Data-driven measurements of the prompt photon identification efficiency in ATLAS

K. Liu[1,2]
(under the supervision of Y. Liu[1] and G. Marchiori[2])

[1]USTC - Hefei
[2]LPNHE - Paris

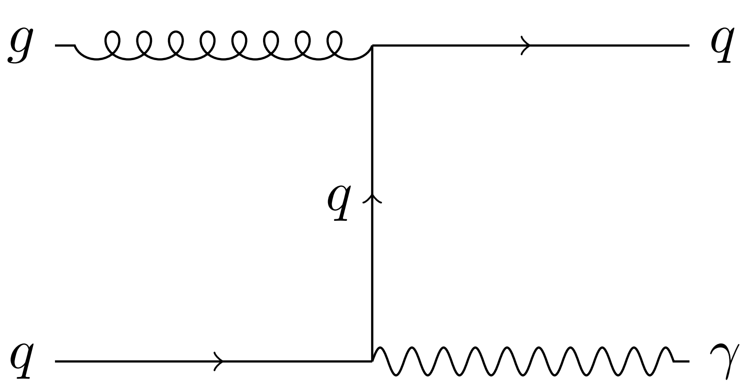
FCPPL workshop, 21/03/2012

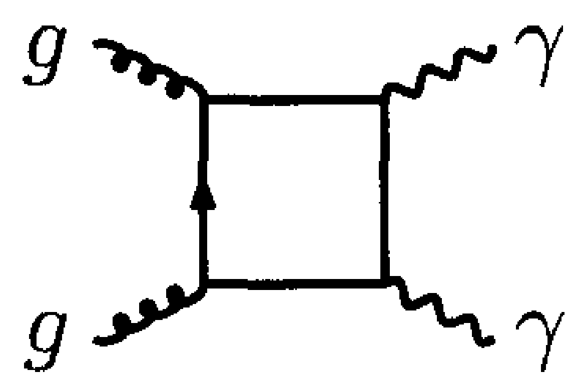
A few words about Kun

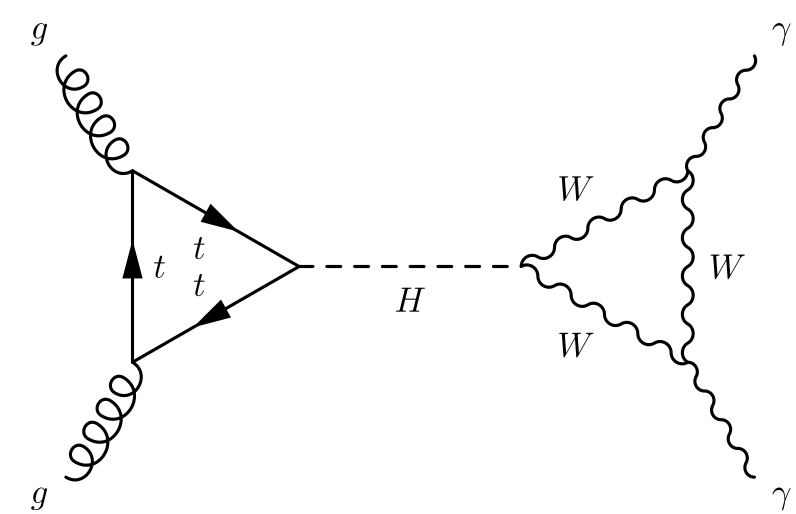
- Master at USTC
- Came to LPNHE May-June 2011 with Yanwen for a pre-thesis stage
- Stayed at LPNHE till mid Feb. 2012 - now at USTC for a few months
- Started his co-tutorship Ph.D. at UPMC (Paris-6) in September 2011
- Worked hard on
  - photon efficiency
  - photon trigger optimization for 2012 running
  - ATLAS authorship qualification work (trigger software development/debugging)
  - unfolding of SM diphoton cross section with ATLAS 2011 data

Motivation

- Prompt photons = photons not originated from hadron decays
- Measurements with prompt photons (examples):
  - inclusive γ/γγ and γ+jet xsections: tests of pQCD, extract gluon PDF
  - diphoton resonances: search for Higgs, Graviton, ...

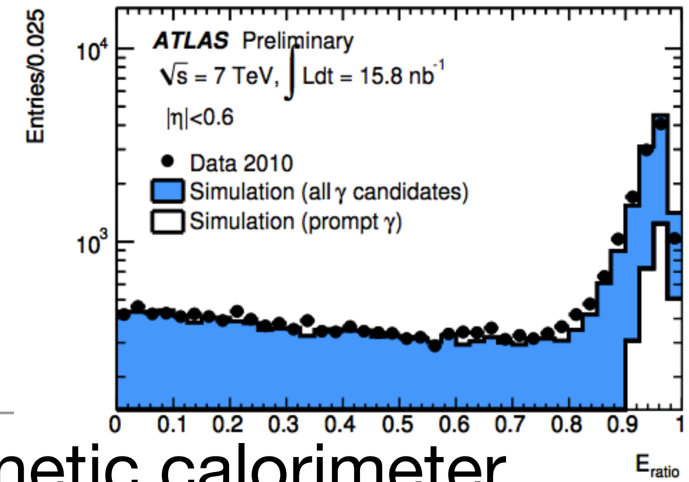


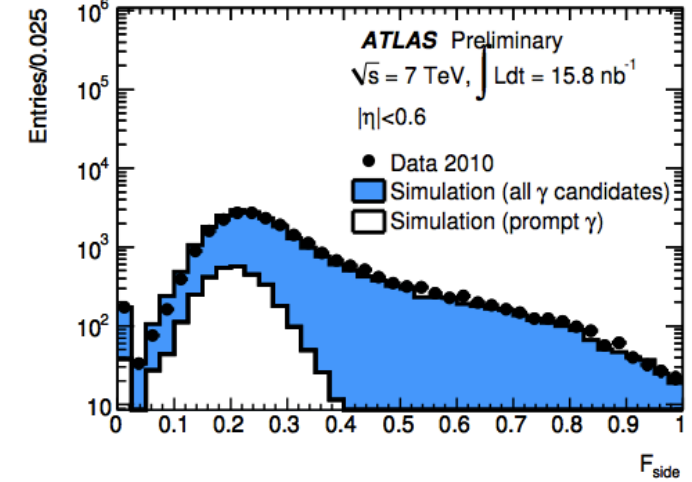




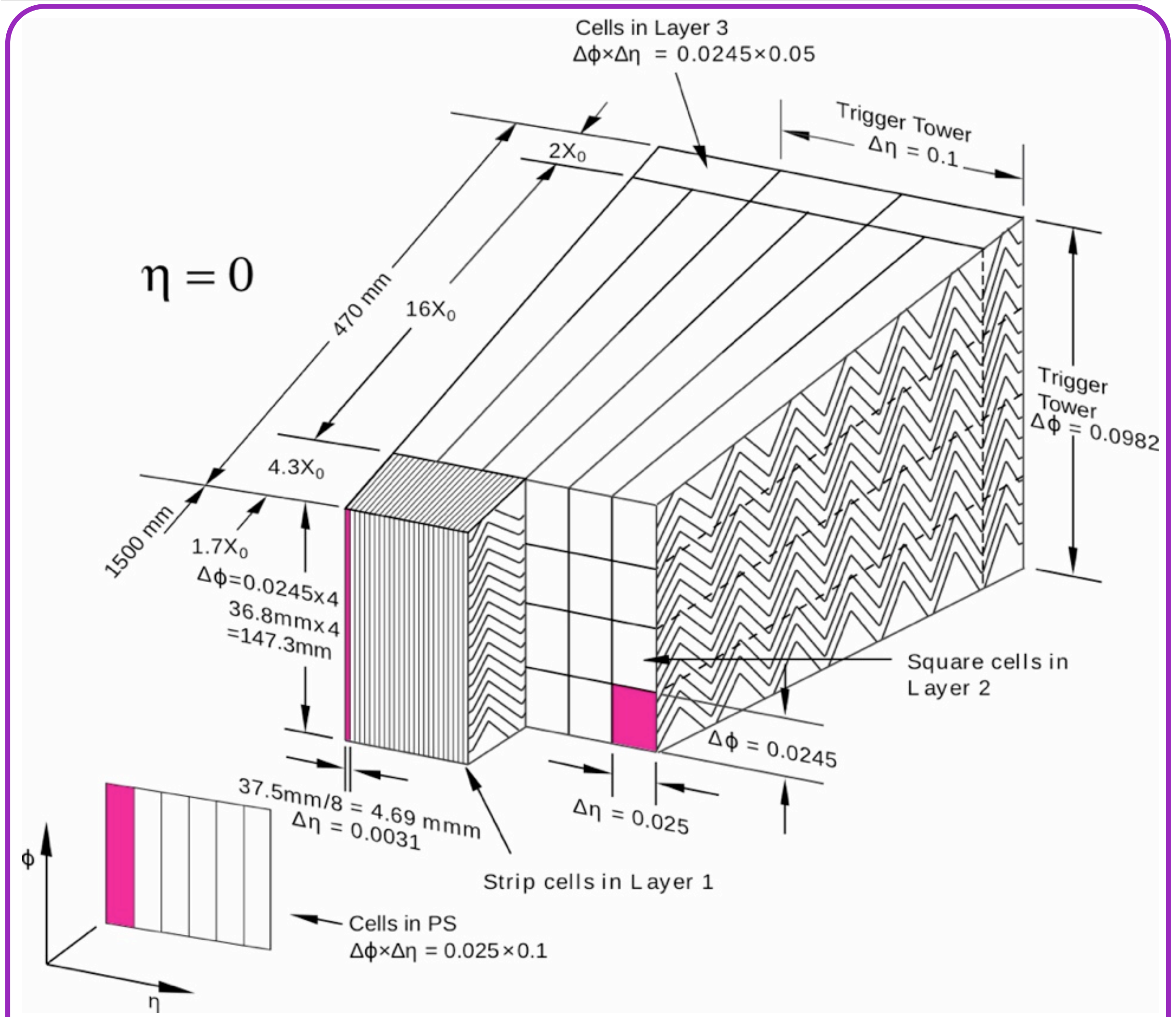
- Photon identification (ID): discriminate against large bkg from jets ($\pi^0 \rightarrow \gamma\gamma$)
- Photon (identification) efficiency needed to compute xsections and compare them to predictions (pQCD; SM vs fermiophobic Higgs; ...)

