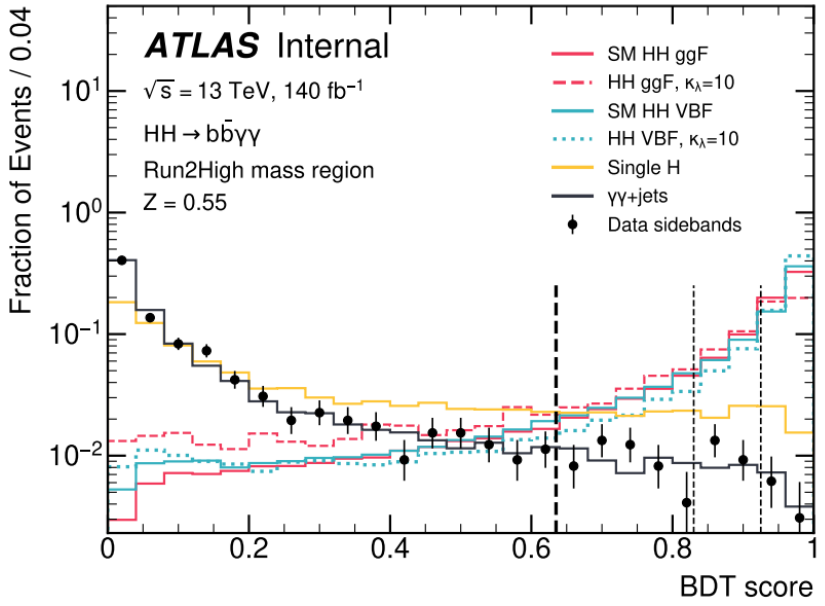
	Legacy analysis	Run 2 + p Run 3
Significance	0.54	0.67 **(+ 24%)**
UL on μ_{HH}	4.86	3.72 **(- 24%)**
κ_λ limit	[-2.7, 7.6]	[-2.4, 7.6]
κ_{2V} limit	[-1.1, 3.3]	[-0.9, 3.1]

- My contributions
  - BDT/categorization study
    - Input sample optimization
    - Hyperparameter check
  - Systematics evaluation (photon conversion systematics)

BDT and categorization

- We use BDT after preselection to separate signal (HH) and background (H+ $\gamma\gamma$+jets)
  - High mass region ($m^*_{b\bar{b}\gamma\gamma} > 350$ GeV): Target for **SM signal**
  - Low mass region ($m^*_{b\bar{b}\gamma\gamma} \leq 350$ GeV): Target for **LM signal**
- Categorized $m_{\gamma\gamma}$ fit region to make counting significance maximum
  - 3 regions for HM: Target for $\boldsymbol{\kappa_\lambda = 1}$ HH signal
  - 4 regions for LM: Target for $\boldsymbol{\kappa_\lambda = 10}$ HH signal



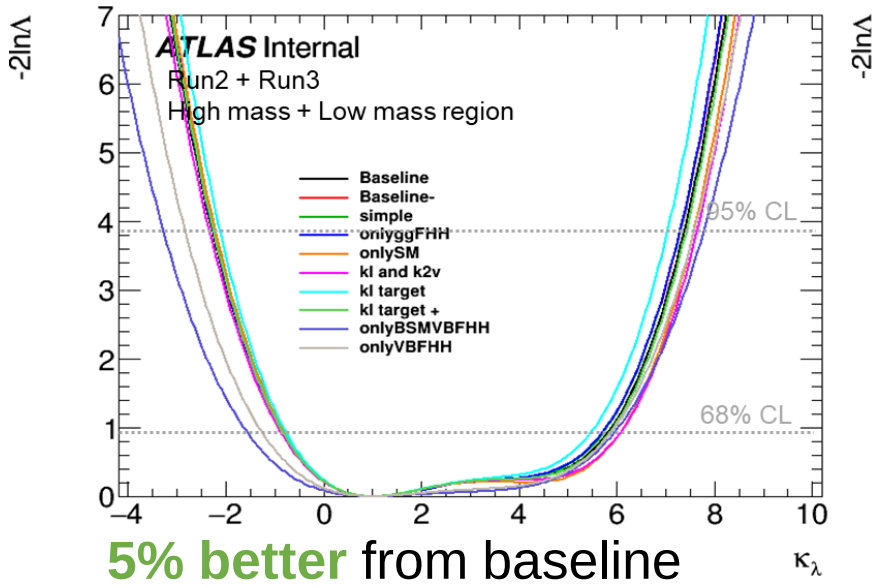


Need to optimize BDT input signal and categorization target to get better sensitivity for κ_λ?

