D0 Note 4672

November, 2006

(version 9)

The Isolated Photon Cross Section in the Central Rapidity Region at 1.96 TeV.

Dmitry Bandurin[1], Michael Begel[2]

[1] *Joint Institute for Nuclear Research, Dubna, Russia*

[2] *University of Rochester, Rochester, USA*

Abstract

A measurement of the inclusive cross section for production of isolated photons is presented for transverse energies in the range of 23−300 GeV for the central region pseudorapidity of $|\eta| < 0.9$. The results are based on a data sample of 380 pb^{-1} of integrated luminosity accumulated during 2002−2004 in $p\bar{p}$ collisions at $\sqrt{s}$ =1.96 TeV and recorded with the DØ detector at the Fermilab Tevatron Collider. The obtained results are compared with the next-to-leading order QCD predictions using CTEQ6.1M and MRST2004 parton distribution functions.

1 Introduction.

The production of isolated photons in high energy hadronic collisions has been studied intensively at experiments [1, 2, 3, 6, 7, 8, 9, 4, 5] and theoretically (see for example [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]) during the last 20 years.

In high energy $p\bar{p}$ collisions the dominant source for production of photons with moderate and high transverse momentum p_T is direct (or prompt) photons. They are called direct since they are produced directly from parton-parton interactions and not from the hadron decays (such as π^0, η, K_s^0). These photons come unaltered from the hard process and therefore can give us a clean test of the hard scattering dynamics. In the region up to $p_T^\gamma \simeq 150$ GeV the direct photons are mainly produced through the Compton scattering $q+g \rightarrow q+\gamma$ [1)] and thus their production cross section is sensitive to the gluon density inside the colliding hadrons. A fractional contribution of different QCD subprocesses to the production of isolated prompt photons [2)] is shown in Fig. 1. A high center of mass energy at Tevatron and the statistics accumulated currently in Run II allows us to test QCD and gluon distribution in the region of large Q^2 (1 order of magnitude higher than reached at HERA) and wide range of x_T: $0.02 < x_T < 0.30$. This process is complimentary to deep inelastic scattering and to '$W/Z + jet$' production. The measurements of isolated photon cross section also allows testing the next-to-leading order (NLO) and resumed QCD calculations, phenomenological models of gluon radiation, studies of photon isolation and the fragmentation process.

Photon identification is free from the uncertainties caused by the parton fragmentation or by experimental issues related to jet identification and energy measurement and thus has an advantage over jet production measurement.

In addition, photons in the final state may be an important sign of new particles and/or physics beyond the standard model. Thus, first of all, it is useful and necessary to study and to understand the "conventional" sources of photons.

This note presents a first measurement of the cross section for production of isolated photons in $p\bar{p}$ collisions at $\sqrt{s}$ =1.96 TeV (Run II) in the central pseudorapidity region of $|\eta| < 0.9$ [3)] and covers a much wider p_T range than Run I measurements at CDF [3] and DØ [6].

[1)] while for higher p_T^γ the annihilation process $q+\bar{q} \rightarrow g+\gamma$ becomes dominating (see e.g. [16]). Recent CDF data [21], calculation with DIPHOX package compared with preliminary Run I D0 data [22] as well as estimation done with PYTHIA event generator show that relative contribution from diphoton production is small enough ($\lesssim 1\%$).

[2)] with fraction of hadronic energy $\epsilon_h \leq 10\%$ in the cone of $R = 0.4$.

[3)] Pseudorapidity is defined as $\eta = -ln\ tan\theta/2$, where θ is the polar angle with respect to the proton beam.

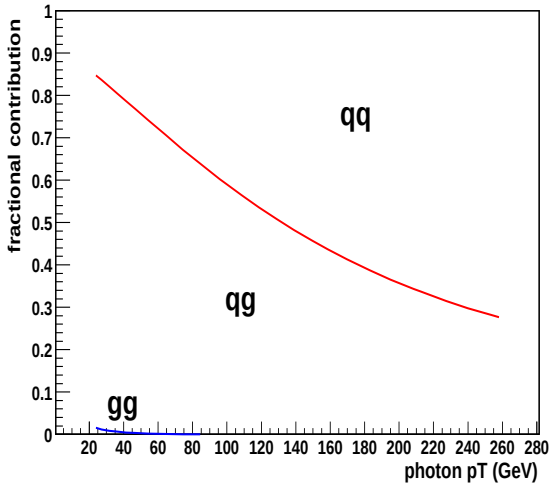

Figure 1: A fractional contribution of QCD subprocesses to the production of isolated prompt photons.

2 Data sample.

For the analysis we use data preselected by Common Sample group according to the '1 EM' skim definition with the following criteria [24]:

- EM object $|ID| = 10, 11$;
- $p_T > 15$ GeV.

The analyzed data correspond to the time period from August 2002 to June 2004 and combine the runs 161973–194566 in global trigger versions of v8–12.

Runs in the data sample are required to be declared as GOOD or REASONABLE by all quality groups excluding just CTT and MUON groups.

The data have been selected by a combination of the unprescaled (for a given run) triggers [4)] from the list:
EM_HI_SH || EM_HI || EM_MX_SH || EM_MX in trigger versions v8–11 and
E1_SHT20 || E1_SH30 || E1_L50 in trigger version v12.

Just events from good luminosity blocks are taken, (i.e. all the non-normalizable luminosity blocks were excluded from the analysis). In addition, luminosity blocks which were troubled by transient problems in calorimeter and identified as bad by JET/MET groups are removed. After applying all the criteria described above, the integrated luminosity of the dataset is 380 ± 22.8 pb^{-1} [25].

[4)] A more detailed description can be found in section 4.2.1

3 Selection criteria.

Photon candidates were identified in the DØ detector [26] as isolated clusters of energy depositions in the uranium and liquid-argon sampling calorimeter. The electromagnetic (EM) section of the calorimeter is segmented longitudinally into four layers (EM1-EM4) of 2, 2, 7, and 10 radiation lengths, respectively, and transversely into cells in pseudorapidity and azimuthal angle $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ (0.05×0.05 in EM3). In addition, the cluster may also contain the energy deposited in the hadronic portion of the calorimeter located after the EM one [26].

To select photon candidates in data and Monte Carlo (MC) we have used the following selections:

$\diamondsuit$ Common Sample group preselection cuts (section 2).

$\diamondsuit$ Primary vertex (PV) selection: The event vertex was required to be within 50 cm of the nominal center of the detector along the beam ($|Z_{vtx}| < 50\ cm$) and should have at least 3 associated tracks.

Main photonic criteria:

$\diamondsuit$ EM object is reconstructed by simple cone algorithm [24].

$\diamondsuit$ Just central photons are considered with $|\eta_{detector}| < 0.9$.

$\diamondsuit$ To avoid inter-calorimeter boundaries and cracks EM fiducial cuts are applied [24]

$\diamondsuit$ Each candidate was required to deposit more than 95% of the detected energy in the EM section of the calorimeter (EMfrac > 0.95)

$\diamondsuit$ and to be isolated in the annular region between $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.2$ and $R = 0.4$ around the gravity center of the cluster: $Iso(\Delta R02) < 0.10$. Here $Iso(\Delta R02) = (EisoTot - EisoCore)/EisoCore$, where $EisoTot$ is overall (EM+hadronic) tower energy in (η, ϕ) circle of $R = 0.4$ and $EisoCore$ is EM tower energy in circle of $R = 0.2$.

$\diamondsuit$ Probability to have any track spatially matched to the EM cluster in the event was required to be below 0.001.

$\diamondsuit$ We also reject events having too large missing E_T by the cut $E_T^{miss}/p_T^{\gamma} < 0.7$.

Additional selection variables/criteria.

Many other selection criteria were studied and tested for photon identification [29, 30]. The following variables turned out to be most efficient and consistent from the point of view of testing them on the $Z \to ee$ events:

$\diamondsuit$ $ncell\_EM1\ (EMclust)$ — number of cells that belong to the EM cluster, are in EM1, and have a cell energy $E_{cell} > 0.4$ GeV.

$\diamondsuit$ $ncell\_EM1\ (\Delta R02)$ — number of cells with the above criteria and within the ring of $\Delta R = 0.2 - 0.4$.

$\diamondsuit$ $sum\_trackPT\ (\Delta R04)$ — scalar sum of track transverse momenta (p_T^{tr}) in the ring of $0.05 \leq R \leq 0.4$. Here we consider only the tracks that are produced within 2 cm from the primary

vertex of the event and have $p_T^{tr} > 0.3$ GeV.
$\diamondsuit$ $sigrphi\_EM3$ — energy weighted EM cluster width in $r \times \phi$ in EM3 layer [24].

4 Calculation of cross section.

The inclusive photon cross section was obtained by the relation:

$$\frac{d^2\sigma}{dp_T^\gamma \, d\eta^\gamma} = \frac{N \; P \; f_{unsm}}{L_{int} \; \Delta p_T^\gamma \; \Delta\eta^\gamma \; A \; \epsilon_t \; \epsilon_s} \tag{1}$$

where
N is number of photon candidates in the selected sample;
P is photon purity (or fraction of single photons in the sample);
f_{unsm} is unsmearing correction factor caused by finite energetic resolution for EM objects;
L is the integrated luminosity;
Δp_T^γ and $\Delta\eta^\gamma$ are the bin sizes in transverse momentum and pseudorapidity;
A is the geometric acceptance;
ϵ_t is the trigger efficiency;
ϵ_s is the efficiency of selection criteria.

4.1 Electromagnetic jets.

Unfortunately for Monte Carlo studies, only about 5 out of 10,000 QCD jets survive loose (and basic) photon ID cuts (such as EMfrac > 0.90, $Iso(\Delta R02) < 0.15$ and "No track matched"). To apply tighter cuts and to have still a statistics sufficient for the analysis (e.g. for estimation of background efficiency and purity) we need $n \cdot 10^6$ QCD events. This is, of course, too time and CPU consuming. The solution is, as in Run I, a making preselection on the generator level and with a following reconstruction of just electromagnetic jets (see [6]).

Some special cuts need to be defined for this aim. The study was performed (in earlier release-versions: p10.14.xx, p13.06.xx which had sufficient statistics of QCD MC events) on the QCD jets which were identified as EM objects after passing basic EM ID criteria: EMfrac> 0.90, $Iso(\Delta R02) \leq 0.15$. Looking at the particles at the generator level some characteristic variables were defined. These variables include

• maximal p_T^γ in $R \leq 0.2$ – p_T of the most-energetic photon/electron in the ring of $R = 0.2$ (R02) around center of gravity of EM cluster;

• $sPT(R < 0.2)/\hat{p}_\perp^{min}$ – p_T scalar sum of all particles in R02 divided by maximal p_T of a final state parton in the hard interaction;

• $sPT(R \leq 0.2)_{em}/sPT(R \leq 0.2)$ – as above but just e/m particles are considered in R02;

• $sPT(R \leq 0.2)_{had}/sPT(R \leq 0.2)$ – as above but just hadrons are considered in R02;

• $sPT(\Delta R02)/sPT(R \leq 0.2)$ – p_T scalar sum of all particles in the ring of $\Delta R = 0.2$ around EM cluster divided by $sPT(R \leq 0.2)$;

$$\hat{p}_\perp^{\,min} = 40 \text{ GeV}$$

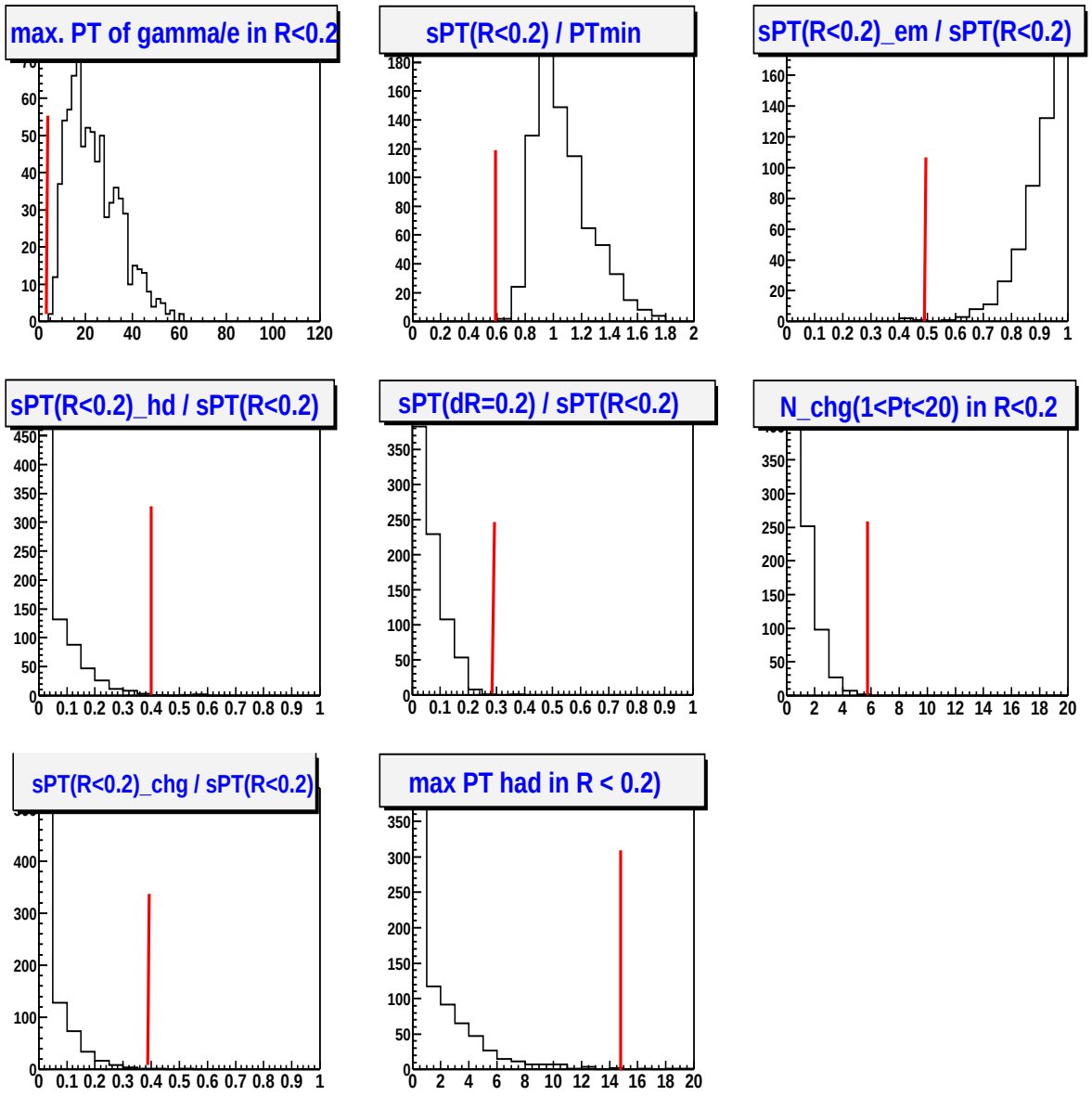

Figure 2: Distributions over characteristic variables and position of cuts to select em-jets are shown. A case of the full MC generation of 'γ^{dir}+jet' events with $\hat{p}_\perp^{\,min} = 40$ GeV is considered. The shown cut position is slightly depends on the minimal parton transverse momentum $\hat{p}_\perp^{\,min}$.

• $N_{ch}(R \leq 0.2)$ – number of charged particles with $p_T > 1$ GeV in R02;

• maximal hadron p_T in $R02$ – p_T of the most energetic hadron in R02.