Photon ID in ATLAS

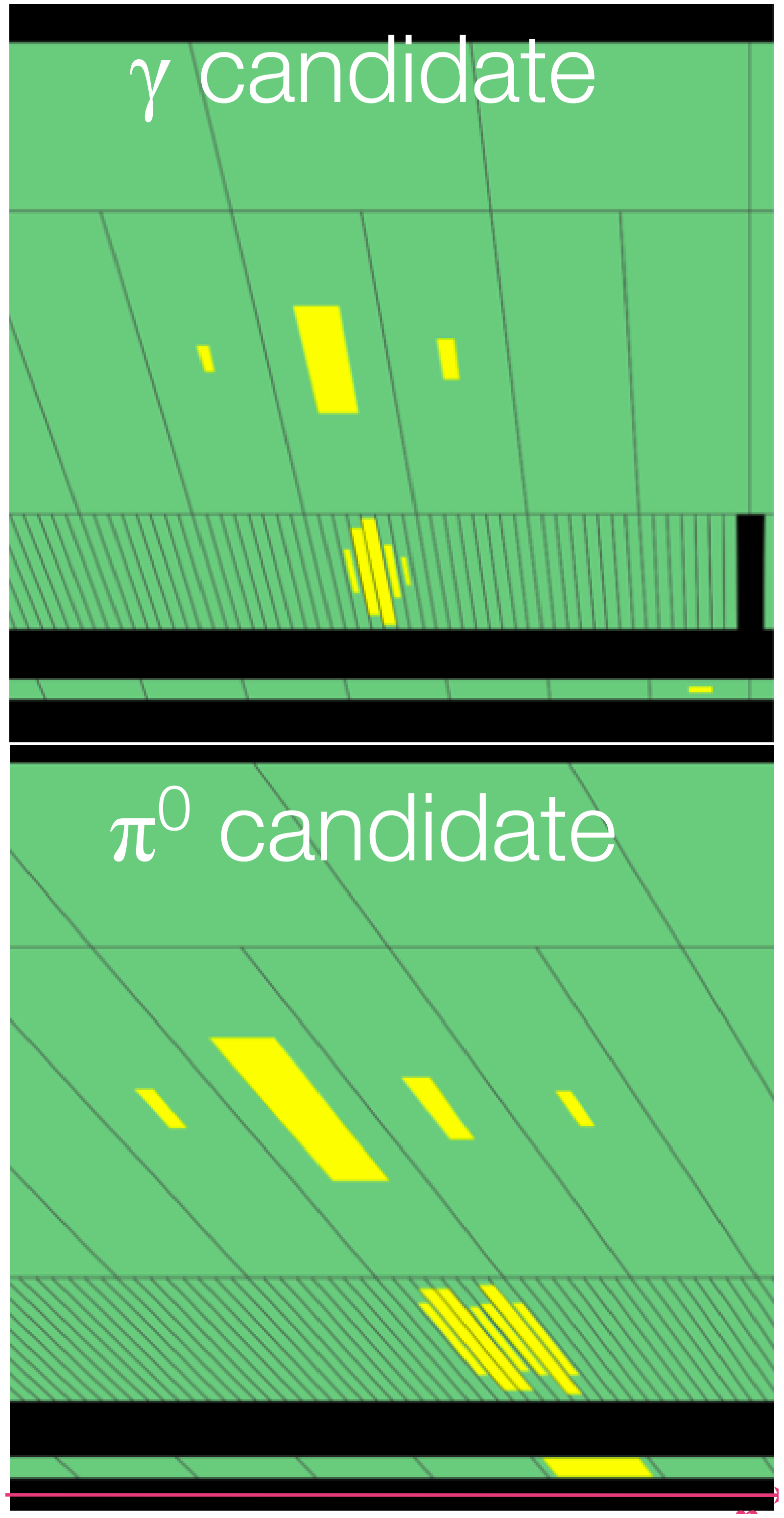




- Exploit segmentation of electromagnetic calorimeter
- Based on 9 discriminating "shower shape" variables

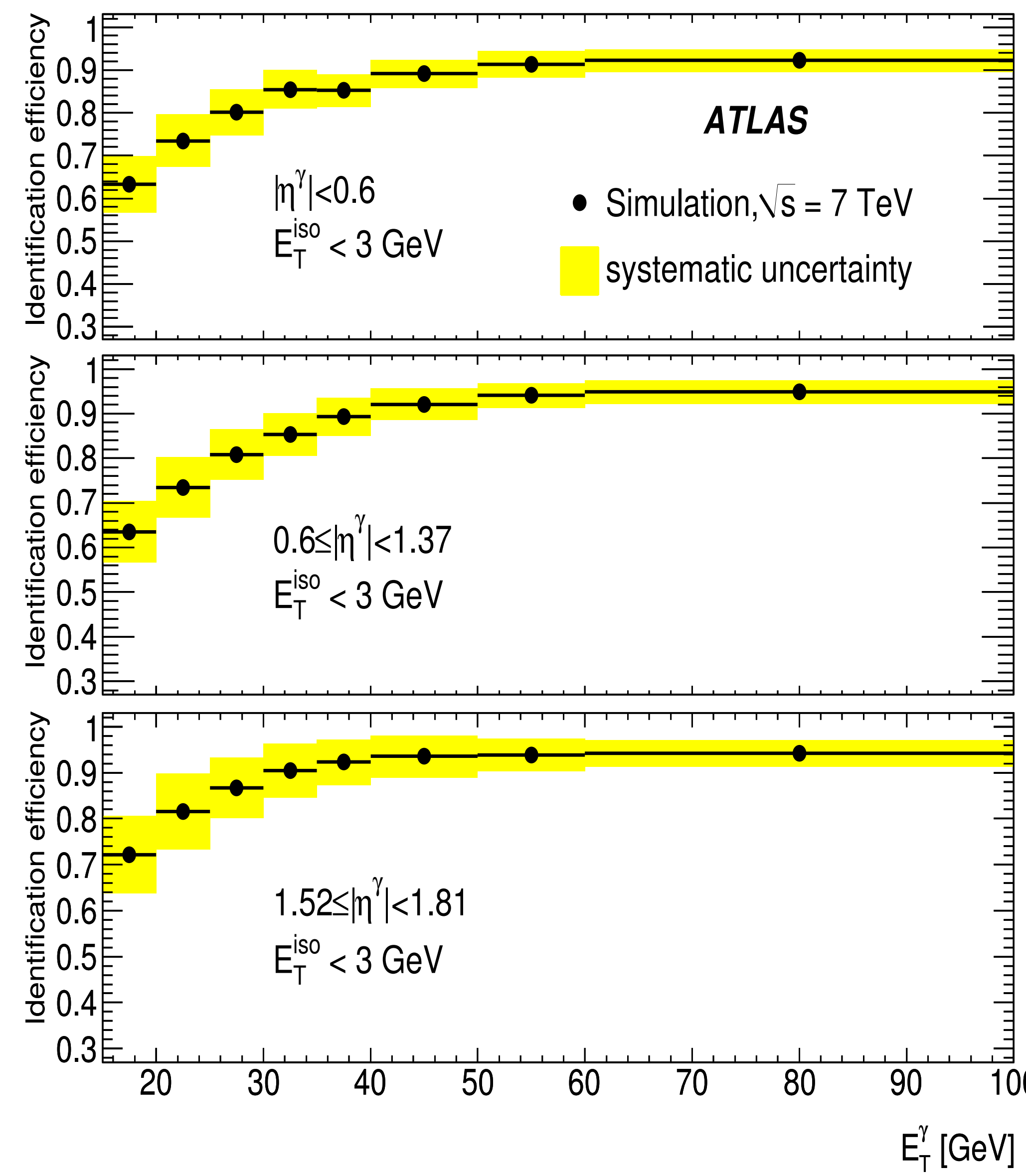


- Pb-LAr EM calorimeter
  - very fine layer 1 segmentation up to |η|=2.37
  - ✓ γ energy/direction measurement
  - ✓ π^0/γ discrimination (shower shape)