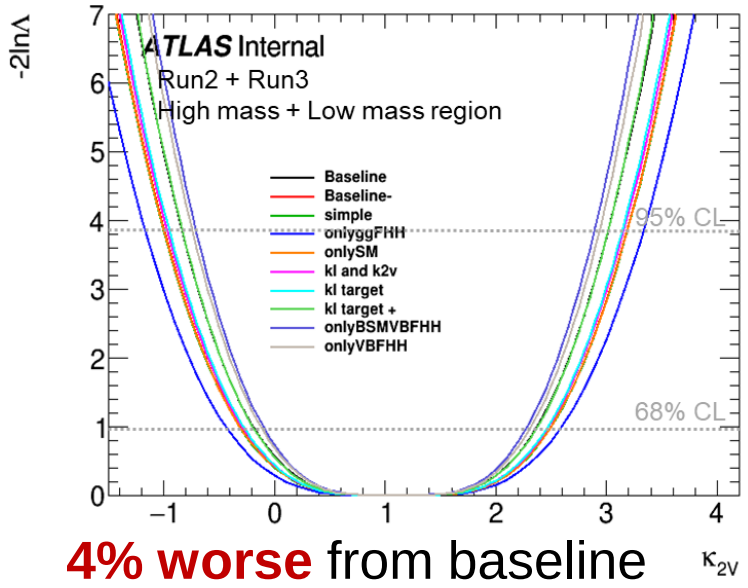
Input sample optimization for BDT

We tested a simpler “kl target” scenario where HM corresponds to κ_λ=1 and LM to κ_λ=5

	High mass input	Low mass input
Legacy analysis (baseline)	◆ SM ggFHH κ_λ=1 ◆ SM VBFHH κ_λ=1 ◆ BSM VBFHH (κ_λ, κ_{2V}, κ_V) = (0, 1, 1) (1, 1.5, 1) (1, 3, 1) (-5, 1, 0.5) (10, 1, 1)	◆ BSM ggFHH (κ_λ=5, 10) ◆ BSM VBFHH (κ_λ, κ_{2V}, κ_V) = (0, 1, 1) (1, 1.5, 1) (1, 3, 1) (-5, 1, 0.5) (10, 1, 1)
kl target scenario	◆ SM ggFHH κ_λ=1 ◆ SM VBFHH κ_λ=1	◆ BSM ggFHH **(κ_λ=5)** ◆ BSM VBFHH (κ_λ, κ_{2V}, κ_V) = (**5, 1, 1**) reweight



5% better from baseline



4% worse from baseline

* We tested several scenarios: detail

If we used “kl target” set for BDT and categorization, it will be improved 5% for κ_λ

But we decided to set same input sample as legacy analysis…

Photon conversion systematics

- We apply different MVA for converted and unconverted photon when reconstructing energy of photon
- Conversion fractions of data and MC are different
  → Single systematic uncertainty acting on the photon energy scale in E/Gamma tool
- Evaluated photon energy bias for Run 3

$$\text{Bias}\,(|\eta|, p_T) = \left\langle \frac{E}{E^{true}}(|\eta|, p_T) \right\rangle_{Z\to ll\gamma\ data} - \left\langle \frac{E}{E^{true}}(|\eta|, p_T) \right\rangle_{Z\to ll\gamma\ MC}$$

We use Z → lly for evaluation
- ☺ high purity
- ☹ **Limited eta, p_T range**

- Data-like E/Etrue can be reproduced by scaling conversion fractions in MC to those in data.
- Conversion fraction of Z→llγ can be evaluated by fitting E1/E2 from calorimeter energy ratio value
- Single photon MC sample is used for evaluation: scale conversion fraction to match to Z →llγ mc and data

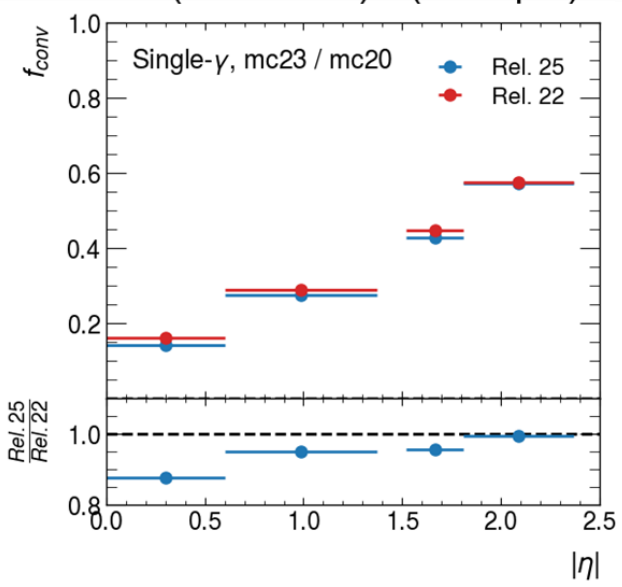
This effect will be used for systematics of $m_{\gamma\gamma}$ peak position 0.3-0.7%