Distributions of those characteristic variables and position of cuts to select em-jets are shown in Fig. 2. This is shown for the full MC generation of 'γ^{dir}+jet' events with $\hat{p}_{\perp}^{min} = 40$ GeV, and the cut position only slightly depends on the minimal parton transverse momentum $\hat{p}_{\perp}^{min}$ [5]. As we see, those cuts are weak enough and leave of order of 99% of the events selected after application of the basic reco-level EM ID cuts (see above). It is worth emphasizing that since in a photon analysis we are going to apply tighter reco-level photon cuts (that the basic ones) we have a safety factor that we did not miss any QCD events.

After application of the found cuts the QCD (in fact, EM) jets in the cone of $R = 0.7$ were formed, written to a file and then fully simulated by the packages DØGstar [32], DØSim, and DØReco.

About 10^5 em-jet events were generated for 17 p_T intervals from 23 to 300 GeV to study background rejections and photon purities after the used cuts.

We have classified (and simulated) three separate types of em-jets: π^0-, η- and $(K_s^0 + \omega)$-em-jets. They are classified according to the particle that induce the em-jet (i.e. the most energetic particle in the core of em-jet with at least one photon in the final state) [29].

4.2 Calculation of efficiencies.

4.2.1 Trigger efficiency.

All of the events in this analysis must have fired one of the unprescaled single high p_T EM triggers. For global CMT trigger versions from 8 to 11 (runs $\leq$ 178721) the possible triggers are EM_HI_SH || EM_HI || EM_MX_SH || EM_MX. For global CMT trigger version 12 (runs $>$ 178721) it is E1_SHT20 || E1_SH30 || E1_L50.

Trigger EM_HI_SH (EM_HI) at level 1 requires one calorimeter EM trigger tower with $p_T >$10 GeV, at level 2 requires one EM candidate with $p_T >$12 GeV and at level 3 selects EM clusters with $p_T >$20 (30) GeV satisfying transverse shower shape requirements. Trigger EM_MX_SH (EM_MX) differs in that the level 1 requirement is one calorimeter EM trigger tower with $p_T >$15 GeV.

Trigger E1_SHT20 (E1_SH30) at level 1 requires one calorimeter EM trigger tower with $p_T >$11 GeV, has no limitations at level 2 and at level 3 selects EM clusters with $p_T >$20 (30) GeV satisfying tight (loose) shower shape requirements. E1_L50 differs from E1_SH30 by selection of EM clusters with $p_T >$50 at level 3.

Since we prefer to use the data to calculate the efficiencies of the trigger combinations (and since there is no a source of pure photons in data) and we believe that the calorimeter based trigger

[5] One needs to note here that those cuts are highly correlated, i.e. in considering QCD jets a limitation of one of those variables would strongly affect the distribution over the other listed variables.

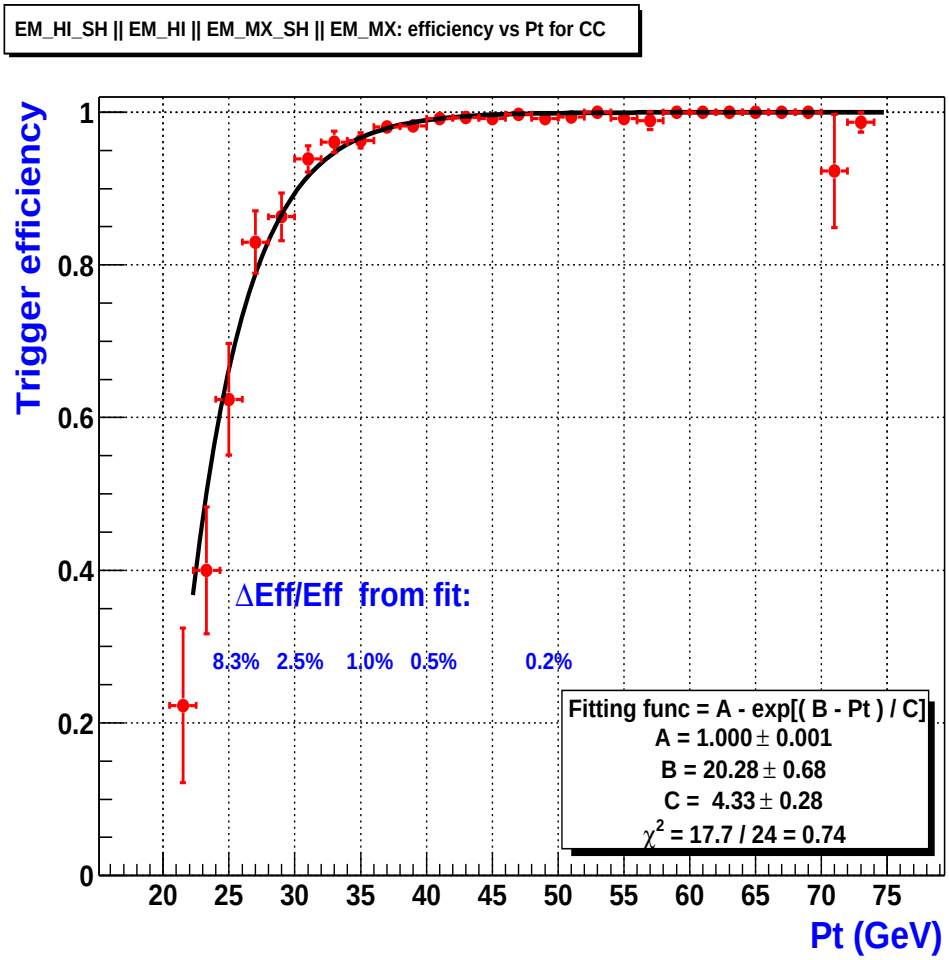

Figure 3: Trigger efficiency as a function of p_T for the trigger combination EM_HI_SH || EM_HI || EM_MX_SH || EM_MX. The fitting function as well as the relative fit errors for some points are shown on the plot.

requirements for an electron should equally apply to a photon, we use $Z \to ee$ events [6)]. To estimate efficiency of the trigger combinations vs. p_T, the tag and probe method [31] has been used, the found points as well as the fit parameters are shown on Figs. 3 and 4 for the global trigger versions v8–11 and v12 respectively. The found fit with the errors will be used in formula (1) to calculate the cross section.

4.2.2 Acceptance.

The acceptance A takes into account the events lost due to various geometric selection criteria which are aimed at keeping just EM clusters in the fiducial regions in η and ϕ of the calorimeter

[6)] In addition, since the number of radiation lengths in Run II in CC region before calorimeter varies as 3.5–5.5 X_0 for (the larger the larger η) the differences between electron and photon should even smaller.

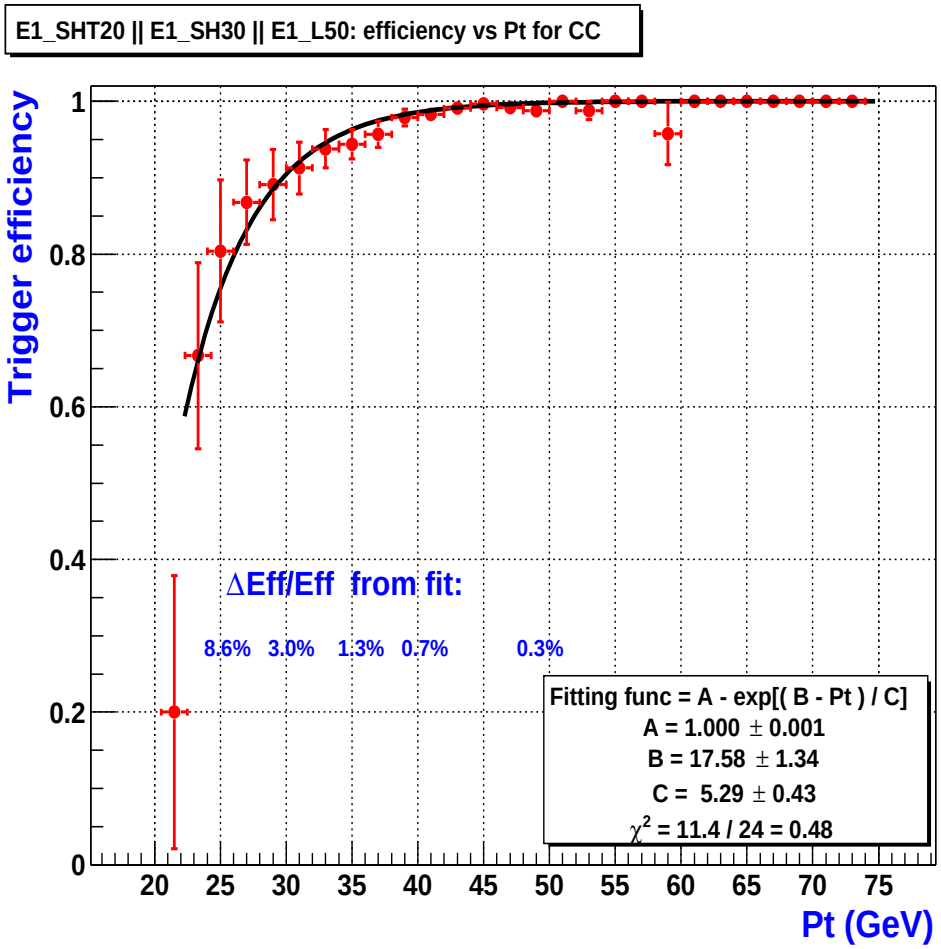

Figure 4: Trigger efficiency as a function of p_T for the trigger combination E1_SHT20 || E1_SH30 || E1_L50. The fitting function as well as the relative fit errors for some points are shown on the plot.

(i.e. to avoid inter-calorimeter section boundaries and edges) [26] where we best understand them.

The acceptance is calculated first from MC "γ^{dir} +jet" events. First, we select direct photons in different p_T^γ intervals that have $|\eta| < 0.9$ on the particle level. We chose the reconstructed EM cluster to be in the same p_T^γ interval [7)] and in correspondence with the photon by requirement $R \leq 0.25$ where R is a distance in the $\eta - \phi$ space between photon and EM cluster. Acceptance is defined for a given p_T^γ interval as a fraction of generated events N_{gen} that pass two reconstruction level cuts N_{pass}: $|\eta_{det}| < 0.9$ and EM fiducial cuts

$$A(\Delta Pt) = \frac{N_{pass}}{N_{gen}} \tag{2}$$

[7)] $\pm\sigma_{em}$ for a given p_T^γ scale (see (18) of Appendix B).

For a wide range the found from Monte Carlo acceptance agrees with the constant 0.784 ± 0.003.

Unfortunately, it turned out that Monte Carlo simulation does not reproduce correctly energy losses in the cracks between azimuthal module boundaries [27]. A particle that enters the calorimeter near the boundary loses some part of its energy to the cracks which shift the cluster centroid in phi towards center of the module (i.e. to a *good* phi fiducial region). To check MC description of energy losses near the phi cracks we have considered $Z \rightarrow ee$ events in MC and data that satisfy criteria: $86 < M^{ee}_{inv} < 97$ GeV/c^2, $Iso(\Delta R02) < 0.10$, EMfrac of $e^{\pm}$ is greater 0.95 and each $e^{\pm}$ has spatially matched track with probability > 0.1. Fig. 5 demonstrates shifts of electron phi cluster position out of the phi cracks to the center of phi modules (Here $phimod$ is defined as $phi * 16/\pi \ \% \ 1.0$). In data this shift is more and thus a larger (than in MC) fraction of electrons are ascribed to a good fiducial region.

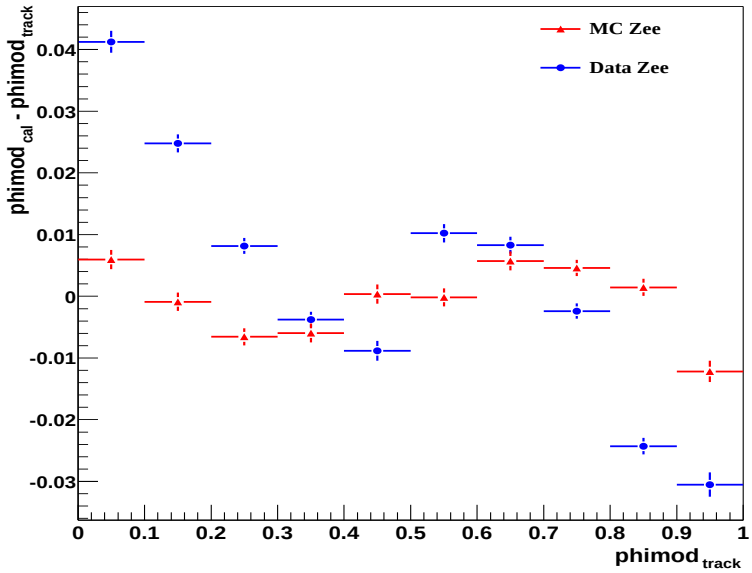

Figure 5: Average shifts of calorimeter (EM3) *phimod* (defined as $phi \cdot 16/\pi \ \% \ 1.0$) w.r.t. track *phimod* as a function of track *phimod*. Two cases that correspond to electrons from $Z \rightarrow ee$ decay in MC (red) and data (blue) are shown.

Fig. 6 presents normalized distributions of the number of events over calorimeter (EM3) *phimod*. Three considered cases correspond to electrons from $Z \rightarrow ee$ in MC (red), data (blue) and to MC photons (purple). One can see a difference in the fraction of events in MC and data on the edges of phi modules ($phimod < 0.1$ or > 0.9). Fig. 7 shows efficiencies w.r.t. event selection 'in phi fiducial' region for three types of events: MC (red), data (blue) electrons from $Z \rightarrow ee$ decay and MC photons (purple). Good agreement of electron and photon MC efficiencies can be noted. Fig. 7 shows efficiencies w.r.t. event selection 'in phi fiducial' region for MC photons. A linear behavior (with fit 0.861 ± 0.004) at whole p^{γ}_T range ($23 \div 260$ GeV) is seen. From Figs. 6 and 7 one can note an equivalent behavior of MC electrons and photons. Fig. 9 shows ratio of MC-to-data 'in phi fiducial' efficiencies. Thus, in data we select, on the average, by 7.6% more events than in MC by 'in phi fiducial' cut.

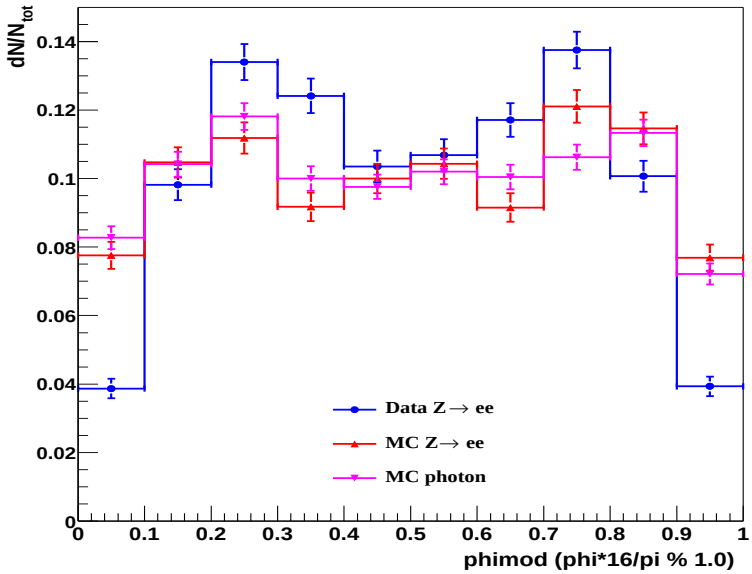

Figure 6: Normalized distributions of the number of events over calorimeter (EM3) *phimod*. Three considered cases correspond to electrons from $Z \rightarrow ee$ in MC (red), data (blue) and to MC photons (purple).