Measuring the photon identification efficiency

- No abundant, clean source of photons to measure efficiencies with tag&probe technique (unlike e,μ leptons: Z→ll)
  - cleanest source: radiative Z decays (Z→llγ) But: low stat (requires > fb^{-1}), limited p_T^γ
  - analysis of 2010 data (37 pb^{-1}): efficiency from simulation
    - correct ("fudge") shower shape distributions by average data-MC difference observed in photon-enriched samples
    - large systematic uncertainties (up to ~15%)
- 2011 dataset: ~ 5 fb^{-1}
  - start to have enough radiative Z decays (for E_T^γ<50 GeV)
  - analyses not statistically limited ⇒ need to reduce systematic uncertainties to few % level

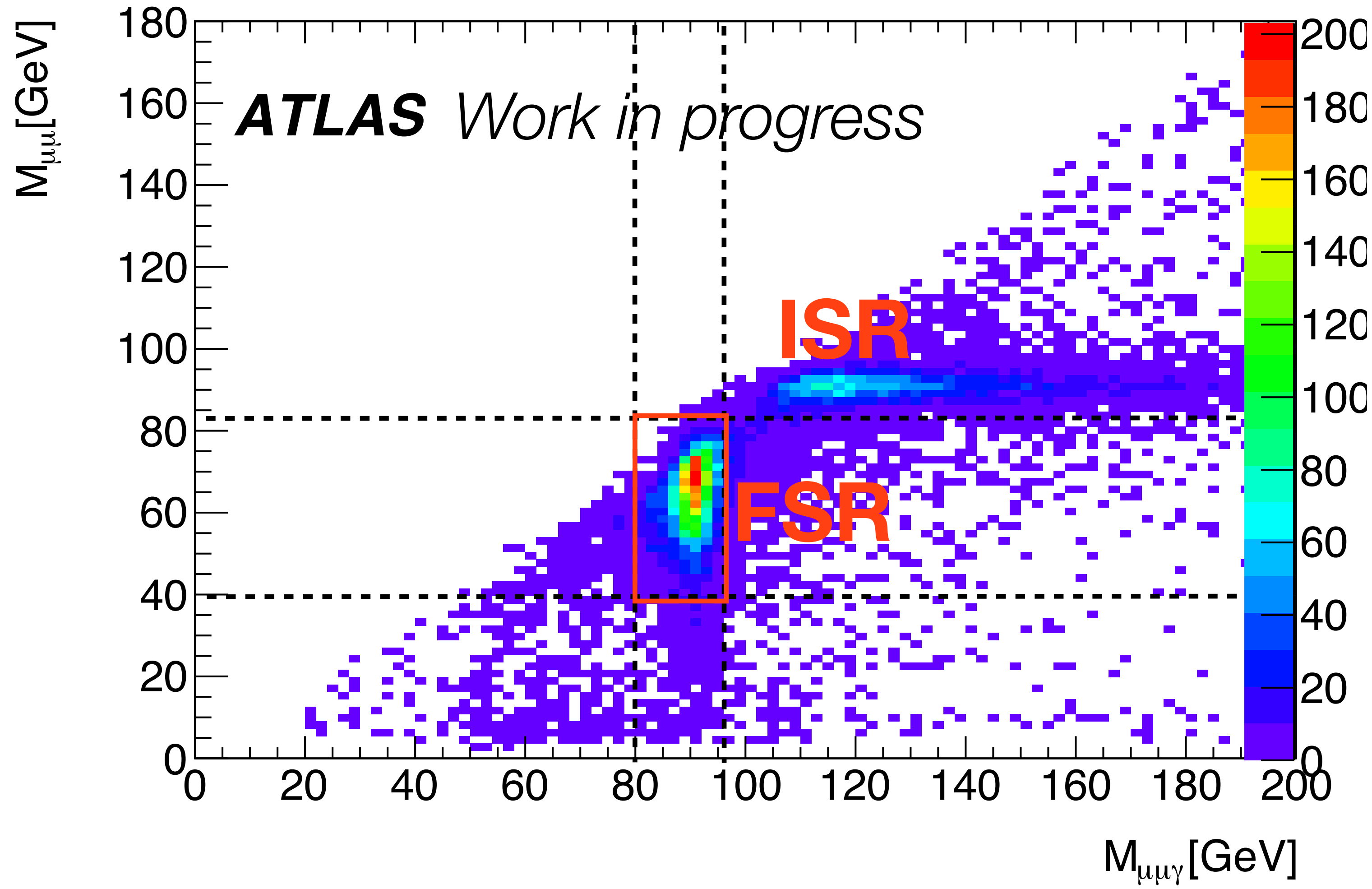


Data-driven measurements of the photon ID efficiency in ATLAS

- Radiative Z decays
  - select $Z \rightarrow ll\gamma$ sample using requirements on leptons and kinematics of $ll\gamma$
  - gamma candidate = clean probe to measure efficiency
- Matrix method
  - select a photon-enriched sample with events passing photon triggers (very loose selection)
  - estimate residual bkg contamination in selected sample, both before and after application of ID criteria, using a discriminant variable for which the different signal and background efficiencies are known
- Electron extrapolation
  - select pure sample of electrons from $Z \rightarrow ee$ (with T&P)
  - “transform” electron shower shapes into photon shower shapes

Photon efficiency from radiative Z decays

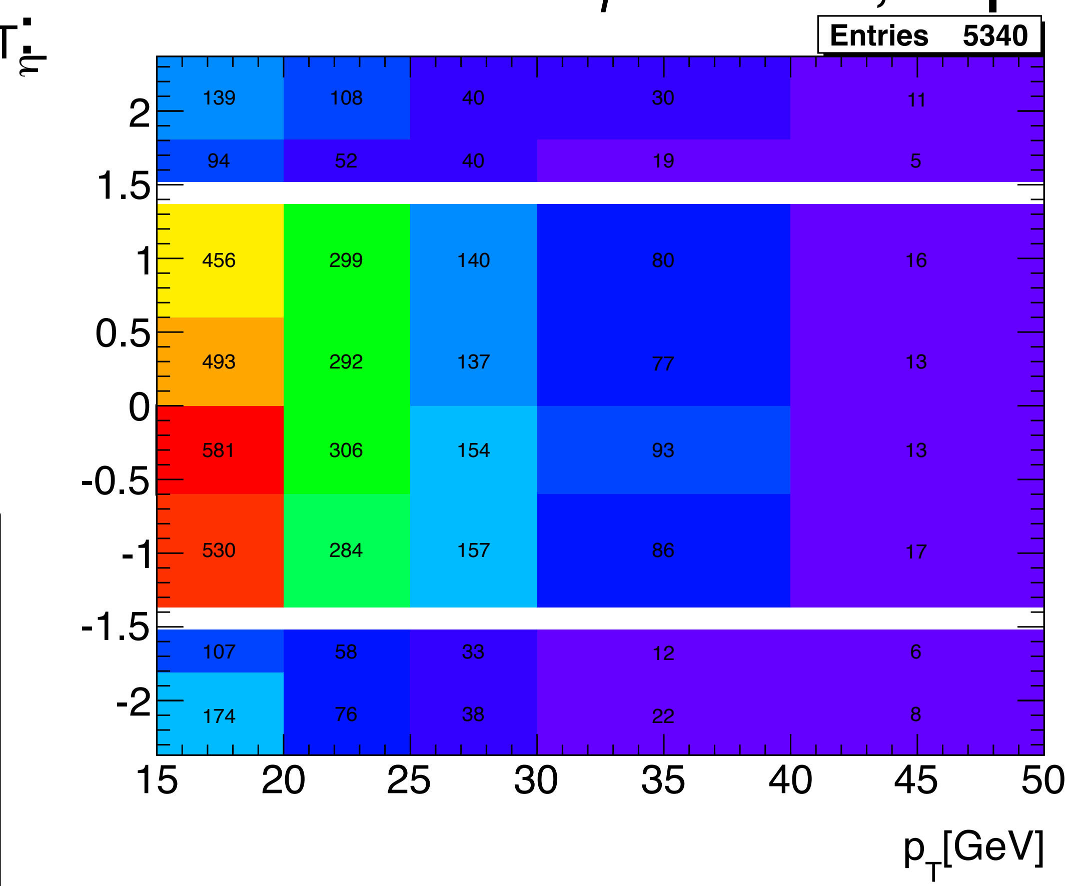
- single electron or muon triggers
- good electron or muon (in tracker, EM calorimeter/muon spectrometer)
    - originating from primary vertex
    - isolated (reject leptons from heavy flavor decays)
    - not too close to the photon (avoid bias on photon shower shapes)
- $E_T^{\gamma,l}$>15 GeV
- 80 < $M_{ll\gamma}$ < 96 GeV
- 40 < M_{ll} < 83 GeV
- ~11k probes in total