Single photon conversion fractions

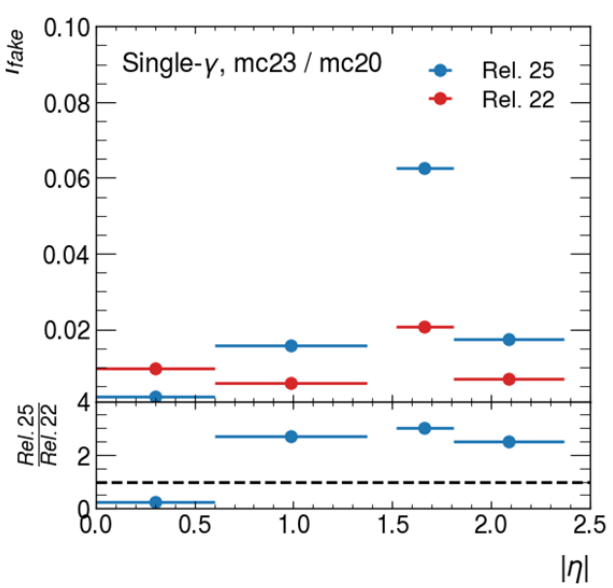
Conversion photon requirement

- 0 mm < Rconv < 800 mm
- From Run 3: Converted photon in TRT barrel region is not regarded as converted photon
  (TRT gas was changed → fake converted photon was increased)

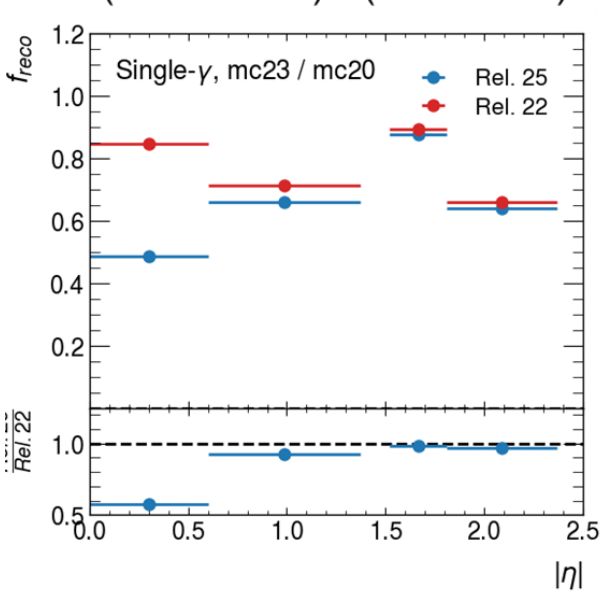
f_{conv}
= N(True conv.)/N(True ph.)



f_{fake}
= N(Reco. conv.)/N(True unconv.)



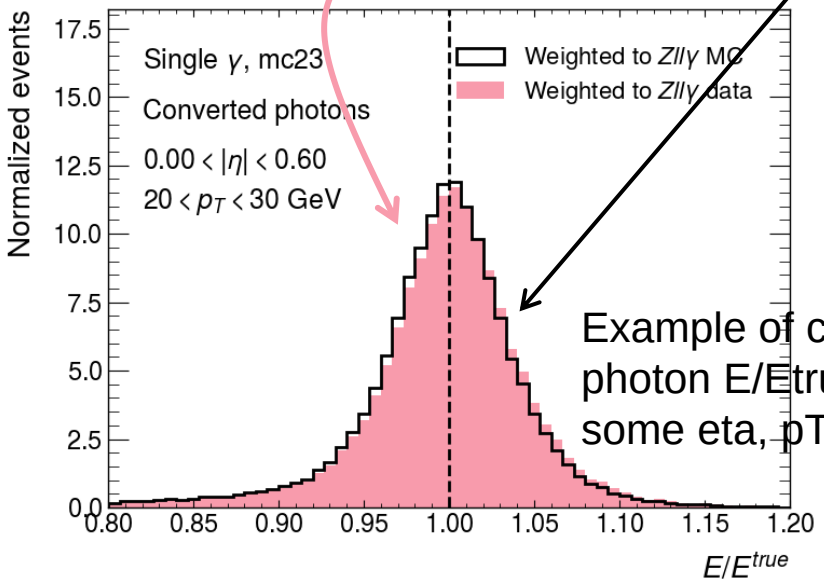
f_{reco}
= N(Reco. conv.)/N(True conv.)