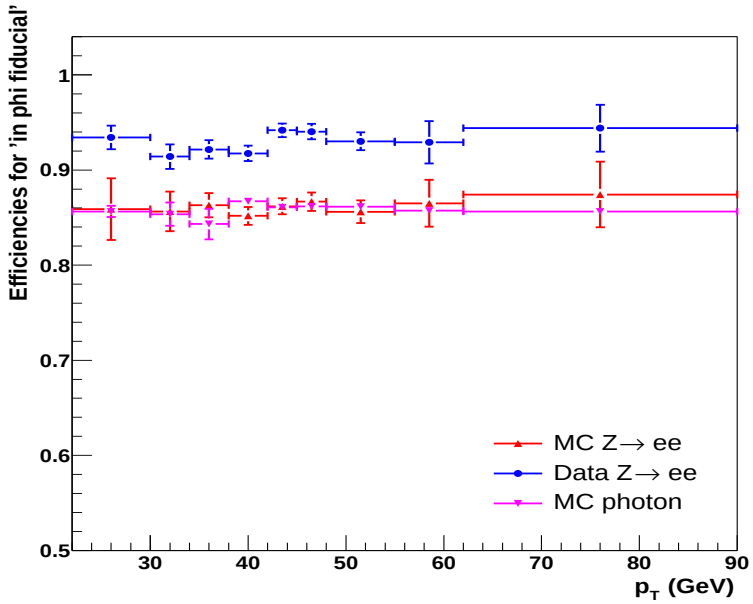

Figure 7: Efficiencies w.r.t. event selection 'in phi fiducial' region for three types of events: MC (red), data (blue) electrons from $Z \rightarrow ee$ decay and MC photons (purple). Good agreement of electron and photon MC efficiencies can be noted.

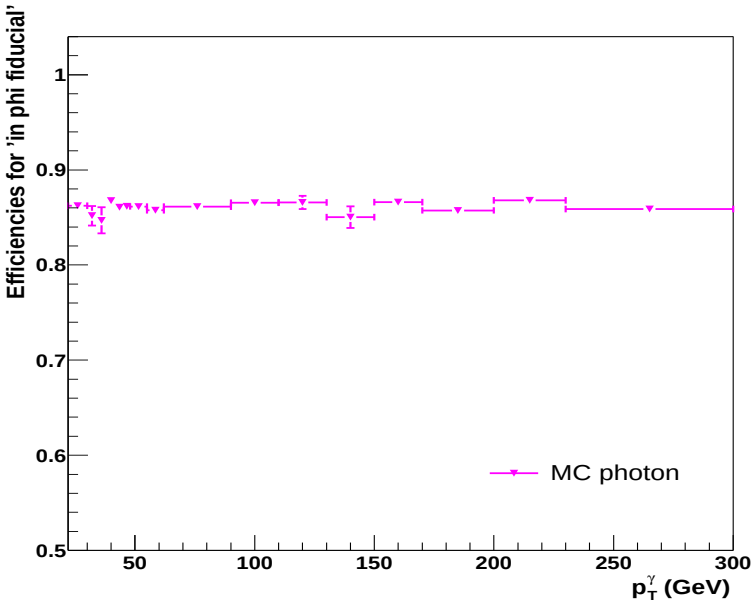

Figure 8: Efficiencies w.r.t. event selection 'in phi fiducial' region for MC photons. A linear behavior (with fit 0.861 ± 0.004) at whole p_T^γ range is seen.

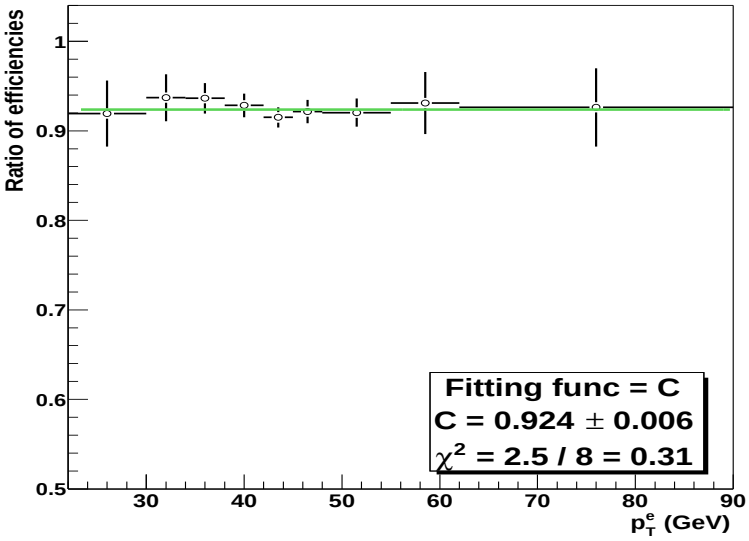

Figure 9: Ratio of MC/data efficiencies w.r.t. event selection 'in phi fiducial' region. The fitting line with fit value is shown on the plot.

The final corrected photon acceptance as a function of p_T^γ is shown in Fig. 10. We see from the plot that it can be fitted by the constant 0.842 ± 0.002.

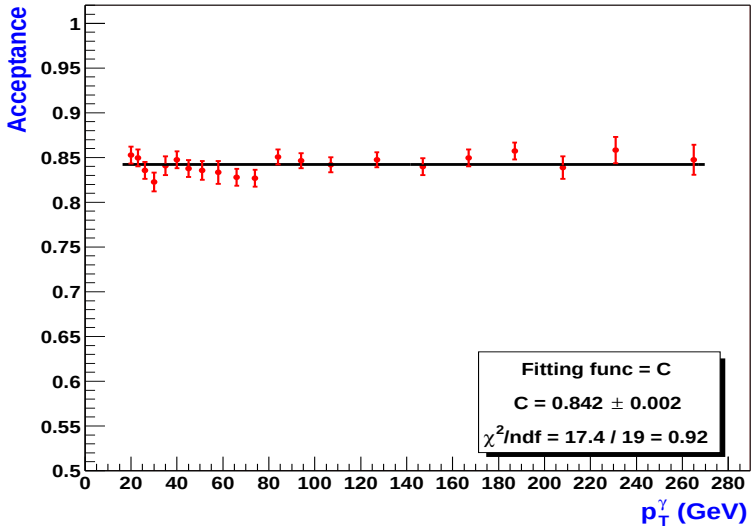

Figure 10: The geometric acceptance as a function of direct photon p_T for $|\eta| < 0.9$ found from MC simulation.

In spite of good agreement between MC 'in phi fiducial' efficiencies for photon and electrons (shown in Fig. 7) we admit a possible difference between them in data within 1%. Taking also into account a possible dependence of acceptance on PDF within 1% [33] our final value of geometric acceptance is 0.842 ± 0.015.

4.2.3 Vertex selection efficiency.

We also keep just the events coming from primary vertex with $|Z_{vtx}| < 50\ cm$ where Z_{vtx} is the distance along the beam axis from the center of the detector. The qualitative requirement on the number of tracks matched to the vertex is imposed: $N_{trk} \geq 3$. The efficiency in z vertex cuts is estimated by taking the ratio of events that passed photon selection cuts (section 4.2.4) with and without z vertex requirements. Here we assume that the QCD background events remained after application of photon selection cuts mostly have 1-jet topology as the signal events.

It was found that efficiency for z vertex cuts depends [28] as on the instantaneous luminosity L_{inst} as on the photon transverse momentum p_T^γ. Dependence of vertex efficiency on the luminosity is presented in Fig. 11. As we see it can be fitted by function

$$\epsilon_1^{vtx}(L_{inst}) = 1 - a_1\ exp[b_1\ L_{inst}] \tag{3}$$

with $a_1 = 0.128 \pm 0.003$ and $b_1 = (-1.422 \pm 0.108) \cdot 10^{-2}$. Since most of the considered here events have $L_{inst} \sim 20 - 30\ cm^2 s^{-1}$ (see Fig. 12) the average efficiency is ~ 0.90.

Dependence of vertex efficiency on p_T^γ is shown in Fig. 13. This dependence is parametrized by the polynomial of second power

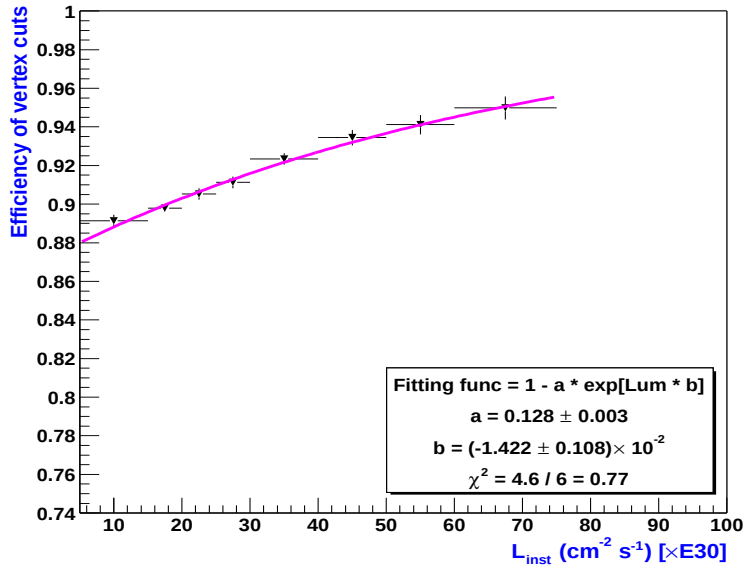

Figure 11: Vertex cuts selection efficiency as a function of instantaneous luminosity.

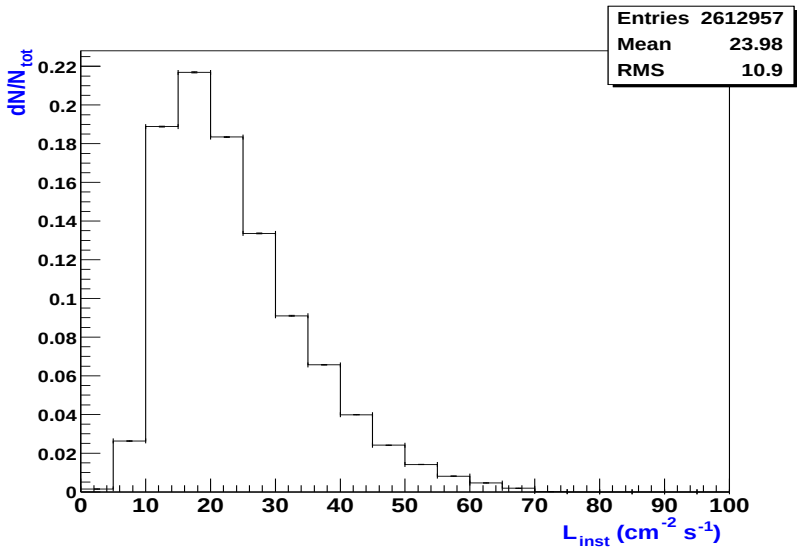

Figure 12: Normalized distribution of the number of selected events over instantaneous luminosity (in units of 10^{30}).

$$\epsilon_2^{vtx}(p_T^\gamma) = a_2 + b_2\, p_T^\gamma + c_2\, (p_T^\gamma)^2 \tag{4}$$

with $a_2 = 0.901 \pm 0.001$, $b_2 = (1.180 \pm 0.384) \cdot 10^{-4}$ and $c_2 = (-1.411 \pm 0.291) \cdot 10^{-6}$ ($\chi^2/ndf = 0.86$).

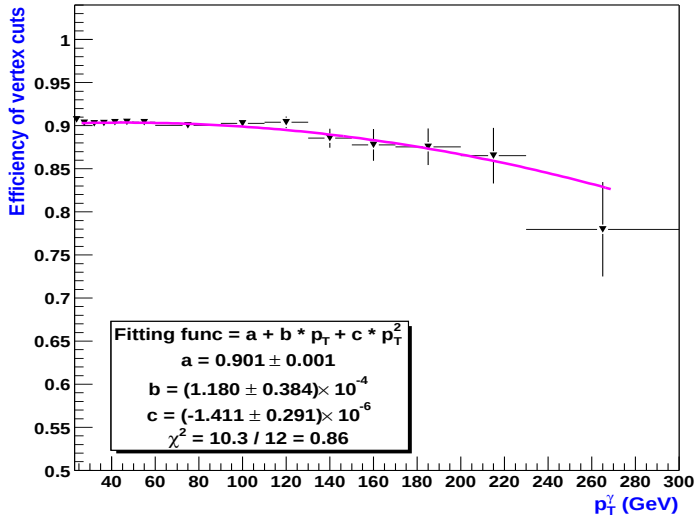

Figure 13: Vertex cuts selection efficiency as a function of photon transverse momentum p_T^γ.

A combined dependence of the vertex selection efficiency on L_{inst} and on p_T^γ can be expressed as

$$\epsilon_{1,2}^{vtx}(L_{inst}, p_T^\gamma) = \epsilon_1^{vtx}(L_{inst}) \cdot \epsilon_2^{vtx}(p_T^\gamma)/\epsilon_2^{vtx}(\langle p_T^\gamma \rangle) \tag{5}$$

where $\langle p_T^\gamma \rangle$ is mean p_T^γ value in the selected sample (what is $\simeq 30.2$ GeV).

One needs to note here that since we select events without track match to the photon EM cluster, there is some probability to misidentify the event primary vertex. Study of systematic error in the final cross section due to the choice of vertex position is given in Appendix E.

4.2.4 Photon selection efficiency.

Since the cross section of direct photons is about 10^3 lower than for jets we have to apply powerful criteria for the background suppression. Below we define two sets of criteria.

Main set.

The main set of criteria used for such a suppression is the following [8]:

- EMfrac > 0.95;

[8] see also section 3.

- $Iso(\Delta R02) < 0.10$;
- Probability for the EM cluster to have any spatially matched track is less than 0.001;
- $E_T^{miss}/p_T^{\gamma} < 0.70$

The presented criteria should significantly reduce the number of QCD jets that have energetic hadrons in the core and/or noticeable energetic activity around the EM cluster. Since most jets produce charged hadrons one can significantly reject corresponding events by the third criterion [9)]. The last cut rejects events with large missing E_T.

To estimate the efficiency for direct photons to pass those cuts we have used the fully simulated and reconstructed MC "γ^{dir} +jet" events. Only photons with reconstructed energies that are in the same p_T regions as the generated photons are used for the selection efficiency determination. The efficiency to pass cuts 'EMfrac > 0.95' and '$Iso(\Delta R02) < 0.10$' is shown in Fig. 14.

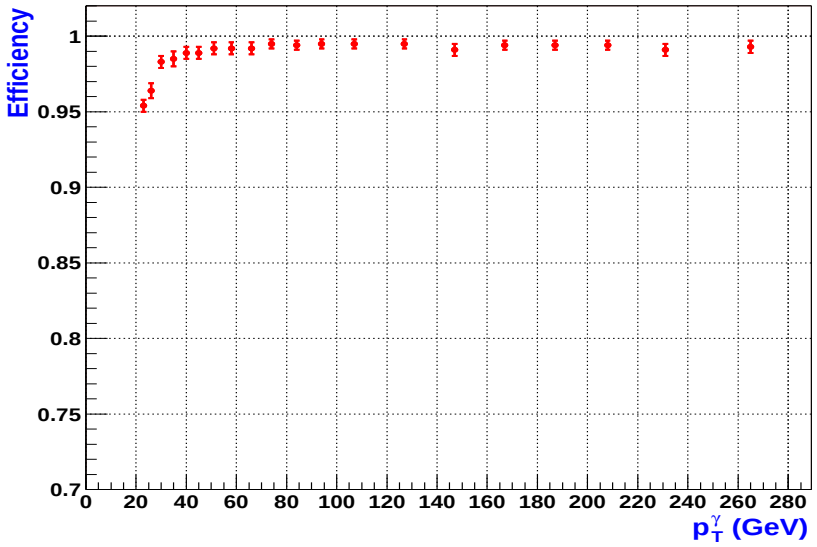

Figure 14: Photon selection efficiency to pass criteria 'EMfrac > 0.95' and '$Iso(\Delta R02) < 0.10$' as a function of p_T^{γ}.