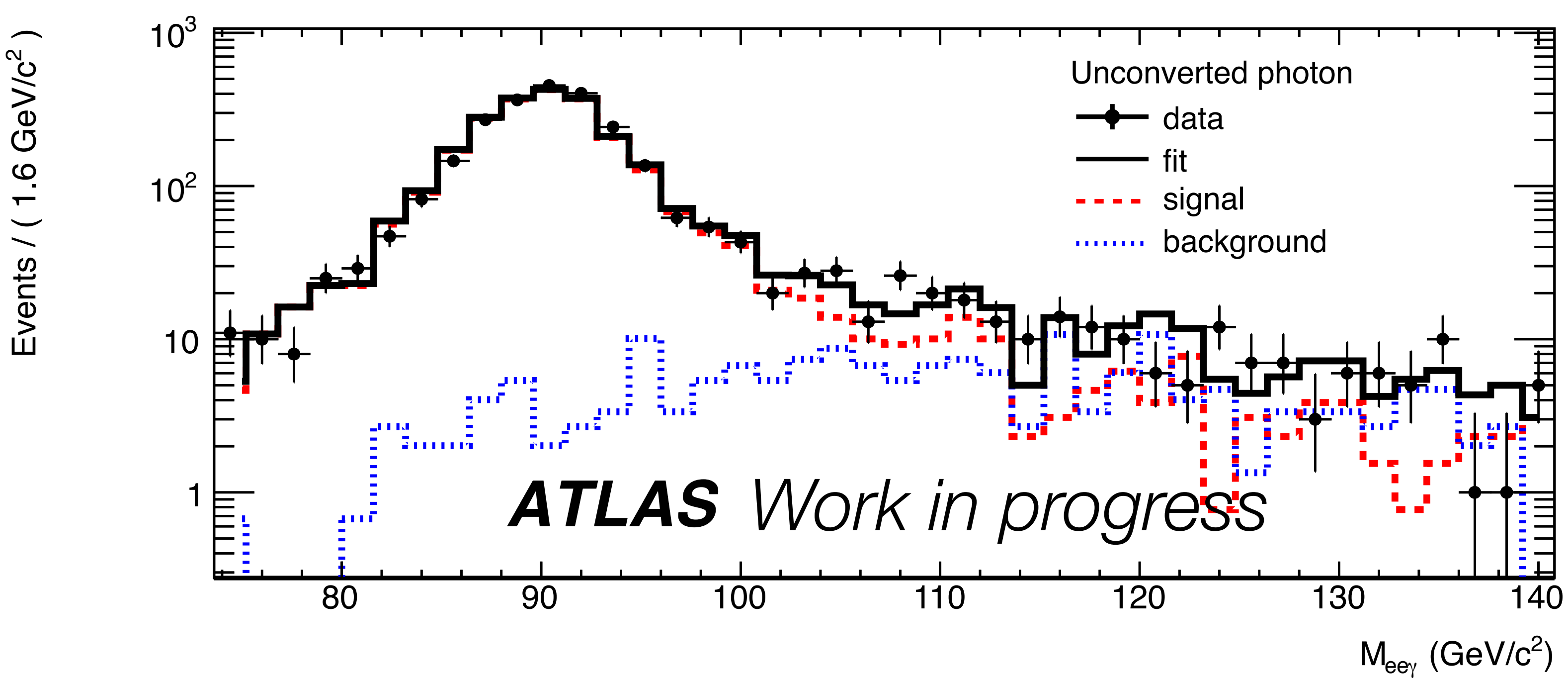
Photon efficiency from radiative Z decays

- Typical probe E_T/eta distribution peaks at low E_T:

unconverted photons, l=μ



- Residual bkg from Z+jets; estimated from template fit to $M_{ll\gamma}$ using MC templates



- purity typically > 98%; error on efficiency when assuming P=1: 1-1.5%

- Results consistent for l=e,μ ⇒ combined. Statistically limited (plots later)

Photon efficiency from matrix method

- After selecting photon-enriched sample (photon triggers: very loose cuts on few shower shapes), use track isolation as discriminating variable to count signal (S) and background (B) that pass/fail the identification (T) criteria

- System of 4 equations in 4 unknowns (if track iso efficiencies are known)

$$N^T_{pass} = N^S_{pass} + N^B_{pass}$$
$$N^T_{fail} = N^S_{fail} + N^B_{fail}$$

pass track isolation cut ⇒

$$N^{Iso}_{pass} = \varepsilon^s_p * N^S_{pass} + \varepsilon^b_p * N^B_{pass}$$
$$N^{Iso}_{fail} = \varepsilon^s_f * N^S_{fail} + \varepsilon^b_f * N^B_{fail}$$

$$N^S_{pass} = \frac{N^{Iso}_{pass} - \varepsilon^b_p * N^T_{pass}}{\varepsilon^s_p - \varepsilon^b_p}$$

$$N^S_{fail} = \frac{N^{Iso}_{fail} - \varepsilon^b_f * N^T_{fail}}{\varepsilon^s_f - \varepsilon^b_f}$$

⇒

$$\varepsilon^{tight-ID} = \frac{N^S_{pass}}{N^S_{pass} + N^S_{fail}}$$

$$P = \frac{N^S_{pass}}{N^T_{pass}} = \frac{\varepsilon_p - \varepsilon^b_p}{\varepsilon^s_p - \varepsilon^b_p}$$

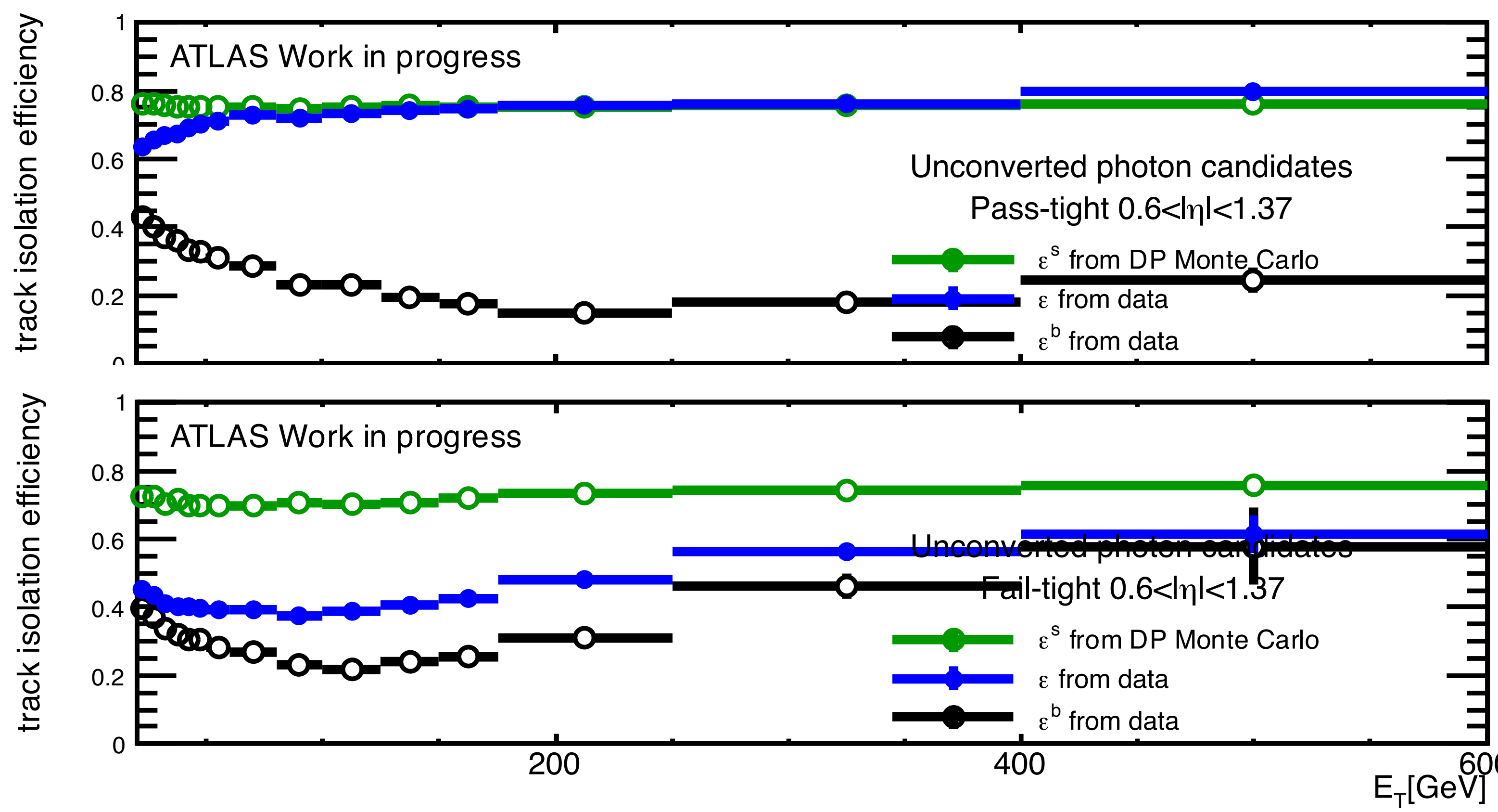
$$F = \frac{N^S_{fail}}{N^T_{fail}} = \frac{\varepsilon_f - \varepsilon^b_f}{\varepsilon^s_f - \varepsilon^b_f}$$