- fconv is similar between rel. 22 and 25: expected
- freco and ffake difference in 0 < |eta| < 0.8 bin comes from changed TRT requirement

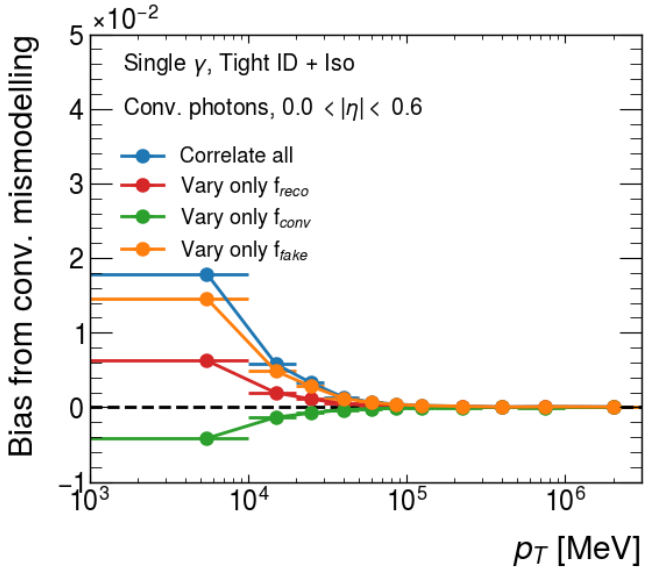
Energy scale bias

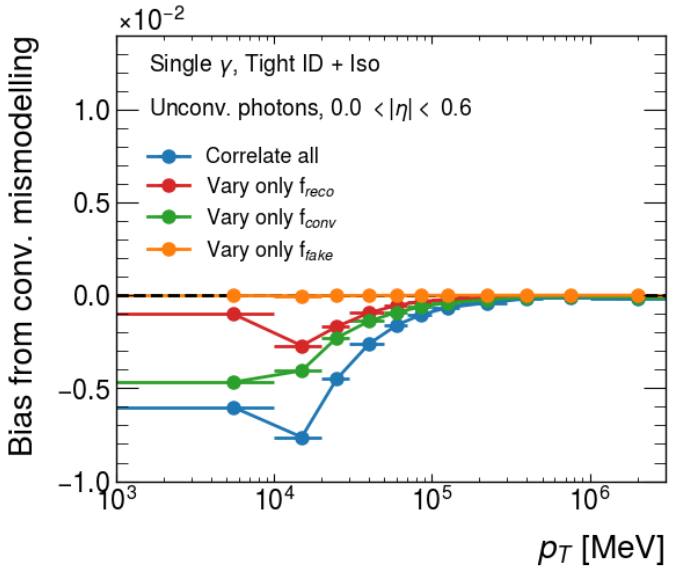
$$\text{Bias}(|\eta|, p_T) = \left\langle \frac{E}{E^{true}}(|\eta|, p_T) \right\rangle^{Z\to ll\gamma\ data} - \left\langle \frac{E}{E^{true}}(|\eta|, p_T) \right\rangle^{Z\to ll\gamma\ MC}$$



Example of converted photon E/Etrue in some eta, pT bin

We evaluated **energy scale bias**





- ✓ Similar order values and pT trend were obtained including error of Z →llγ fit data
- ✓ Conversion fraction of Z →llγ in Rel. 25 is larger than Rel. 22
  - affect when reweighting → make larger(smaller) energy scale bias
  - We are discussing to release “conservative” value as a pre-recommendation

Summary

Run 2 + partial Run 3 $HH \to b\bar{b}\gamma\gamma$ analysis is ongoing

BDT / categorization

- Evaluated sensitivity with changing BDT and categorization target
- “kl target” scenario has best sensitivity for κ_λ limitation

(This strategy wasn’t used for this analysis)

Systematics evaluation

- Evaluated photon conversion systematics in Run 3
- Difference of Run 2 and 3 conversion fractions are observed in TRT barrel region, the other feature is similar to rel. 22
- Something is unclear: We will release conservative value as a pre-recommendation

I’d like to start investigating for full Run 3 analysis (and HL-LHC)

- Photon: new photon ID and trigger for getting better signal efficiency
- Event selection: Make several BDTs for each κ_λ target