Here we assume that electromagnetic shower initiated by the photon is modeled well in the program that simulates the response of the DØ detector [32]. Comparison of the EM fraction deposited by $e^{\pm}$ from Z decay [33] in MC and data shows good agreement. One only has to accept that the small difference between photon and electron induced showers are reproduced well by D0gstar [32] in order to conclude that the photon efficiency w.r.t. the cut EMfrac> 0.95 found from the MC simulation is accurate.

Others[33] have noted that there may be a deviation between the efficiency of the isolation criterion in data and MC due to the calorimeter noise in data which is not modeled in MC. In this

[9)] Besides, the neutral hadrons with multiphoton final state of their decay channel would have a bigger conversion probability to $e^{\pm}$ pair that it is for a single photon.

case the efficiency found in MC w.r.t. this cut would be higher than in the data. To estimate the difference in efficiency of the isolation cut we have used the $Z \to ee$ MC and data events. These events were selected by the following criteria: $86 < M^{ee}_{inv} < 97$ GeV/c^2, EMfrac of $e^{\pm}$ is greater 0.95 with spatially matched track for each of $e^{\pm}$. The found efficiencies to pass criterion $Iso(\Delta R02) < 0.10$ is shown in Fig. 15 for MC and data electrons/positrons.

Their ratio can be parametrized (with $\chi^2/ndf = 0.56$) by

$$r_{eff} = 0.952 \pm 0.016 + (0.745 \pm 0.274) \times 10^{-3} \times p_T^{\gamma}. \quad (6)$$

We have applied this correction to the total selection efficiency in the interval of $23 \leq p_T^{\gamma} \leq 60$ GeV (i.e. in the region where this MC/data disagreement is observed).

The corrected efficiency is shown in Fig. 16. It also includes requirement of 'no matched track' with 'track_match_spatialchi2prob() < 0.001'.

A loss of photon events due to the tracking criterion 'track_match_spatialchi2prob() < 0.001' may be caused by a photon conversion to $e^{\pm}$ pair in the tracker media with a following track reconstruction and assignment of this track to the EM cluster [10)]. One can expect that such a probability is very small. There is also the probability that a random track in the event (from an additional $p\bar{p}$ interaction, ISR/FSR effects, from spectator quarks or even QCD jet) overlaps with the photon EM cluster.

To build confidence that the simulated MC is close to reality, we have compared MC/data track reconstruction efficiencies for $e^{\pm}$ from Z^0 decay. For this aim we have chosen $e^{\pm}$ by the criteria described in the previous paragraph and calculated the efficiencies using formula

$$\epsilon_{trk} = \frac{2\ N_{2trk}}{2\ N_{2trk} + N_{1trk}} \quad (7)$$

where N_{2trk} is number of $Z^0 \to ee$ events with two (electron and positron) EM clusters spatially matched to tracks with probability ≥ 0.001 and N_{1trk} is that with only one of (electron/positron) EM clusters spatially matched to a track with probability ≥ 0.001. The found values are 0.938 ± 0.004 in data and 0.944 ± 0.003 in MC, i.e. with just 0.6% difference on the average [11)]. Such a small difference when multiplied by the conversion probability leads to a completely negligible effect.

To estimate the number of additional tracks matched to EM cluster we have again taken EM clusters produced by the $e^{\pm}$ from Z^0 decay. The total number of tracks matched to one EM cluster is shown in Fig. 17. Additional tracks were found in 0.72$\pm$0.13% of events in MC and in 1.10$\pm$0.16% in data.

Given the confidence built by these comparisons of MC and data, we rely on the MC for finding efficiency with respect to the "track_match_spatialchi2prob() < 0.001" criterion.

[10)] Note that this is not the same as the track reconstruction for $e^{\pm}$ produced in the event vertex.

[11)] Analogous values for looser cuts, EMfrac> 0.90 and are $Iso(\Delta R02) < 0.15$ are 0.924 ± 0.005 for data and 0.942 ± 0.003 for MC.

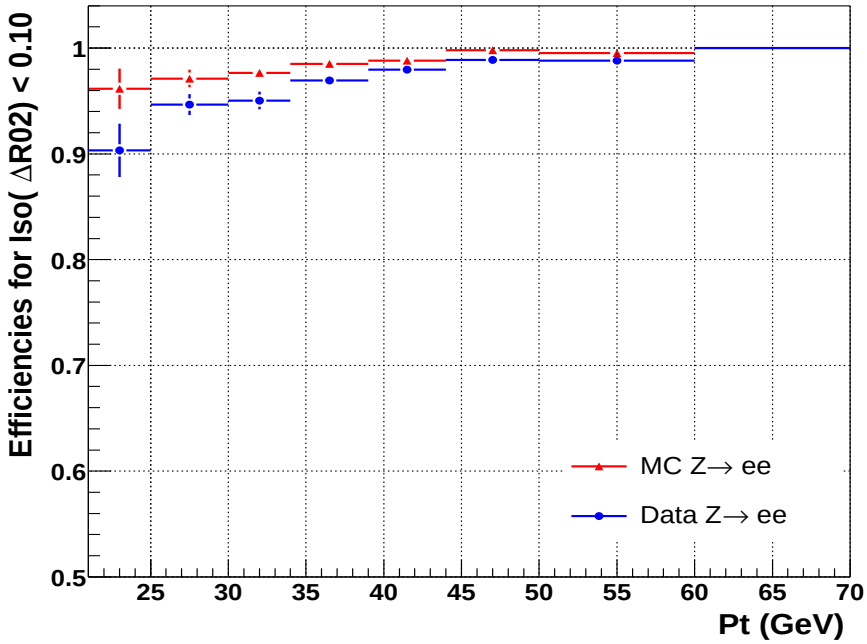

Figure 15: Selection efficiency in MC (red) and data (blue) w.r.t. the $Iso(\Delta R02) < 0.10$ cut as a function of p_T.

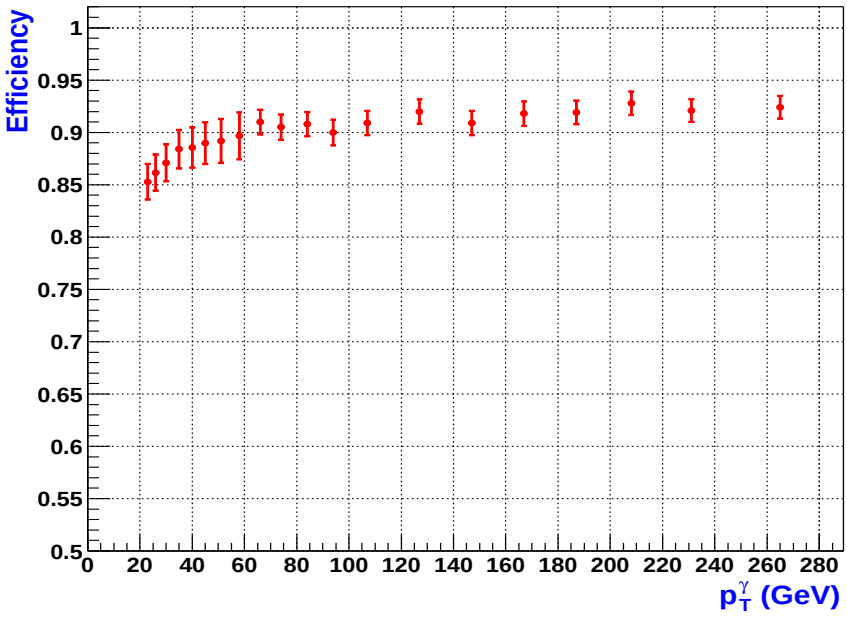

Figure 16: Photon selection efficiency to pass criteria 'EMfrac > 0.95', '$Iso(\Delta R02) < 0.10$' and 'track_match_spatialchi2prob() < 0.001' as a function of p_T (corrected after $Z \rightarrow ee$ MC/data comparison).

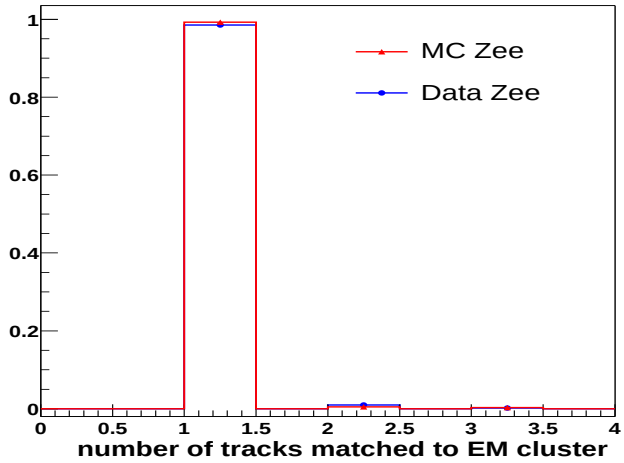

Figure 17: Total number of tracks matched to one EM cluster in $Z^0 \to ee$ event: MC is red and data is blue histograms. Number of events with 2 (and more) matched track is 0.72±0.13% in MC and in 1.10±0.16% in data.

The events that contribute to the inclusive photon production (mainly with direct photon in final state) are expected to have small missing transverse energy $E_{_T}$ (Example of a good candidate to 'direct photon + jet' event (XY view) in data is shown in Fig. 18). Thus we need to apply this cut to get rid of events with large missing $E_{_T}$ ($E_{_T}^{miss}$). The sources of such events can be $W \to e\nu$ events [12)] and events that are not caused by $p\bar{p}$ interaction, for example by cosmics. Example of such event (XY view) in data is presented in Fig. 19. It was found that contribution of events with large missing $E_{_T}$ is especially big at high p_T^γ [13)]. Distribution of the number of events in data over $E_{_T}^{miss}/p_T^\gamma$ after application of 'EMfrac > 0.95', '$Iso(\Delta R02) < 0.10$' and 'track_match_spatialchi2prob() $<$ 0.001' cuts is shown in Fig.20. Values of p_T^γ intervals are shown on the plots. Anomalous contribution of events in the region of $0.8 < E_{_T}^{miss}/p_T^\gamma < 1.4$ is clearly seen (that becomes especially noticeable for $p_T^\gamma > 70$ GeV). One can also observe that fraction of these events increases with growing p_T^γ.

We have introduced $E_{_T}^{miss}/p_T^\gamma < 0.70$ criterion in order to cut off events with large missing $E_{_T}^{miss}$. The distributions over $E_{_T}^{miss}/p_T^\gamma$ for data and and the same intervals of MC events are shown in Fig. 20 and 21. We have used data to estimate the efficiency w.r.t. the $E_{_T}^{miss}$ cut. Just events with $p_T^\gamma > 50$ GeV (i.e. events having relatively good photon purity) have been used for this aim. We have fitted "signal" part of the $E_{_T}^{miss}/p_T^\gamma$ distribution by convolution of the Gaussian and Landau functions while the cosmic part (around unity) just by the Gaussian (see Fig.22). Then we have subtracted the events due to the Gaussian cosmic tail from the signal tail at $E_{_T}^{miss}/p_T^\gamma < 0.70$. The resulting efficiency is shown in Fig.23. It is worth emphasizing that the efficiency obtained in this way differs from that found from the MC "$\gamma + jet$" events (which

[12)] One needs to note that fraction of $W \to e\nu$ background events is small as compared with QCD events [50, 49].

[13)] This fact confi rms their cosmic origin since the slope of p_T^γ spectrum of is much steeper than that for cosmic rays. Thus their relative fraction should grow with energy.

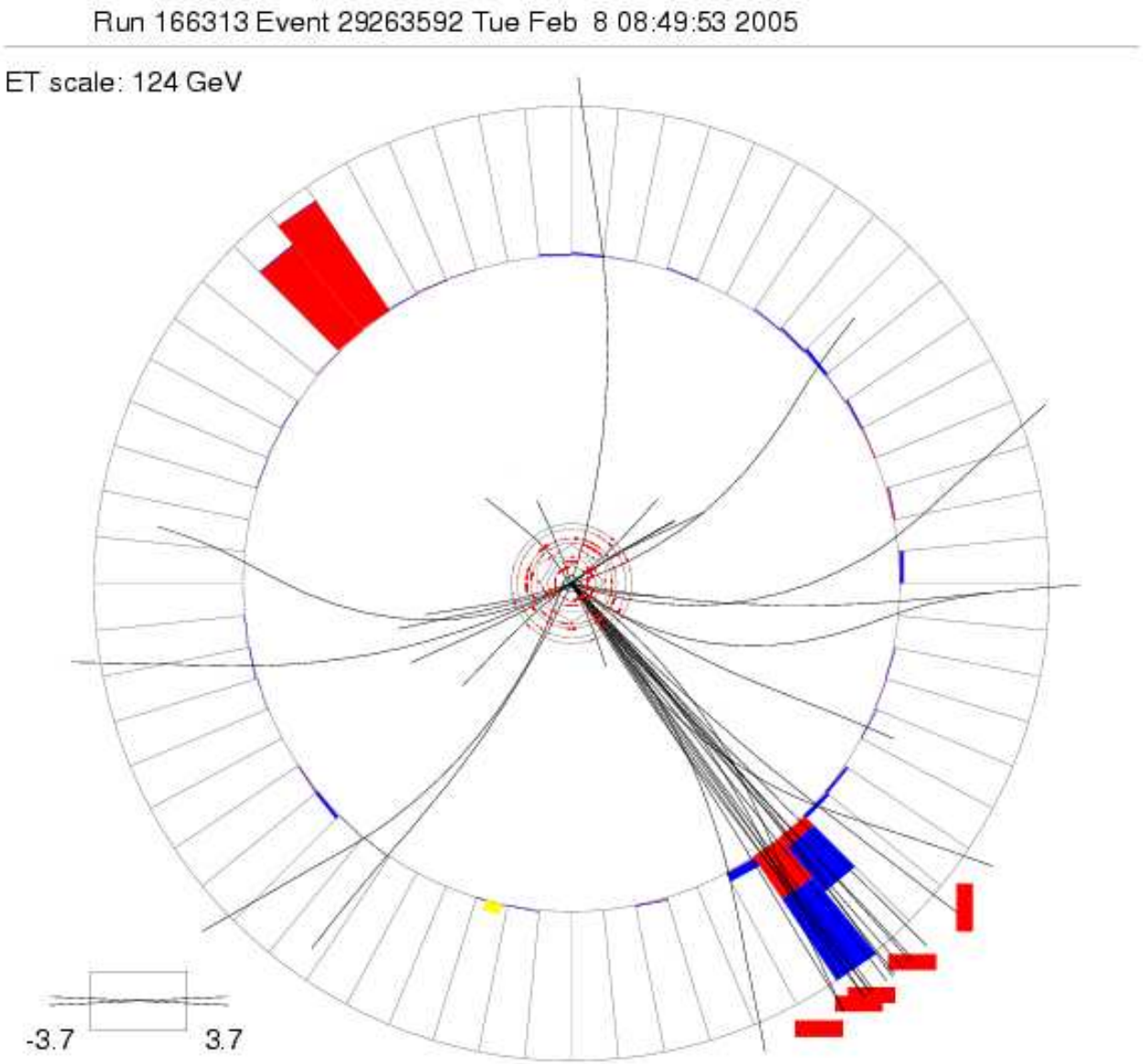

Figure 18: Example of "live" "$\gamma + jet$" event (XY view) obtained from event display. Run and event numbers are shown on the plot.