- Prompt photon track iso efficiencies from signal MC (systematic uncertainty from data/MC efficiency difference for electrons from $Z \rightarrow ee$)

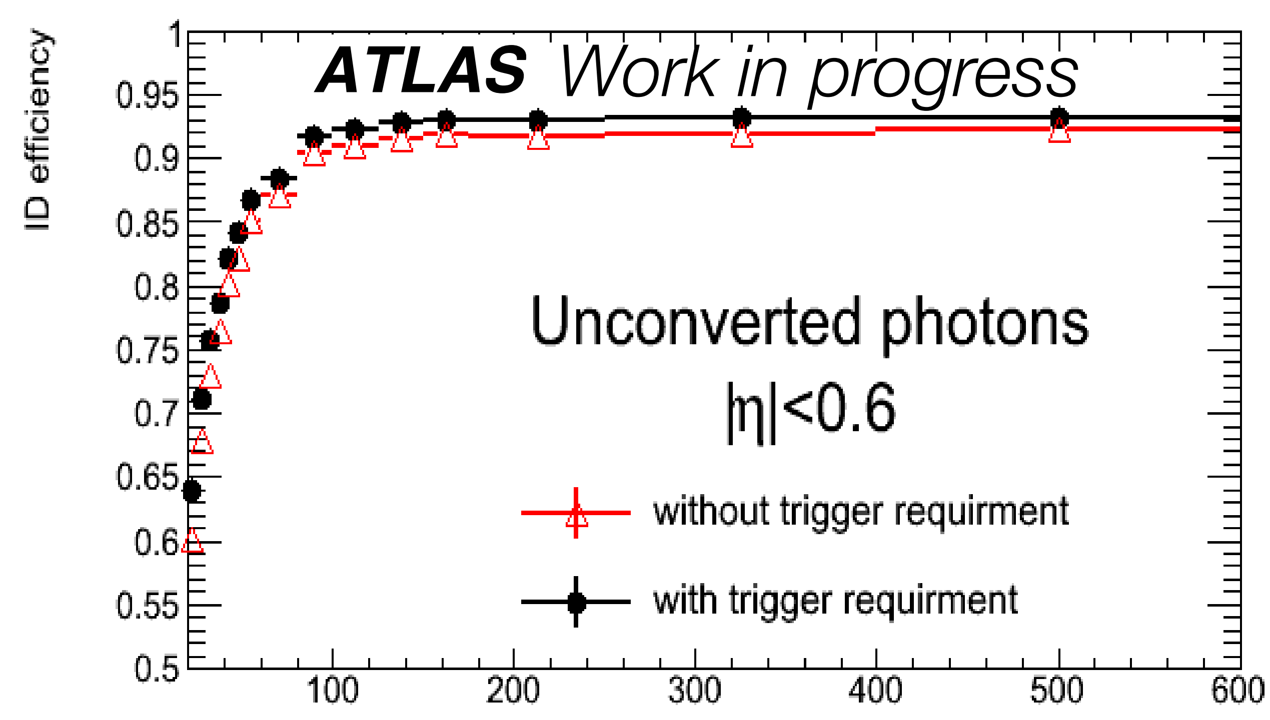
- Background track iso efficiencies from data control sample enriched with jets (systematic uncertainty from closure tests on MC)

Ingredients for efficiency measurement with matrix method

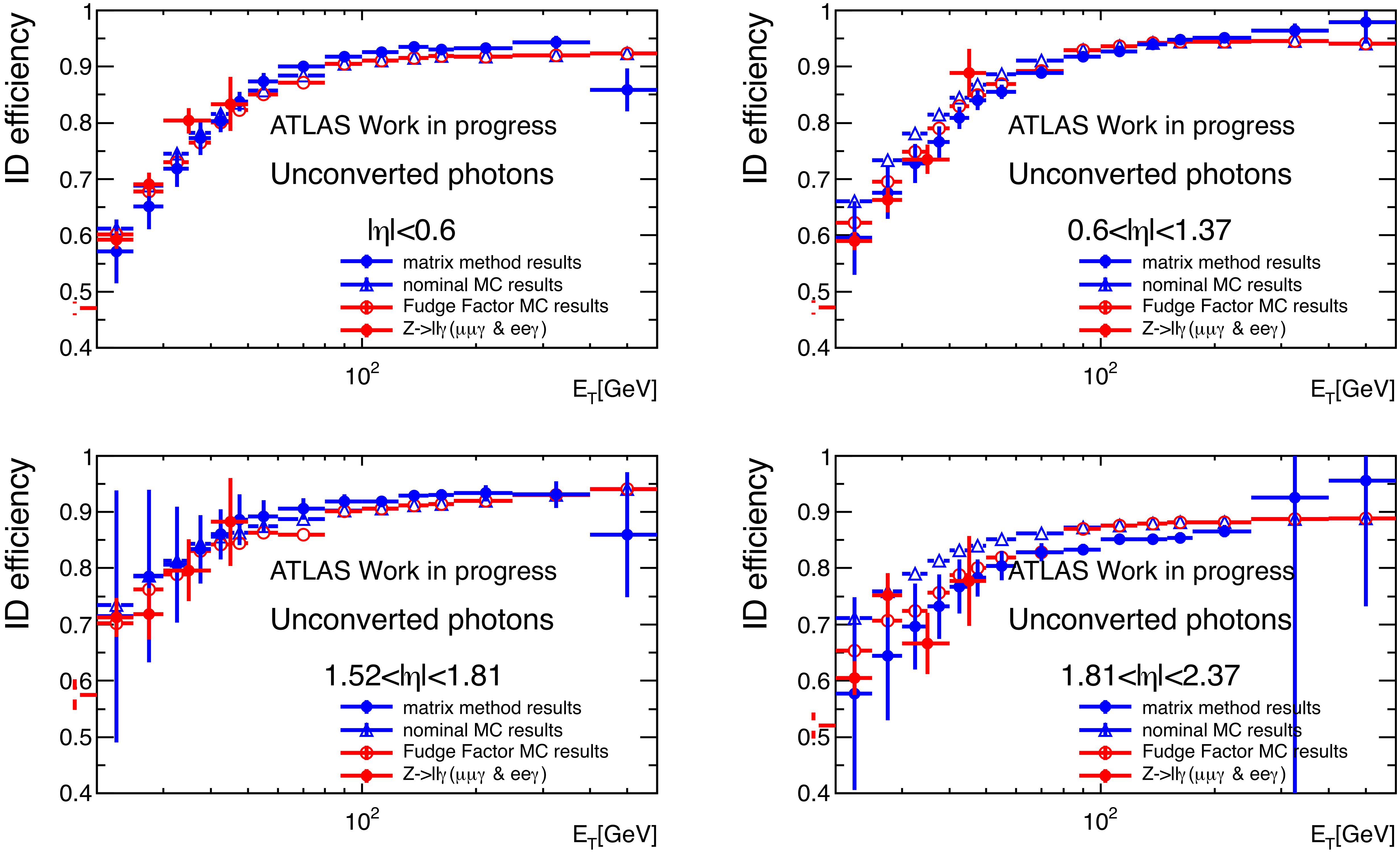
- Track isolation efficiencies



- Efficiency of trigger preselection (from MC; to be checked with radiative Z decays at low ET)