Backup

Run 2 + partial Run 3 $b\bar{b}\gamma\gamma$ analysis

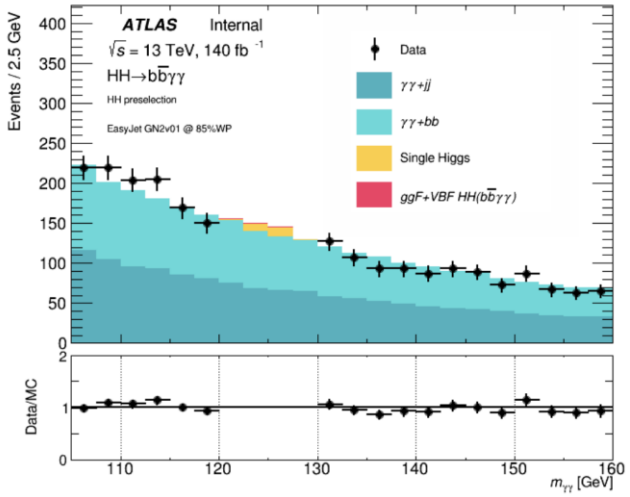
- Overview
  - Target: HiggsPairs2025 (May 2025)
  - Data: Run 2(data15-18) + Run 3(data22, 23) (data24 is under discussion)
  - Status: 2nd EB request on 25th Feb. → Unblined on 14th Mar.

Glance, Int note

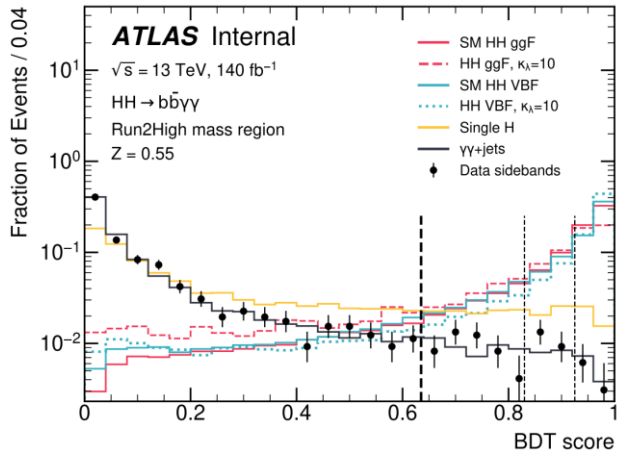
- Analysis flow

Preselection → **Event selection by BDT** → **$m_{\gamma\gamma}$ Fit**

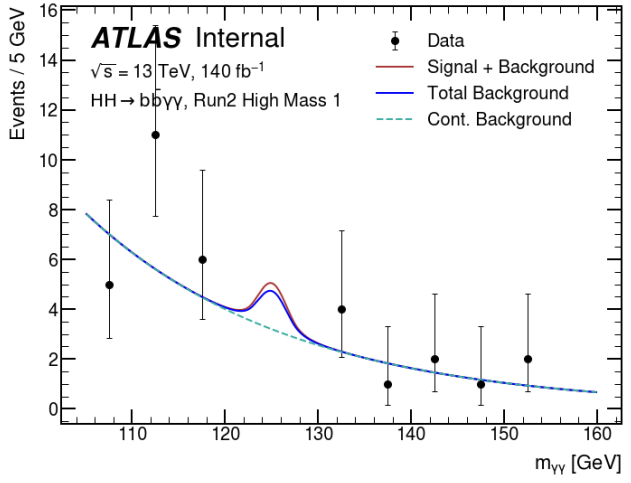
- Exact 2 iso & tight photons
- At least 2b jets @ GN2 85%



- Separate to HM (SM target) and LM (BSM target) region at $m^*_{\gamma\gamma} = 350$ GeV
- HM x3, LM x4 region