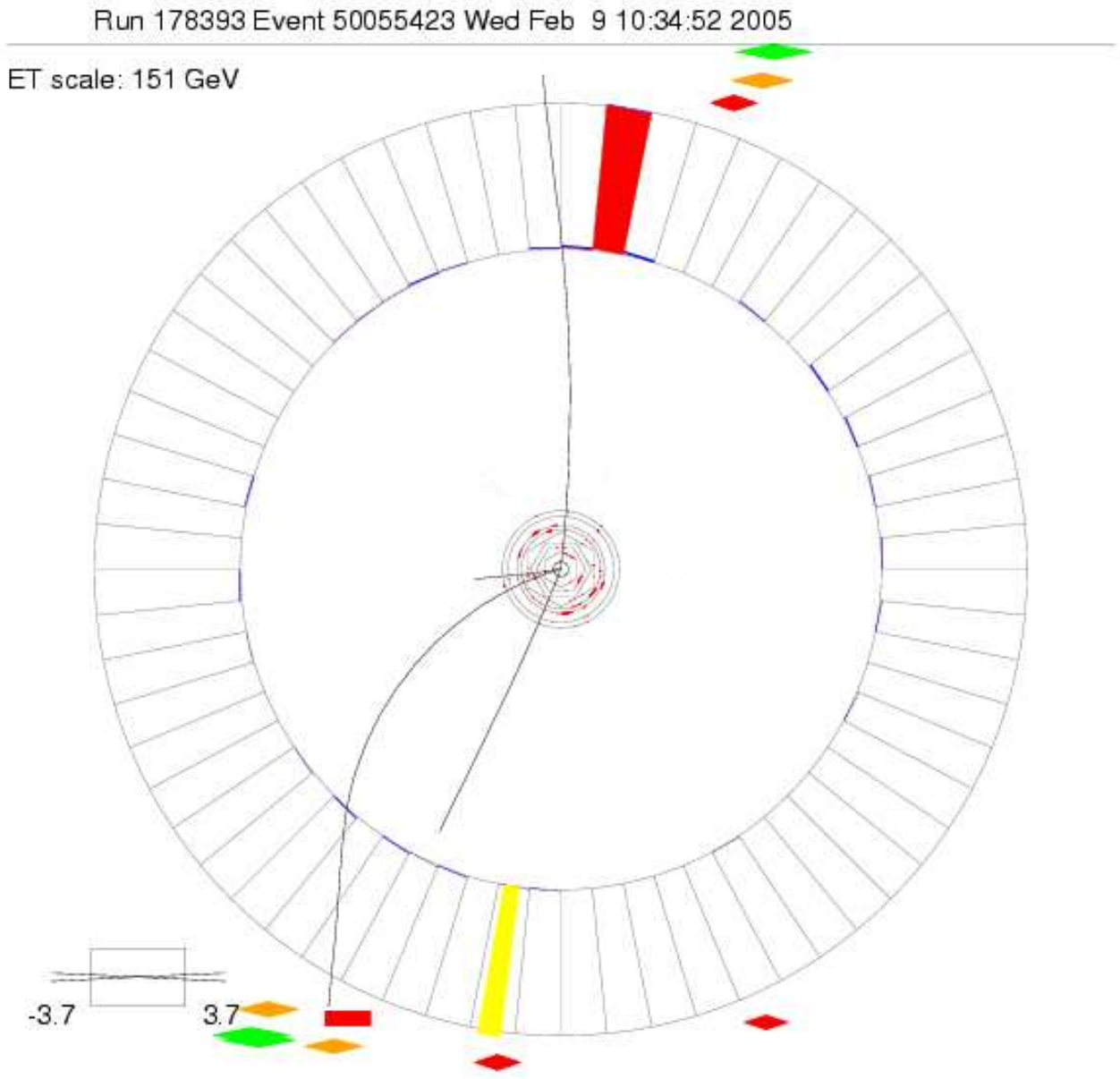

Figure 19: Example of a candidate to cosmic event (XY view) obtained from event display. Run and event numbers are shown on the plot. Missing transverse energy is clearly seen (yellow) in the opposite side w.r.t. EM cluster – photon candidate (red).

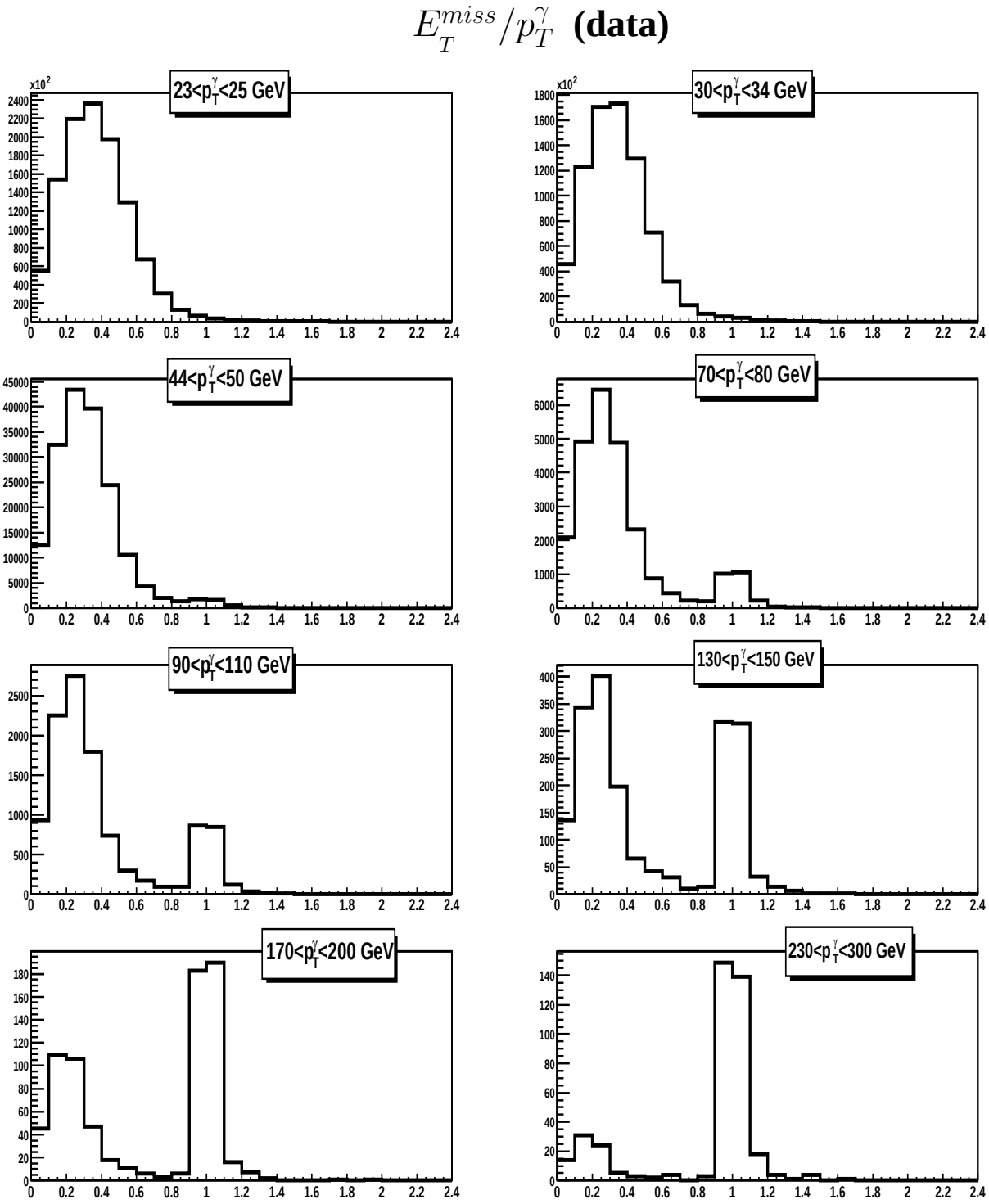

Figure 20: Distribution of the number of events *in data* over E_T^{miss}/p_T^{γ} after application of 'EM-frac > 0.95', '$Iso(\Delta R02) < 0.10$' and 'track_match_spatialchi2prob() < 0.001' cuts. Values of p_T^{γ} intervals are shown on the plots.

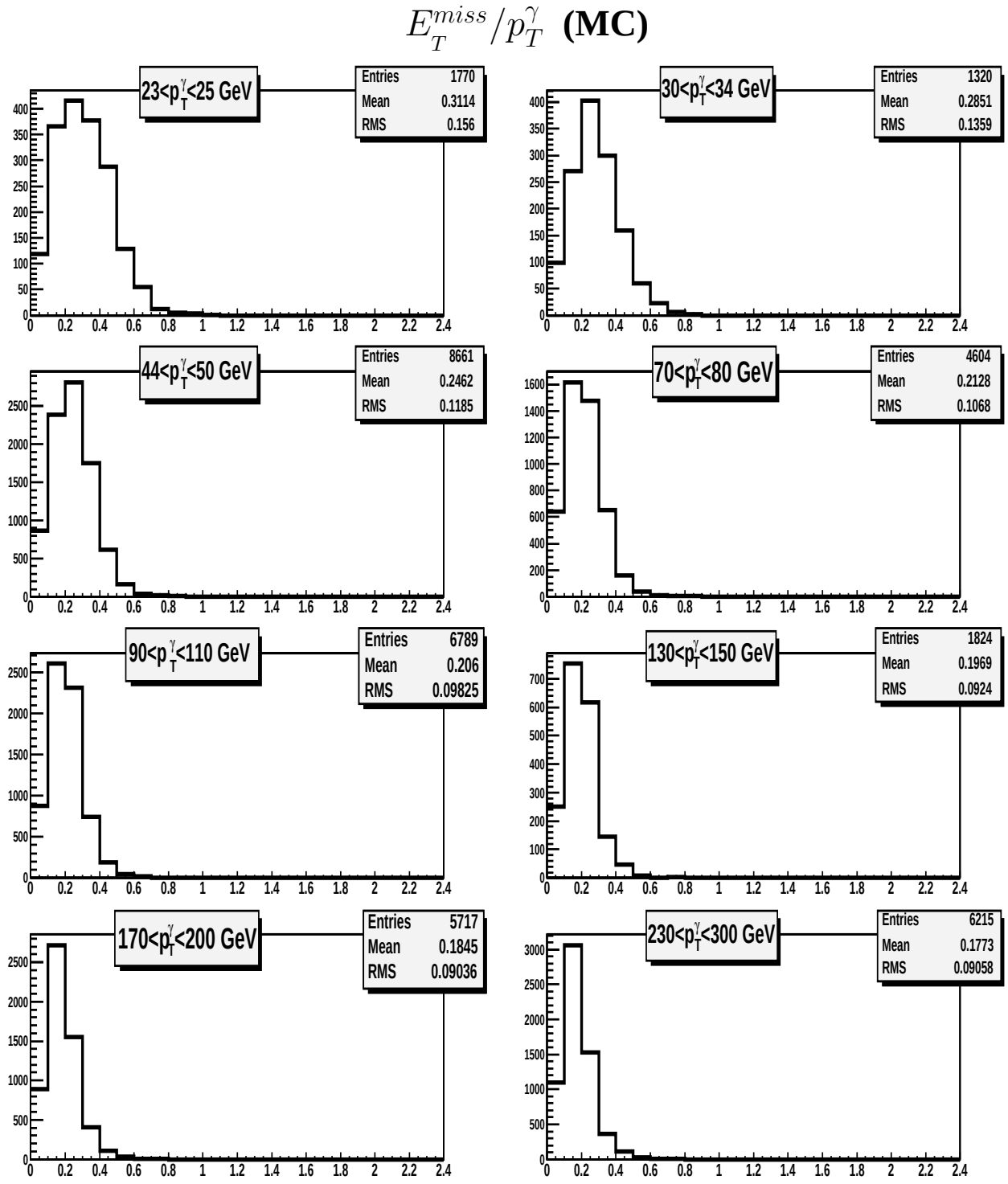

Figure 21: Distribution of the number of events *in MC* over E_T^{miss}/p_T^{γ} after application of 'EM-frac > 0.95', '$Iso(\Delta R02) < 0.10$' and 'track_match_spatialchi2prob() < 0.001' cuts. Values of p_T^{γ} intervals are shown on the plots.

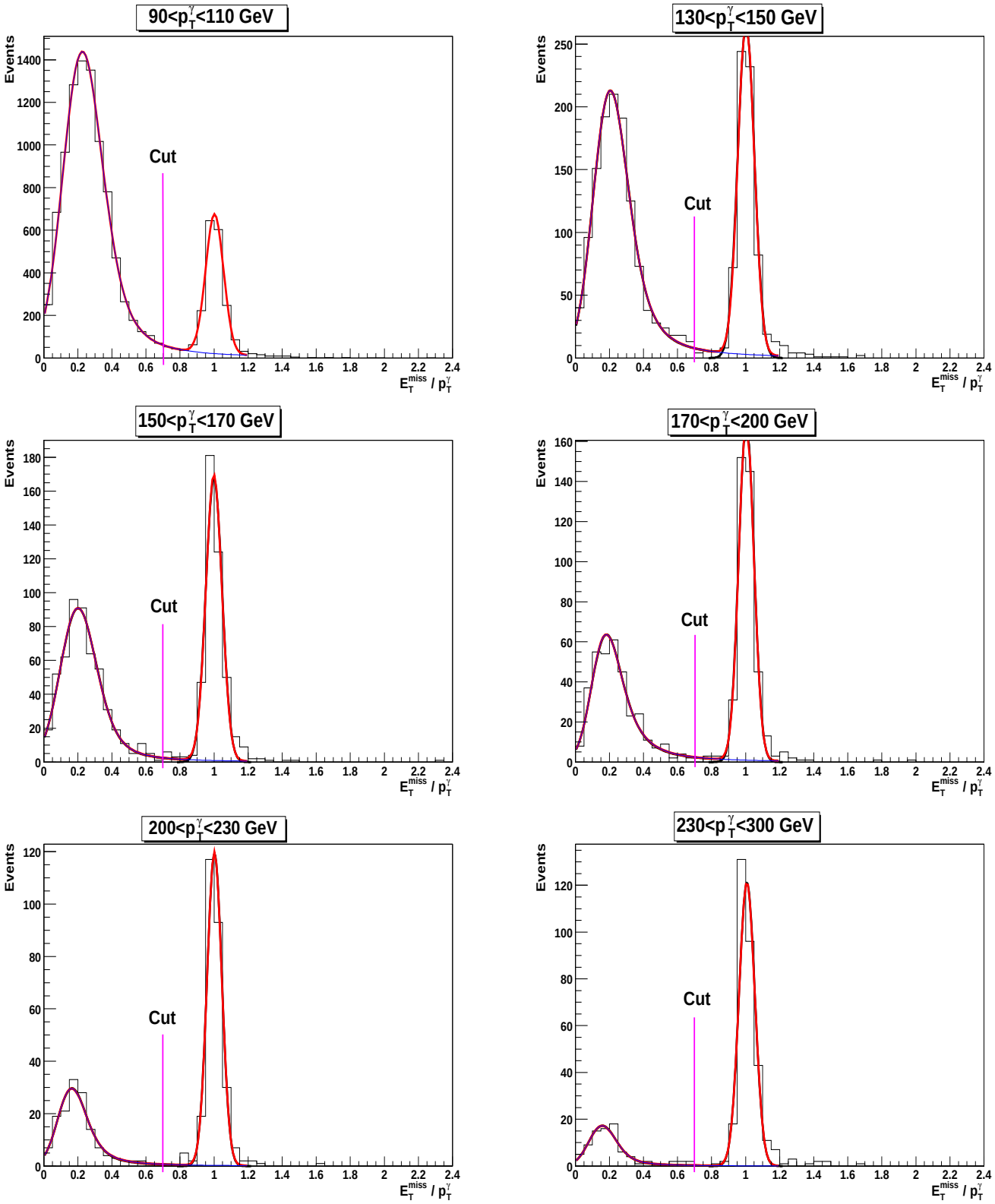

Figure 22: Determination of the pT^{miss}/pT^{γ} cut efficiency from the data.