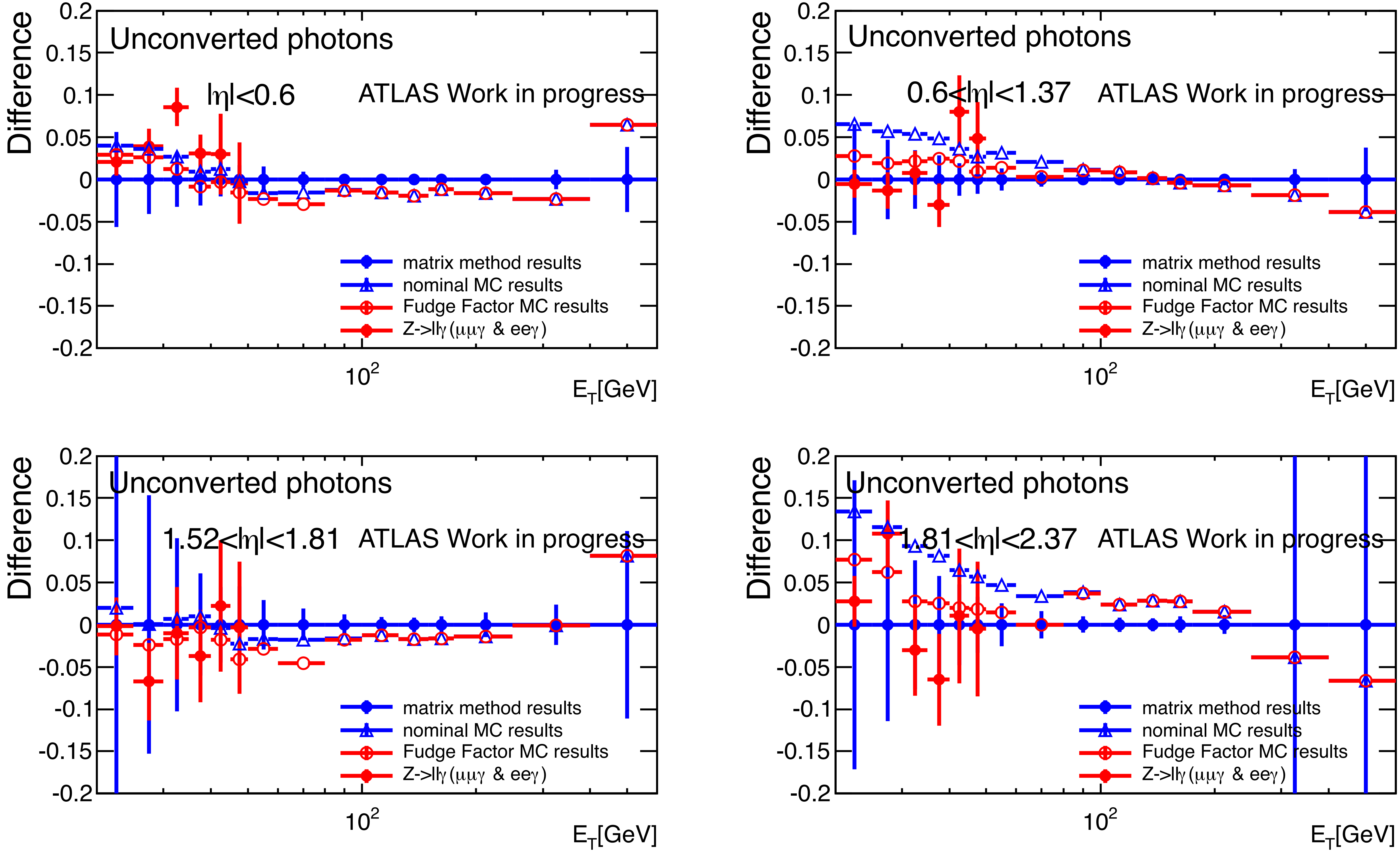


Results



ID efficiency
ATLAS Work in progress
Unconverted photons
|η|<0.6
matrix method results
nominal MC results
Fudge Factor MC results
Z->llγ (μμγ & eeγ)
10^2
E_T[GeV]
ID efficiency
ATLAS Work in progress
Unconverted photons
0.6<|η|<1.37
matrix method results
nominal MC results
Fudge Factor MC results
Z->llγ (μμγ & eeγ)
10^2
E_T[GeV]
ID efficiency
ATLAS Work in progress
Unconverted photons
1.52<|η|<1.81
matrix method results
nominal MC results
Fudge Factor MC results
Z->llγ (μμγ & eeγ)
10^2
E_T[GeV]
ID efficiency
ATLAS Work in progress
Unconverted photons
1.81<|η|<2.37
matrix method results
nominal MC results
Fudge Factor MC results
Z->llγ (μμγ & eeγ)
10^2
E_T[GeV]

Results


Unconverted photons
|η|<0.6
ATLAS Work in progress
0.6<|η|<1.37
1.52<|η|<1.81
1.81<|η|<2.37
Difference
E_T[GeV]
10^2
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
matrix method results
nominal MC results
Fudge Factor MC results
Z->llγ (μμγ & eeγ)

Conclusion

- Prompt photon identification efficiency measured with 2 data-driven methods
  - radiative Z decays ⇒ precise at low E_T (limited at high E_T by statistics)
  - matrix method ⇒ precise at high E_T (limited at low E_T by purity uncertainty)
- Results from 2 methods consistent within few %
  - can reduce systematic uncertainties significantly wrt 2010 analyses
- MC estimates typically within 5% from data results
  - confirms validity of 2010 results based on MC efficiencies (with conservatively large systematic uncertainties)

Backup

ε_p^s and ε_f^s from DP simulated sample

- Central values from truth reconstructed isolated photons

  * DP 17 (Et > 20 GeV)
  * DP 35 (Et > 45 GeV)
  * DP 70 (Et > 90 GeV)
  * DP 140 (Et > 150GeV)
  * DP 280 (Et > 290GeV)
  * DP 500 (Et > 510GeV)

- Systematic uncertainties:

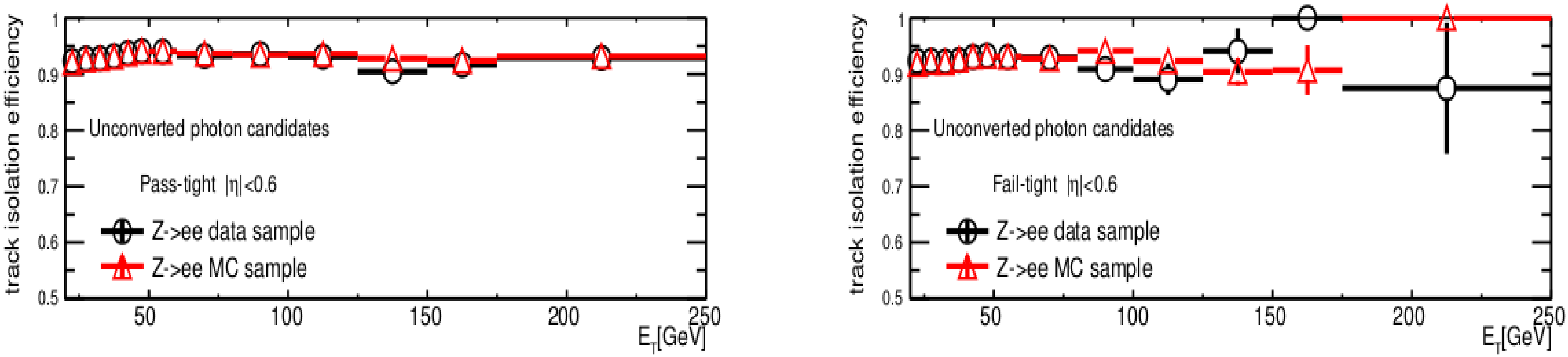
  * Possible data/MC discrepancies:
    * from the difference between ε^s in Z->ee data and MC
    * plots are shown in next slide

  * difference between ε^s for 1 / 2 track conversion
    * only for converted photons
    * from difference between ε^s for unconverted and converted photons