- $\gamma\gamma$jets + single Higgs + HH signal at each region

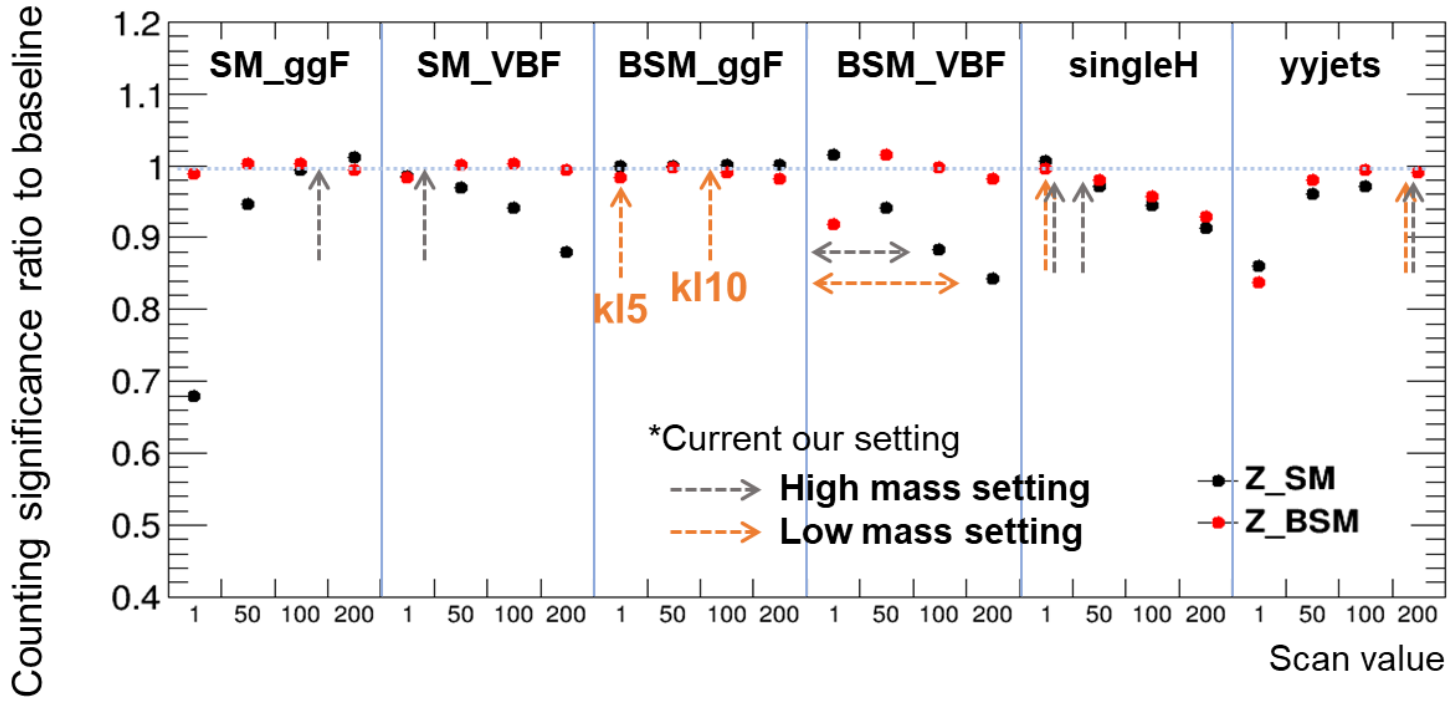


Rough Hyperparamter scan

Detail

- We decided on an analysis strategy for BDT training the same as legacy analysis
  → **Per-process weight and hyperparameters on BDT will be the same as well**
- Something is different from legacy analysis
  The number of events for BDT input, signal target for LM region ggFHH_kl10 → ggFHH_kl5…

Need to evaluate current hyperparameter setting is optimal for this analysis

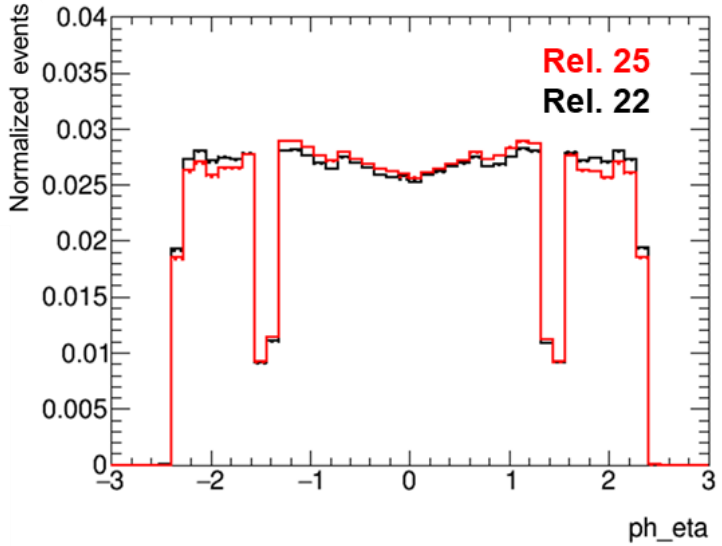


- Performed per-process weight scan and checked counting significance
- BSM VBFHH has a possibility for improving counting significance; future studies