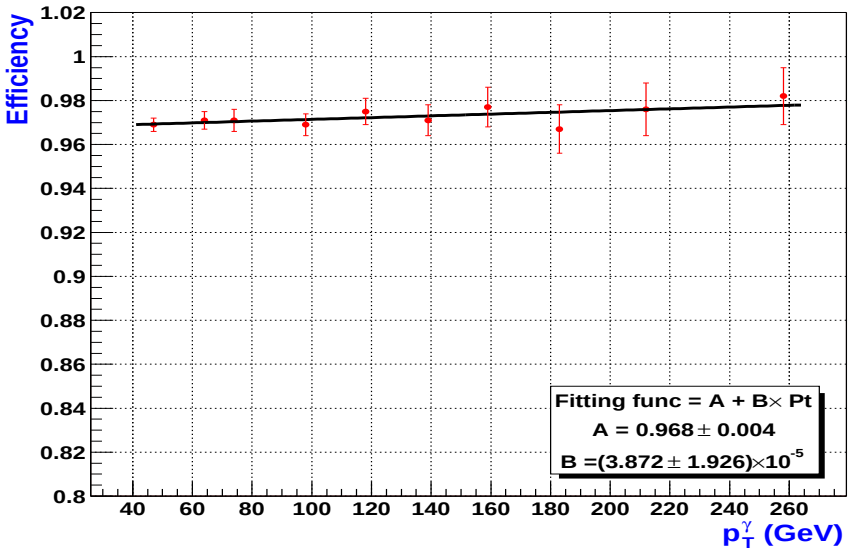

Figure 23: Photon selection efficiency to pass criteria $E_T^{miss}/p_T^\gamma \leq 0.7$ as a function of p_T^γ found from the data (see Fig. 22). The events satisfy conditions 'EMfrac > 0.95', '$Iso(\Delta R02) < 0.10$' and 'track_match_spatialchi2prob() < 0.001'.

is close to 99%) differs just by about 1-2%. The efficiency determined from the data is taken as the base one with 1.5% systematic error.

An account of the last efficiency lead to the final curve for photon selection efficiency that is presented in Fig. 24. It is fitted by $a - exp[(b - p_T)/c]$ with values of parameters:

$$a = 0.895 \pm 0.004, \quad b = -51.23 \pm 7.73 \ \text{ and } \ c = 29.23 \pm 1.53. \tag{8}$$

We admit a dependence of the efficiency on the instantaneous luminosity and add 1% error for this dependence.

Additional set.

The set of additional variables used for a further background suppression was presented earlier in section 3.

It is worth mentioning that these variables are new as compared with Run I and they were introduced, on the one hand, to build a discriminant variable that would allow one to determine a photon fraction (purity) in the selected sample (section 4.4) and, on the other hand, to specify a criterion for increasing the photon purity.

To verify the MC/data agreement with respect to these variables we have tested them on the $Z \to ee$ events. The normalized distributions of those variables for the $Z \to ee$ events in MC and data are shown in Fig. 25.Electrons/positrons from Z^0 decay are required to be within $|\eta_{detector}| < 0.9$ and with $25 < p_T < 75$ GeV. Behavior of efficiencies are shown in Figs. 26, 27 and 28. They are built for each of the variables with p_T of $e^\pm$ in the intervals of $25 < p_T < 40$,

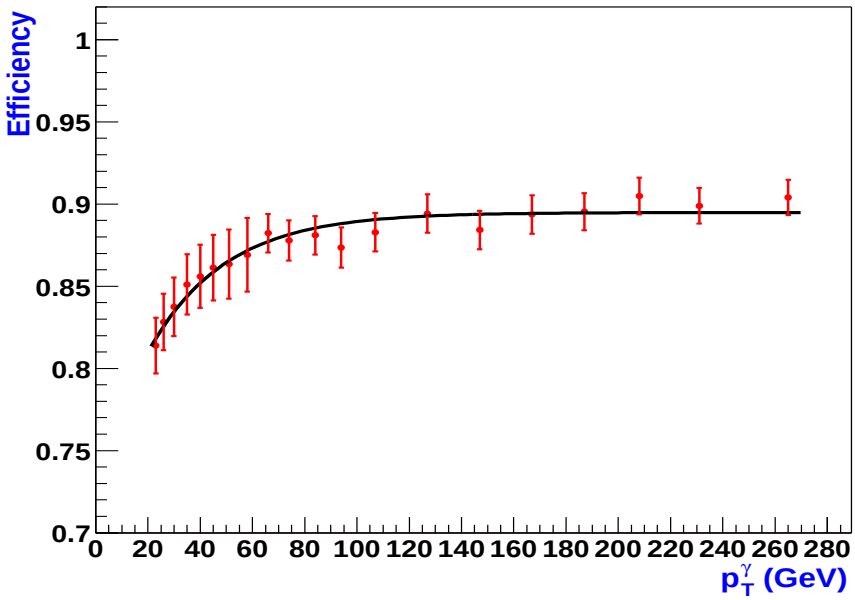

Figure 24: Photon selection efficiency to pass the main cuts: The events satisfy conditions 'EM-frac > 0.95', '$Iso(\Delta R02) < 0.10$', 'track_match_spatialchi2prob() < 0.001' and $E_T^{miss}/p_T^\gamma < 0.70$ as a function of p_T^γ. The fitting function is shown on the plot. The efficiency is fitted by function $a - exp[(b - p_T)/c]$ with $a = 0.895 \pm 0.004$, $b = -51.23 \pm 7.73$, and $c = 29.23 \pm 1.53$.

$40 < p_T < 50$ and $50 < p_T < 75$ GeV, respectively.

After application of such tight limitations as: $ncell\_EM1(EMclust) \leq 2$, $ncell\_EM1\ (\Delta R02) = 0$, $sum\_trackPT\ (\Delta R04) \leq 2$ GeV, and $sigrphi\_EM3 \leq 4$ cm the final efficiency was found to be 75.0% in MC and 73.4% in data. Detailed behavior of efficiencies after step-by-step applications of the cuts is presented in Table 1. Thus, even with such tight selections we do not see noticeable difference between efficiencies in MC and data.

When discriminating between photons and background particles (π^0 as well as the neutral decay channels of η and K_s^0 mesons) we face a typical (for high energy physics) pattern recognition problem. The standard procedure for solving such a problem is the introduction of relevant cuts in the multi-dimensional data. Nowadays the application of a software-implemented artificial neural network (ANN) for pattern recognition is well known and usually gives the results that are superior to conventional approaches [34].

So, instead of direct application of cuts on these variables they were used to build ANN that can accumulate a power of all the four variables and criteria on them. ANN is then applied for an additional selection criterion and calculation of photon purity.

For ANN building we have used JETNET package [35] of version 3.5. ANN is trained to discriminate between direct photon from "γ^{dir} + jet" events and em-jets [14]. The Manhattan algorithm [36] for weight updating was used at the training stage. The network is trained to produce 1

[14] See section 4.1.

Table 1: Selection efficiencies (in %) with tight limitations on the additional variables.

Cut	MC $Z \to ee$	Data $Z \to ee$
$ncell\_EM1(EMclust) \leq 2$	83.2	82.2
$ncell\_EM1\ (\Delta R02) < 1$	77.8	77.1
$sum\_trackPT\ (\Delta R04) \leq 2$ GeV	75.1	73.9
$sigrphi\_EM3 \leq 4$ cm	75.0	73.4

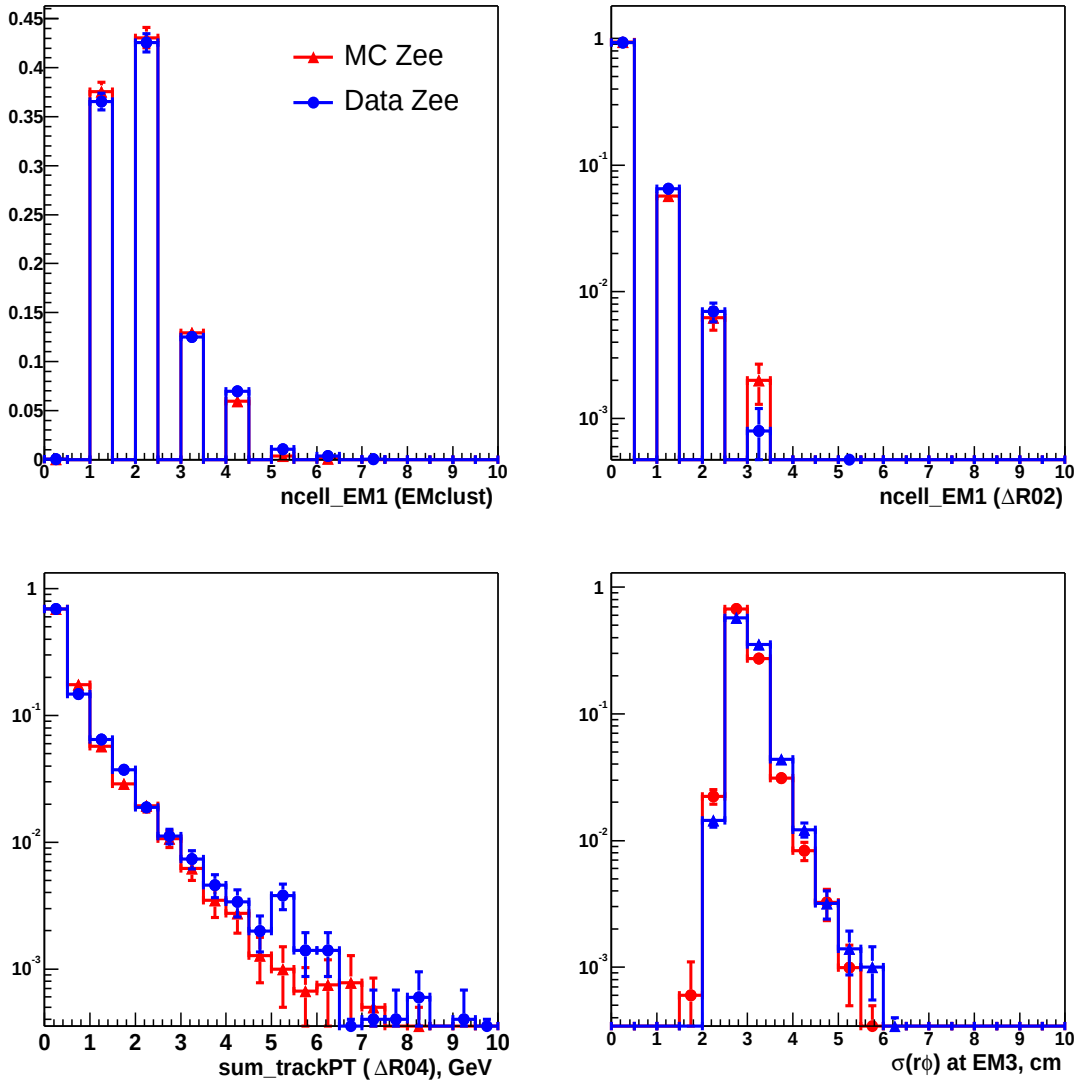

Figure 25: The normalized distributions of the number of the $Z \to ee$ MC (red) and data (blue) events over four *additional* variables of section 3. Electrons/positrons are from the $25 < p_T < 75$ GeV interval.

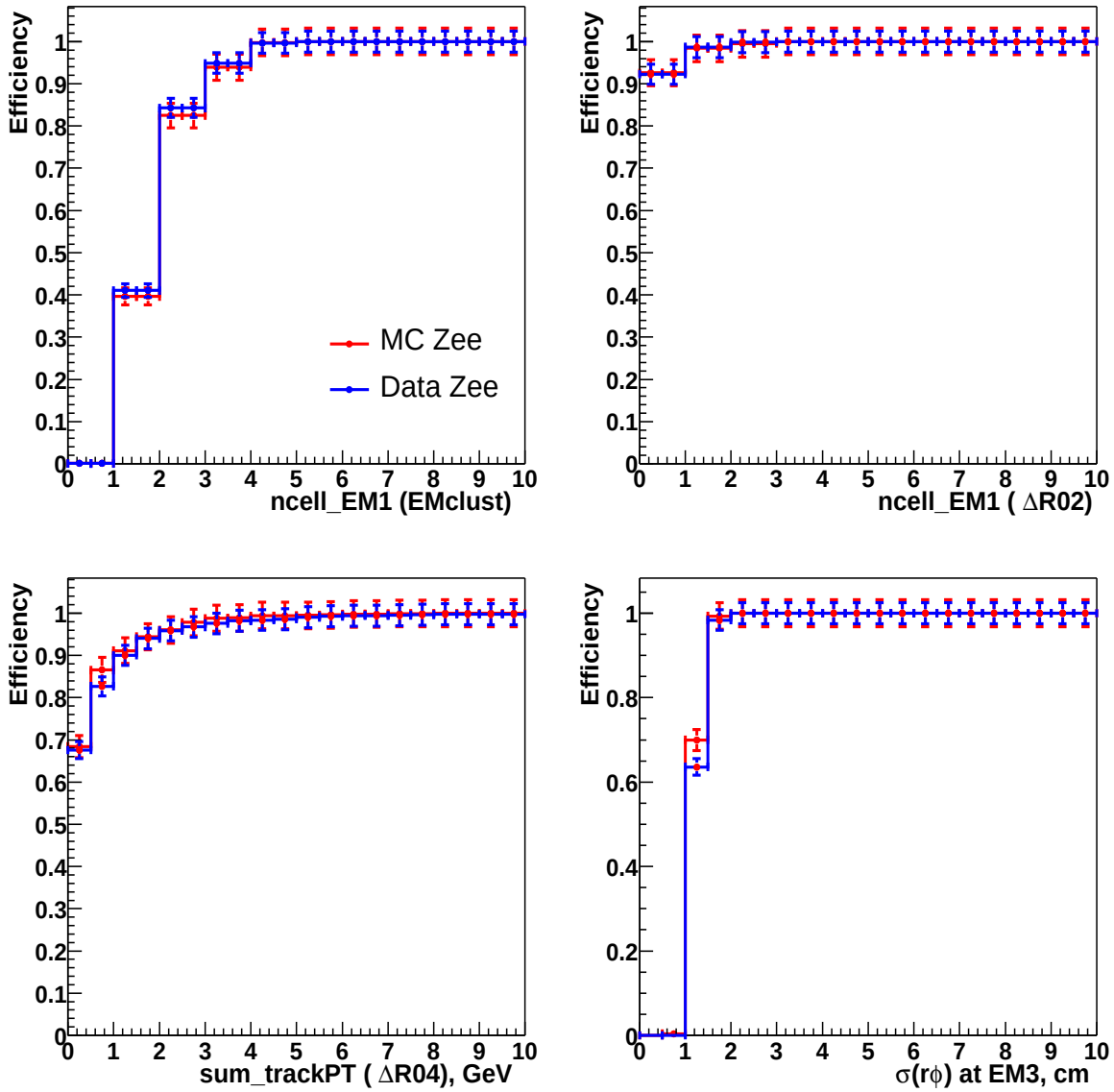

Figure 26: The efficiencies in the $Z \rightarrow ee$ MC (red) and data (blue) events as functions of the four *additional* variables of section 3. Electrons/positrons are from the $25 < p_T < 40$ GeV interval.