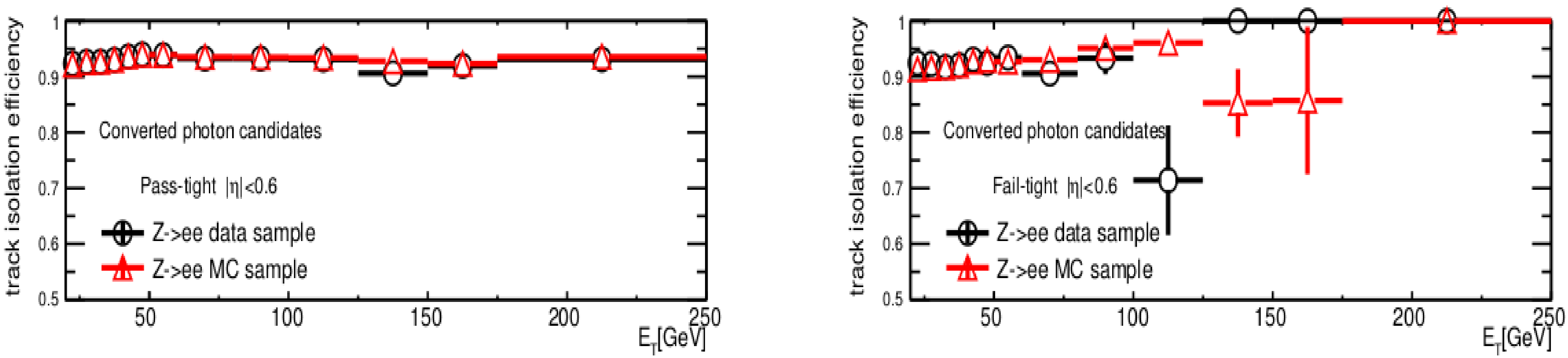
ε_p^s and ε_f^s from DP simulated sample

- Track isolation efficiency from probe electrons in Z->ee data and MC passing or not the unconverted/converted photon identification

unconverted photon

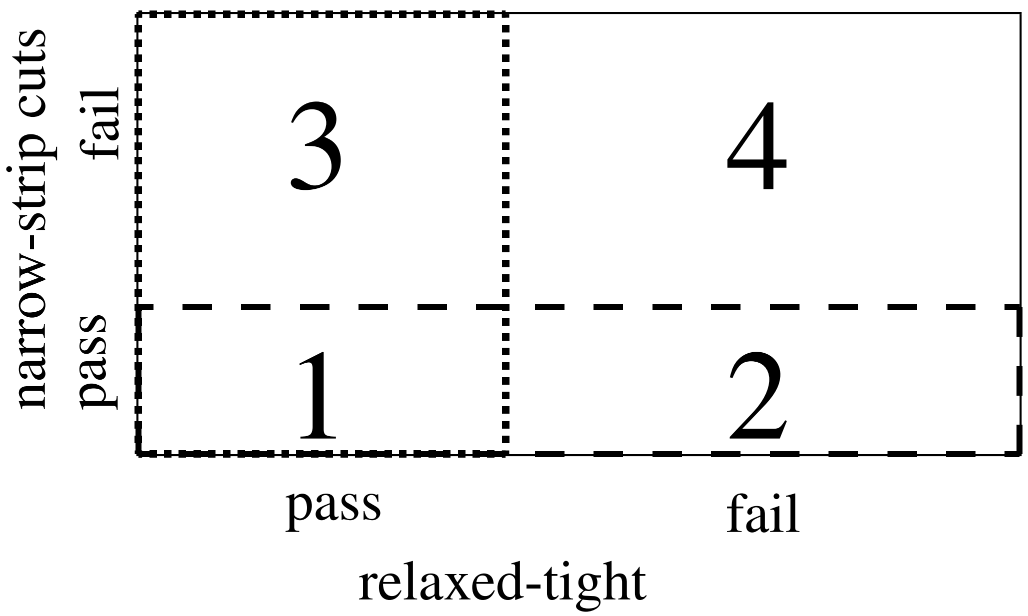


converted photon



ε_p^b and ε_f^b from bkg-enriched data sample

- Select photon candidates in data by reversing narrow strips variables cuts
  * four variables: ΔE , fside , Ws3 and Eratio
  * relaxed-tight cut: ID tight cuts except above narrow-strips variables cuts



pass relaxed-tight cuts

pass all cuts on narrow-strip variables

1: pass tight cuts

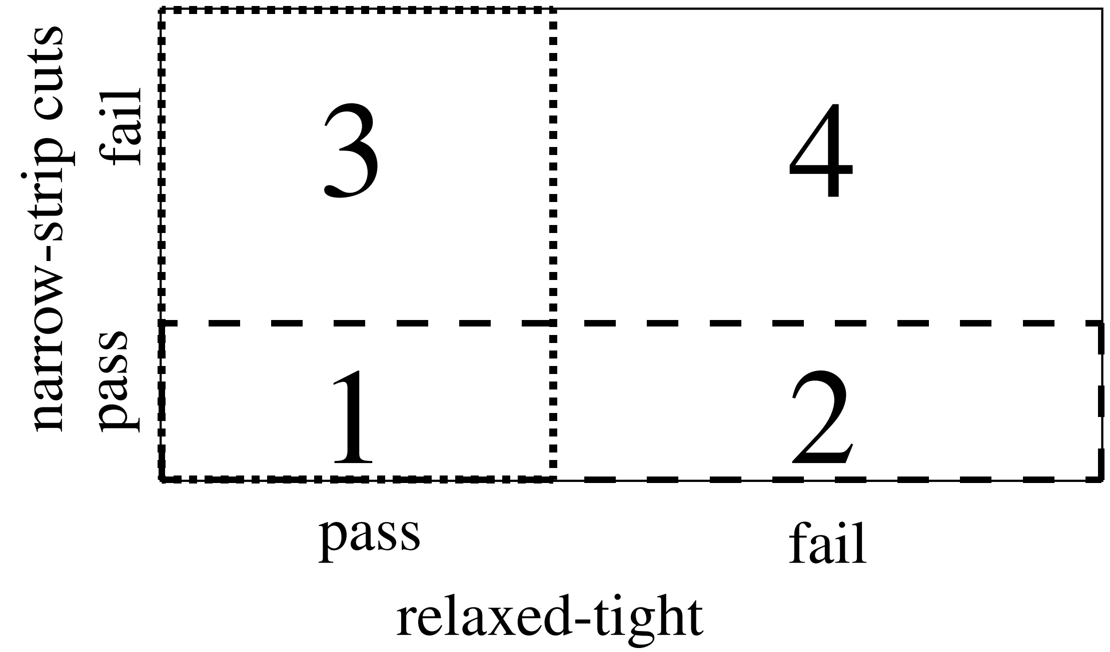
2: pass narrow-strip variable cuts but fail relaxed-tight cuts

3: pass relaxed-tight but fail narrow-strip cuts

4: fail relaxed-tight and fail narrow-strip cutsvariables

  * signal leakage correction (see next slide)

- Systematic uncertainty:
  * bias on ε^b checked with simulated JF MC samples: comparing the difference of ε^b obtained with truth match and with the same procedure used in data
    JF17(Et >20GeV) + JF35(Et>45GeV) + JF70(Et>90GeV) + JF140(Et>150GeV) + JF240(Et>250GeV) + JF500(Et>510GeV)



We introduce a few definitions according to Fig. 3:

- N_A: total number of photon candidates in region 3.
- N_B: total number of photon candidates in region 4.
- ε_p: fraction of photon candidates that are isolated in the tracker after passing tight criteria (region 1).
- ε_f: fraction of photon candidates that are isolated in the tracker after failing tight criteria (region 2+3+4).
- $\varepsilon_p^{b^+}$: fraction of photon candidates in region 3 that are isolated in the tracker.
- $\varepsilon_f^{b^+}$: fraction of photon candidates in region 4 that are isolated in the tracker.

We also remind the previous definitions of:

- N^T_{pass}: total number of photon candidates that pass tight criteria (region 1).
- N^T_{fail}: total number of photon candidates that fail tight criteria (region 2+3+4).