The current setting is almost optimal

Eta comparison Rel. 22 vs 25

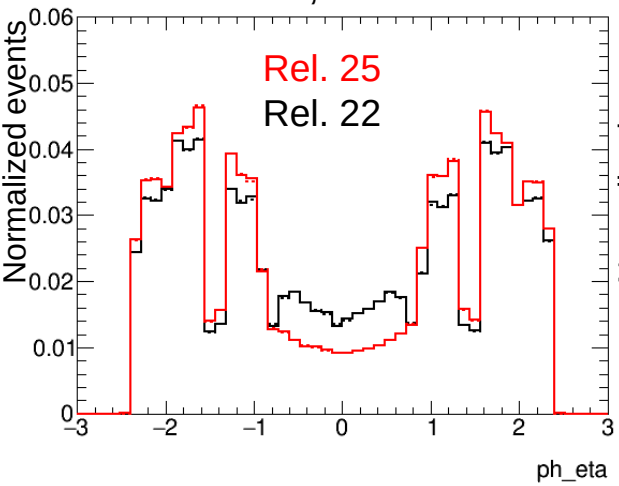
Photon eta which is passed tight & iso cut



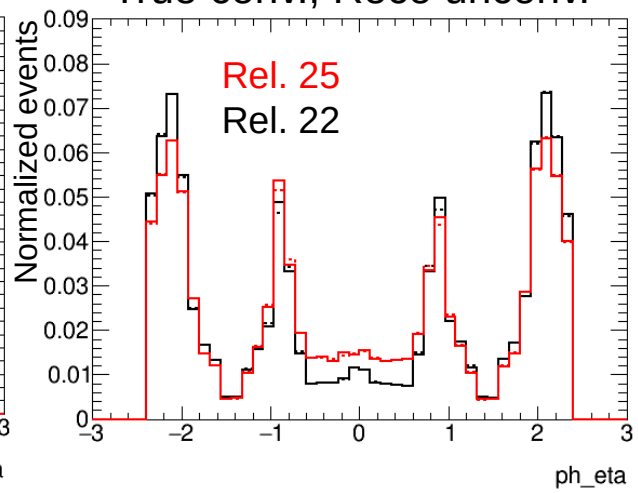
- TRT requirement from Rel. 25 regards converted photon in TRT barrel (0 < |eta| < 0.8?) as unconverted photon
- Number of reco. conv. photon decreased from rel. 22
  → That's due to TRT requirement
- If N(reco. conv.) is decreased…
  - fconv = N(True conv., reco. conv) / N(True converted) will decrease
  - ffake = N(True unconv., reco. conv.) / N(True unconv.) will decrease

freco and ffake difference comes from TRT requirement

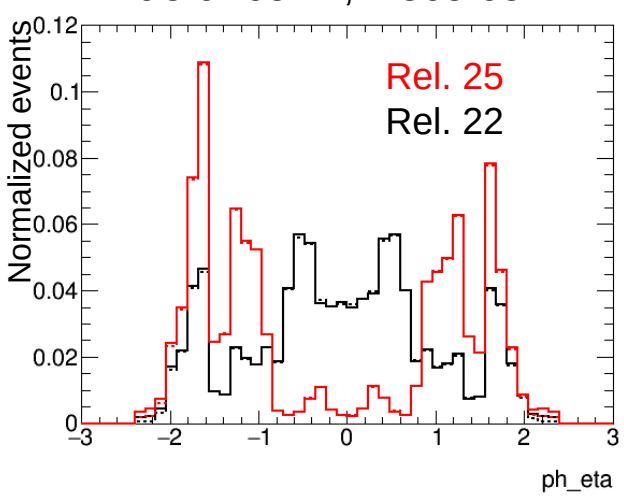
True conv., Reco conv.



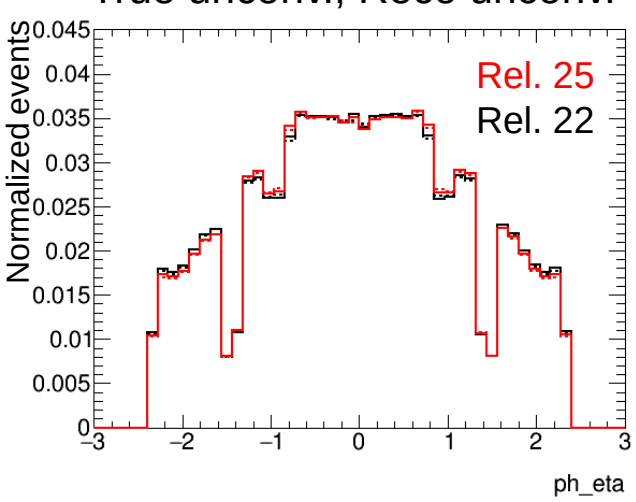
True conv., Reco unconv.