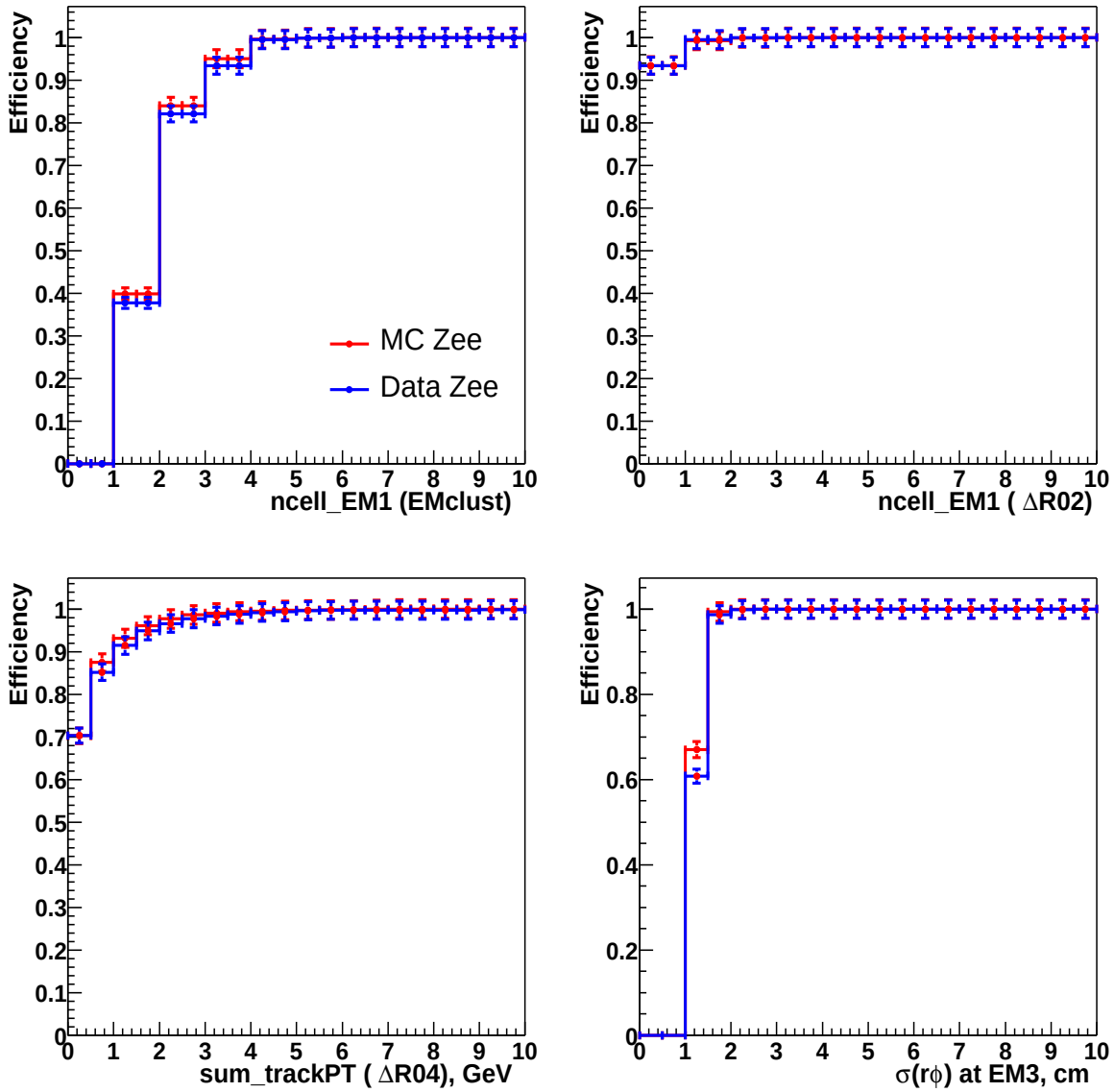

Figure 27: The efficiencies in the $Z \rightarrow ee$ MC (red) and data (blue) events as functions of the four *additional* variables of section 3. Electrons/positrons are from the $40 < p_T < 50$ GeV interval.

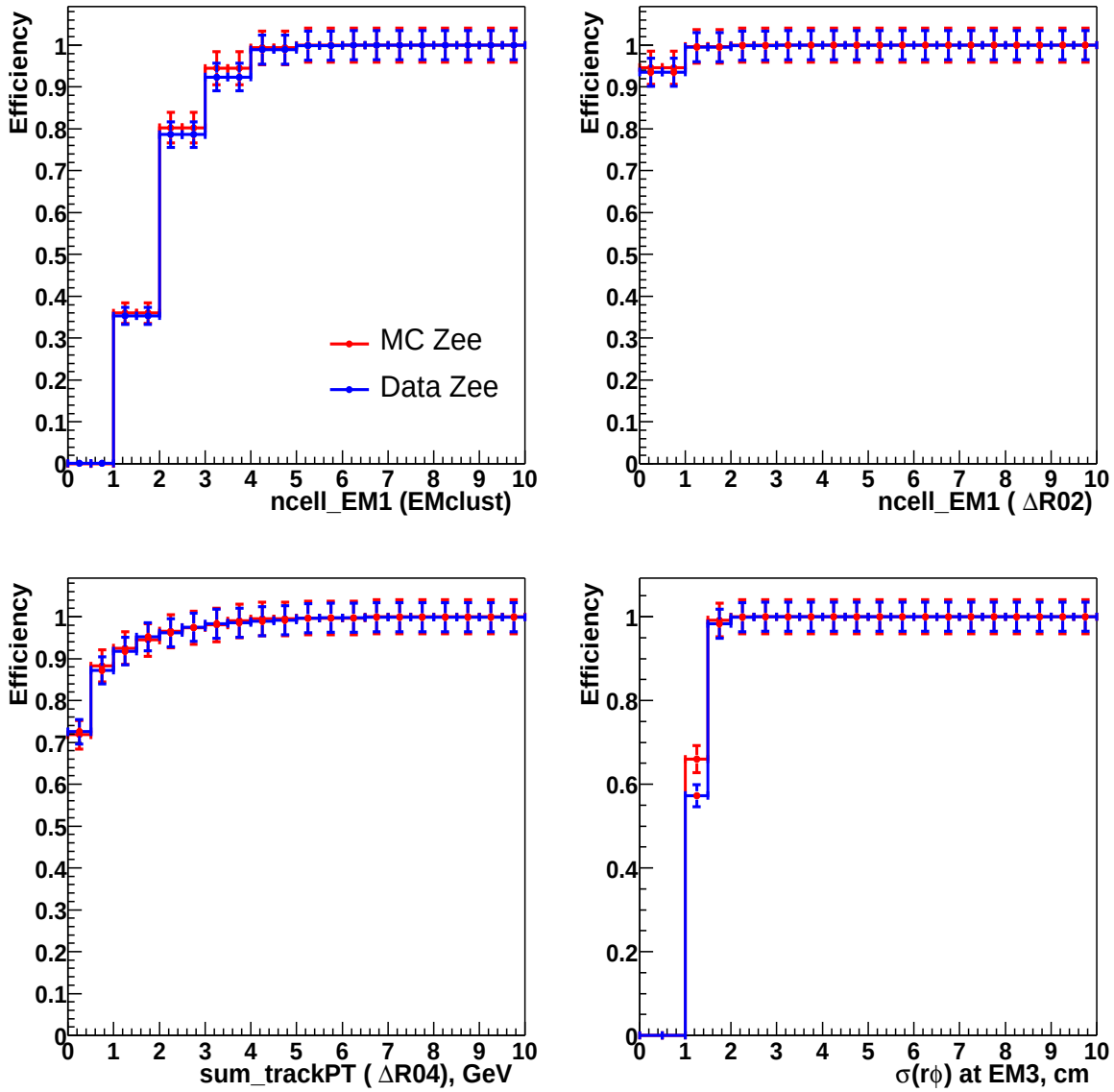

Figure 28: The efficiencies in the $Z \rightarrow ee$ MC (red) and data (blue) events as functions of the four *additional* variables of section 3. Electrons/positrons are from the $50 < p_T < 75$ GeV interval.

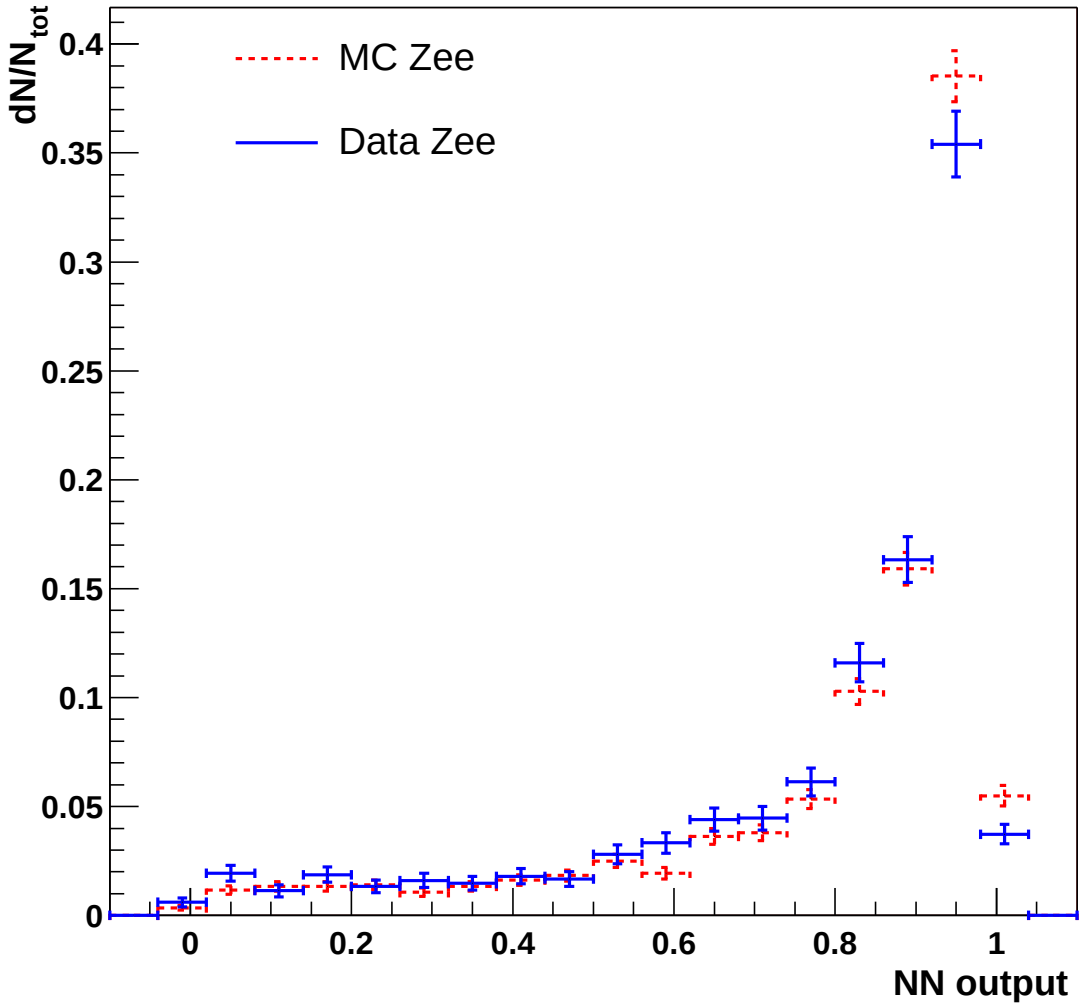

Figure 29: Normalized distribution of ANN output for MC (red) and data (blue) $Z^0 \to ee$ events with $\mathbf{25 < p_T^e < 75}$ GeV and $|\eta^e| < 0.9$ is shown. The ANN was trained to discriminate between photons and em-jets (see section 4.2.4).

Selection efficiencies as a function of NN output

NNout thres.:	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80
MC sel. eff.(%):	100	98.8	97.0	95.0	93.0	90.9	88.2	84.1	77.6
Data sel. eff.(%):	100	98.0	96.3	93.8	91.4	88.5	84.8	80.2	72.7
MC-Data sel.eff.(%):	0.0	0.8	0.6	1.2	1.6	2.4	3.4	3.8	4.8

(unity) in case of signal and 0 (zero) in case background events. For the ANN training we have used MC signal and background events from the interval of $30 < p_T^\gamma < 50$ GeV. Then the built network was tested on the $Z \to ee$ MC/data events (Fig. 29). The difference in the electron selection efficiencies after cuts on NN output is shown below the figure.

We have applied cut on the network output $NN_{output} > 0.5$ as the additional selection criterion. A systematic error assigned for this cut is the difference between the efficiencies for MC and data $Z \to ee$ events and is 2.4% for this cut as is seen from Fig. 29.

Photon efficiency w.r.t. this cut is shown in Fig. 30. It practically does not depend on p_T^γ and agrees with constant 0.937 ± 0.002.

The distributions over NN_{output} for direct photon and em-jets from $44 < p_T^\gamma < 50$ GeV are presented in Fig. 31 with position of the NN_{output} cut. The network output for the data preselected by the same main selection cuts are also shown on this figure (black histogram). The distribution is used in section 4.4 to obtain the photon fraction for a given p_T^γ interval.

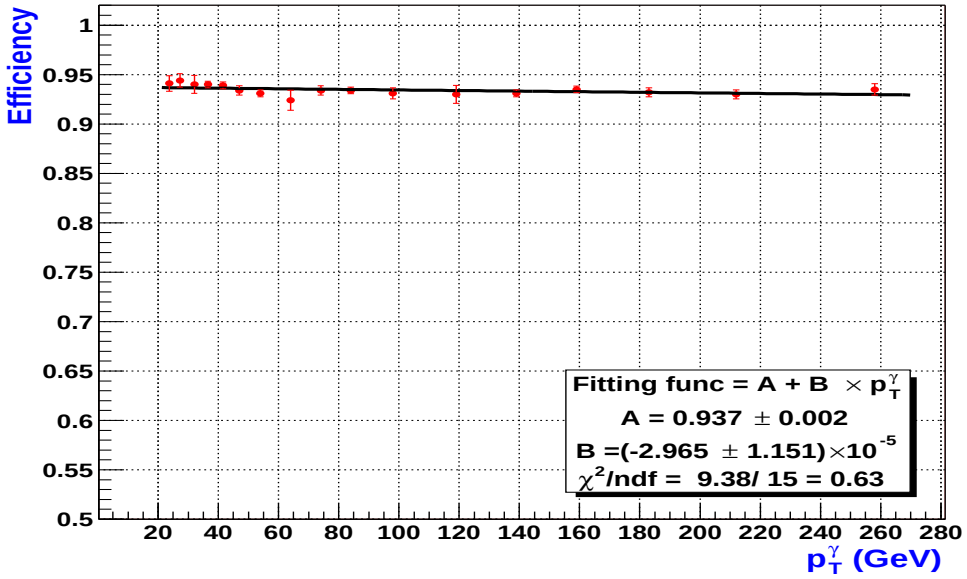

Figure 30: Photon selection efficiency w.r.t. the cut on ANN output $NN_{output} > 0.5$ as function of p_T^γ.