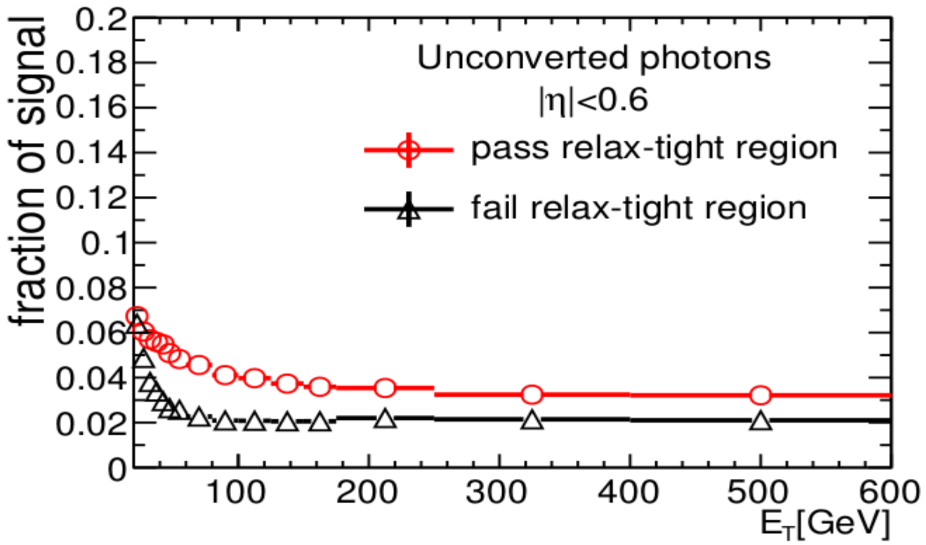
The previous quantities are determined on data.

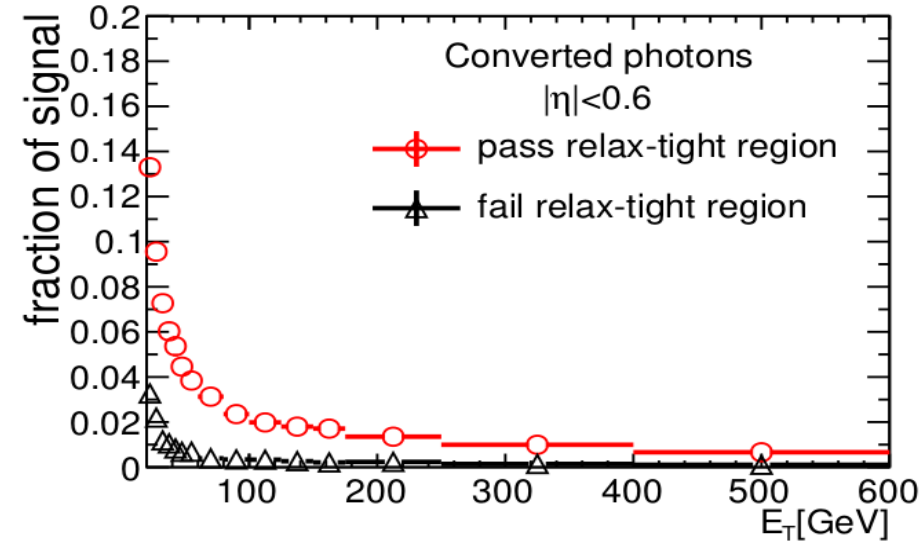
From the prompt photon MC sample we extract the following quantities for true prompt photons:

- f_p: fraction of prompt photons that leak into region 3.
- f_f: fraction of prompt photons that leak into region 4.
- ε_p^s: track isolation efficiency for prompt photons in region 1.
- ε_f^s: track isolation efficiency for prompt photons in region 2+3+4.
- $\varepsilon_p^{s^+}$: track isolation efficiency for prompt photons in region 3.
- $\varepsilon_f^{s^+}$: track isolation efficiency for prompt photons in region 4.

Signal leakage correction

- Fraction of true photons in bkg control sample passing or failing relaxed-tight criteria





- Signal leakage correction procedure: solve three equations to obtain 3 unknowns total real photons in passing and failing tight criteria

$$N_s^{total} = \frac{\varepsilon_p - \varepsilon_p^b}{\varepsilon_p^s - \varepsilon_p^b} * N_P + \frac{\varepsilon_f - \varepsilon_f^b}{\varepsilon_f^s - \varepsilon_f^b} * N_F \quad ....(1)$$

subtract signal leakage from numerator and denominator

$$\varepsilon_p^b = \frac{N_A * \varepsilon_p^{b^+} - N_s^{total} * f_p * \varepsilon_p^{s^+}}{N_A - N_s^{total} * f_p} \quad ....(2)$$

$$\varepsilon_f^b = \frac{N_B * \varepsilon_f^{b^+} - N_s^{total} * f_f * \varepsilon_f^{s^+}}{N_B - N_s^{total} * f_f} \quad ....(3)$$

equations are nonlinear, iterative procedure is used with inputs:

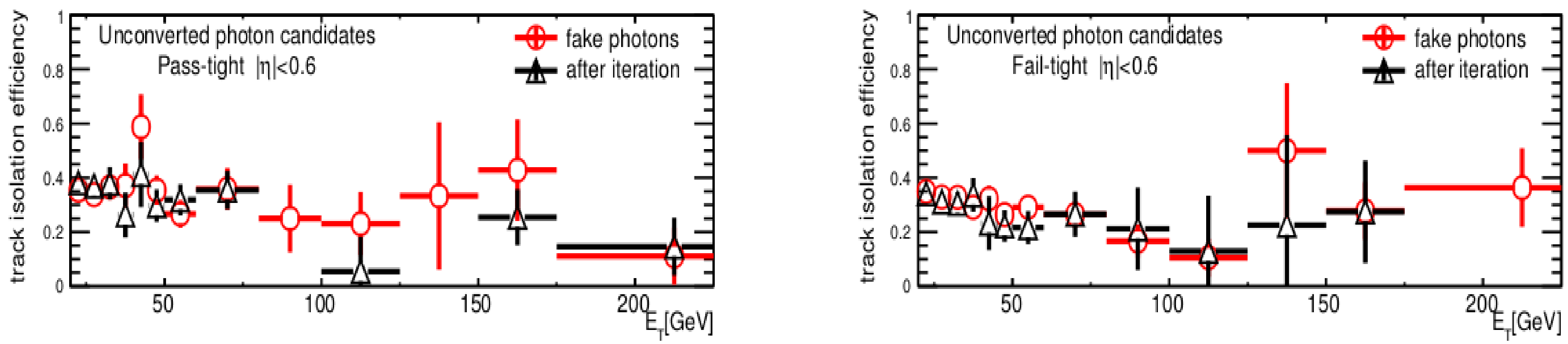
* from DP MC: $\varepsilon_p^{s^+}$ $\varepsilon_f^{s^+}$ f_p f_f

* from JF MC at the beginning: $\varepsilon_p^{b^+}$ $\varepsilon_f^{b^+}$ more and more precise values are solved during iteration (3~5 steps)

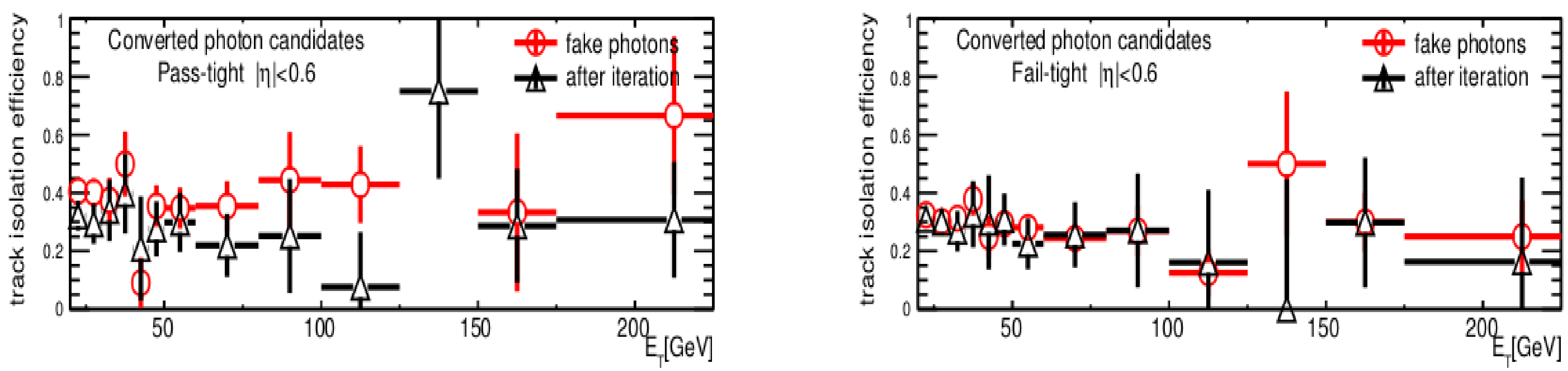
Bias on ε_p^b and ε_f^b in JF Monte Carlo

- In JF MC, compare the difference of ε_p^b and ε_f^b obtained with the truth match anc the procedure used in data(bkg control sample and subtract signal leakage by using iteration)

unconverted photons



converted photons



- These difference is independent with E_T within statistical error. The average difference is taken as systematic uncertainty for each |η| region passing or not tight criteria..

9