True unconv., Reco conv.



True unconv., Reco unconv.

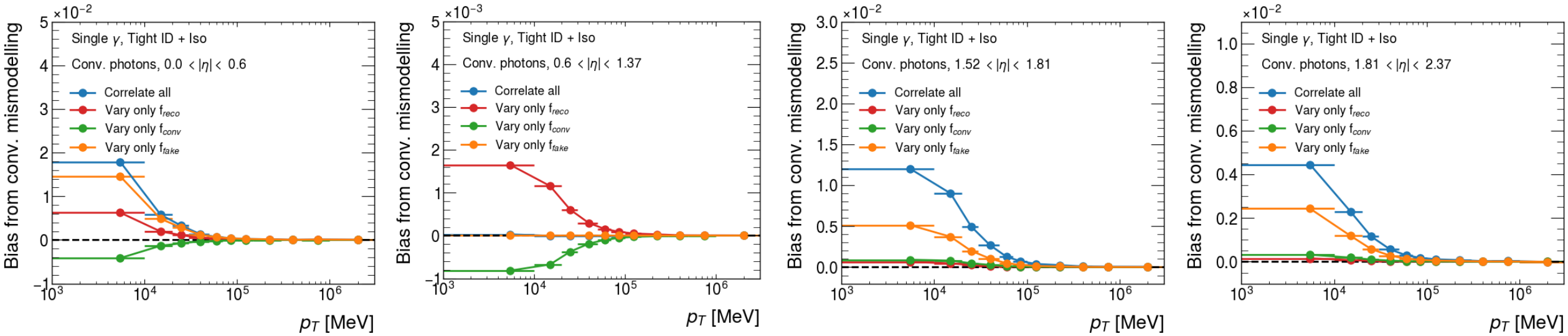


Energy bias – from each fractions

Double-check my result 2

To check which conversion fraction contributes most for energy scale bias, I reweighed only one fraction of single photon sample to Z→lly data and MC

Converted



- Combined result is close to ffake curve
  → Most contribution is ffake in Rel. 25?
- This trend is also seen in Rel. 22 result

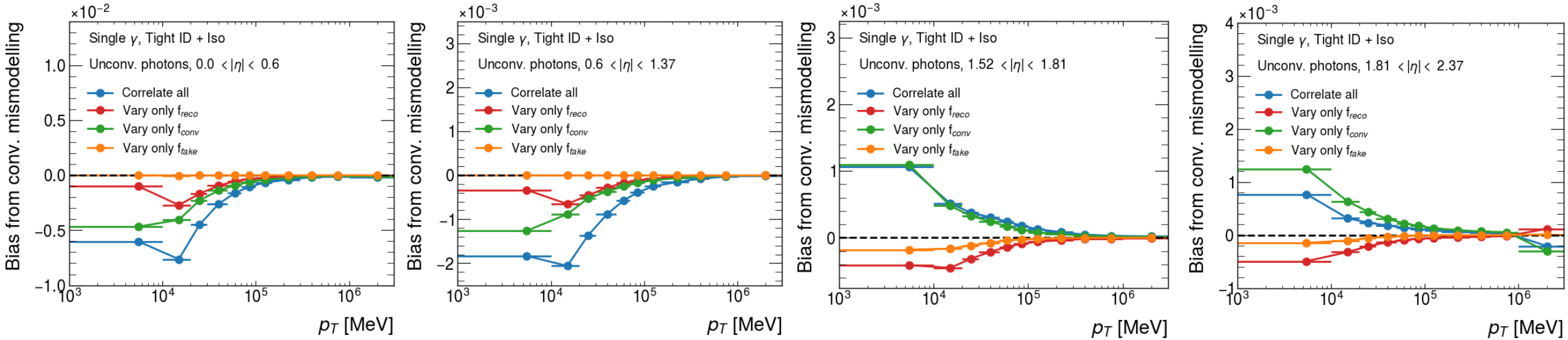
Energy scale bias from converted photon is affected by ffake

Energy bias – from each fractions

Double-check my result 2

To check which conversion fraction contributes most for energy scale bias, I reweighed only one fraction of single photon sample to Z→lly data and MC (This is also performed by Elena before)

Unconverted



- Combined result is close to fconv curve
  → Most contribution is fconv in Rel. 25?
- Rel. 22 result: close to freco (not fconv…)

Energy scale bias from converted photon is affected by freco in Rel. 25