4.3 Photon energy scale correction.

As we know, the calorimeter layer weights used to reconstruct initial energy of the electromagnetic particle were found by using electron based ($Z \to ee, J/\psi \to ee$) events. On the other hand, it is known that photons (due to fundamentally different nature of their interactions) lose noticeably less energy in the material before calorimeter than electrons. This fact can lead to a systematic over-correction in the energy scale for photons and would yield a shift in the cross section [44].

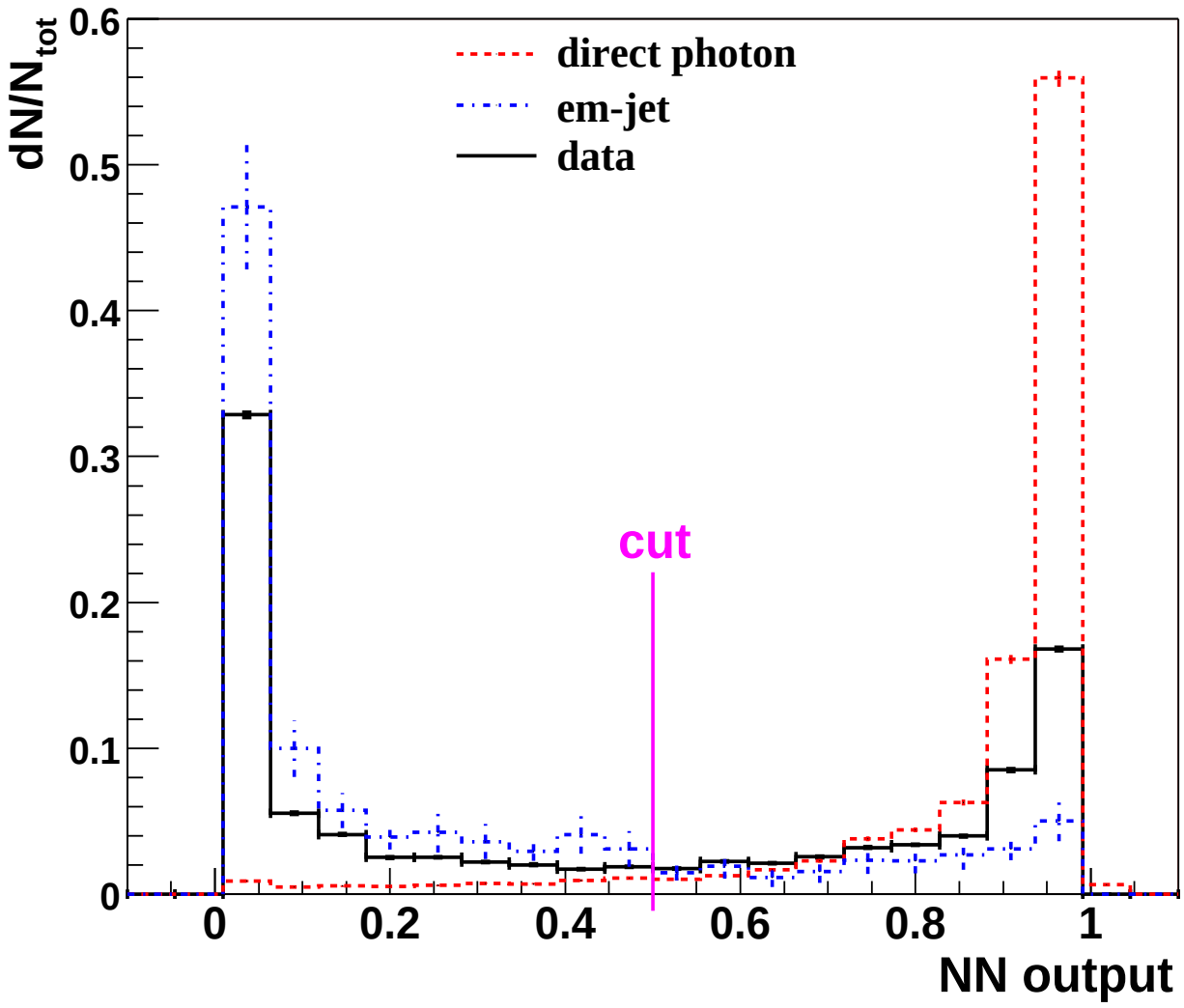

Figure 31: Normalized distribution of ANN output for data, signal and background events from $44 < p_T^\gamma < 50$ GeV after application of the main selection criteria: EMfrac ≥ 0.95, $Iso(\Delta R02) \leq 0.10$, track_match_spatialchi2prob() ≤ 0.001 and $E_T^{miss}/p_T^\gamma < 0.70$. Cut position is shown.

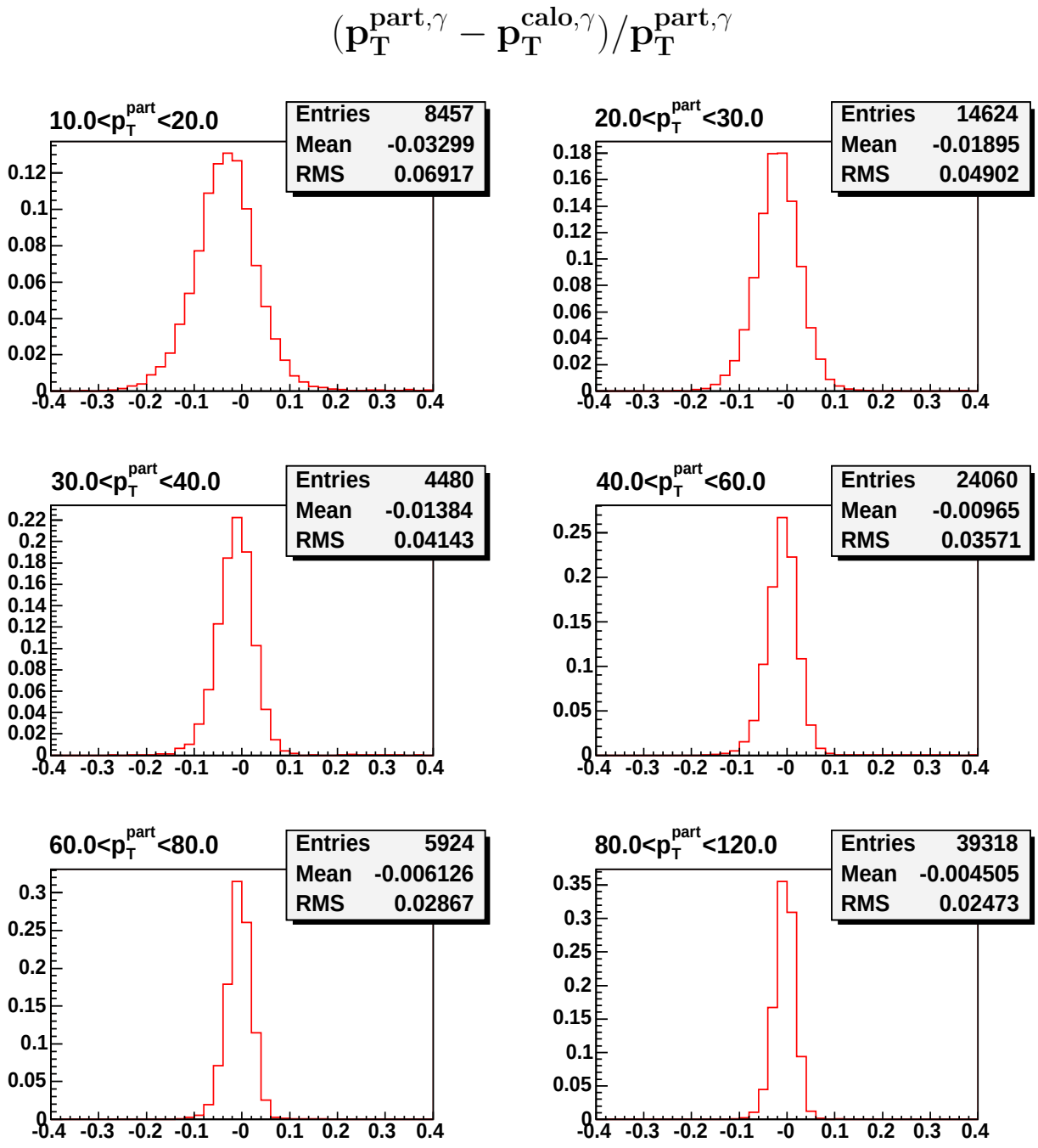

Figure 32: Distributions of the number of events over the relative shift between photon transverse momenta at the particle $p_T^{part,\gamma}$ and calorimeter (EM cluster) $p_T^{calo,\gamma}$ levels (see section 4.3).

Since for the Monte Carlo (MC) simulation the EM-calorimeter layer weights were also (as in the data) determined from the $Z \to ee$ events, we assume that the difference (shift) we would find between true and reconstructed photon p_T can be applied to correct photon p_T in the data [38]. To estimate such a shift, we have considered the "$\gamma^{dir} + jet$" events fully simulated in MC. Then for each event we have taken EM cluster found by the Simple Cone algorithm within $R < 0.20$ from photon on the particle level. Here R is distance in eta-phi space from the photon (particle level) to gravity center of the EM cluster. Distributions of the number of events over the relative shift between photon transverse momenta at the particle $p_T^{part,\gamma}$ and calorimeter $p_T^{calo,\gamma}$ levels, i.e. $(p_T^{part,\gamma} - p_T^{calo,\gamma})/p_T^{part,\gamma}$, is presented in Fig. 32. The $p_T^{calo,\gamma}$ distribution is built after application of cuts EMfrac ≥ 0.95, $Iso(\Delta R02) \leq 0.10$ and track_match_spatialchi2prob() $\leq$ 0.001.

Average shifts of reconstructed p_T of photon EM cluster with respect to partcle p_T as a function of $(p_T^{reco}$ is shown in Fig. 33.

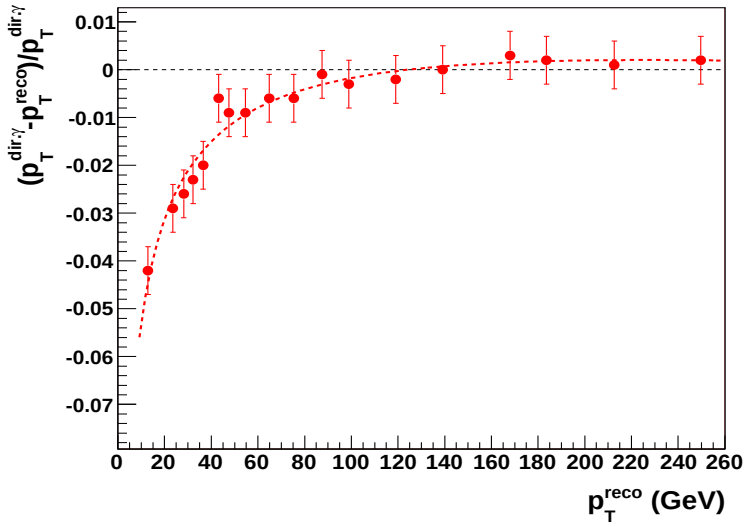

Figure 33: Average shifts of reconstructed p_T of photon EM cluster with respect to particle p_T.

The full size of the correction (absolute value of the difference $(p_T^{part,\gamma} - p_T^{calo,\gamma})$) is taken as a systematic error. On top of it, an additional error of 0.5% on electron energy scale is added in quadrature. One needs also take into account systematic error caused by accuracy of determination of background contribution after the selection cuts (this contribution shifts p_T of photon candidate to less values). As stated above, this error is $\sim 0.1\%$ for the photon purity P varied within 15%.

For the final aim it is necessary to estimate relative error to the photon cross section induced by the photon energy scale systematic error. For this aim two ansatzes, (16) and (17), with parameters (19) from Appendix B are taken. Then the relative error to the cross section, for example, for the first ansatz is equal to

$$|\delta\sigma/\sigma| = \left(par[1]/p_T^{\gamma} + 2par[2]/(\sqrt{s} - 2p_T^{\gamma}) \right) \delta p_T^{\gamma}. \quad (9)$$

The found contribution (in %) to the cross section systematic uncertainty (as an average error for

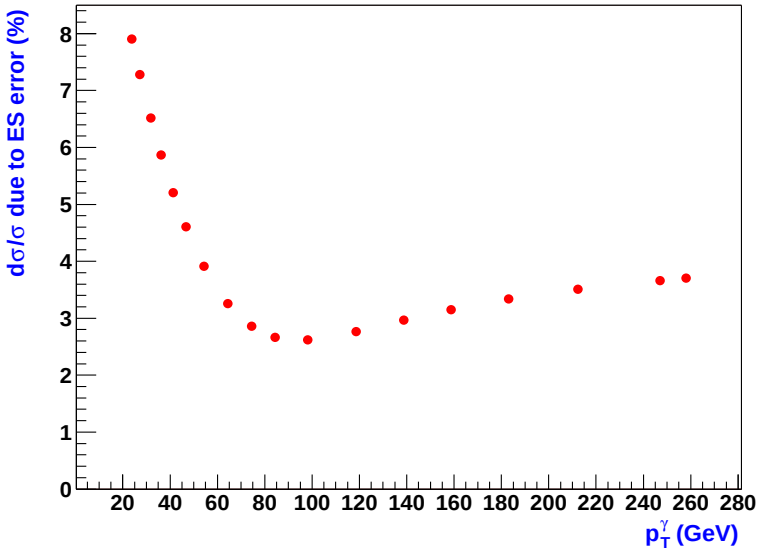

Figure 34: Dependence of the inclusive photon cross section systematic error due to the photon energy scale uncertainty on p_T^γ (see eq.(9)).

the two ansatzes) is shown in Fig. 34 and in summary Table 7 (column " $C_{p_T^\gamma}$ ") [15].

4.4 Photon purity estimation.

As we know, the measured signal from the isolated direct photon is contaminated with a background stemming from QCD jets that have fluctuated into the well-isolated single EM cluster. The cluster is caused mainly by energetic (single or multiple) π^0, η, K_s^0 or ω mesons decaying into photons in the final state [37]. Sometimes these background particles are accompanying by soft hadrons whose energy is mostly deposited in the electromagnetic shower developing within the EM cluster.

To reveal such a background we need the physical variables sensitive to the internal structure of the shower. Additionally, the distributions of these variables in MC events should be very close to those in data. Many possibilities were studied in search of the optimal set [29, 30]. As was described in the previous subsection, we have chosen four variables, three of which ($ncell\_EM1(EMclust)$, $sum\_trackPT$ $(\Delta R04)$ and $sigrphi\_EM3$) are sensitive to the energy distribution inside EM cluster and one ($ncell\_EM1$ $(\Delta R02)$) is an additional variable that characterize the isolation of the EM cluster [16].

As we mentioned in section 4.2.4, the output from artificial neural network, based on these four

[15] Additional error to $d\sigma/\sigma$ is caused by variation of parameters (19) to the ansatz (16) with their errors. It is found to be just $\sim 0.1\%$ what is negligible as compared with $d\sigma/\sigma$ from Table 7.

[16] An additional good variable used for photon discrimination in Run I was the fraction of energy deposited on the first EM layer. Unfortunately in Run II its distribution in MC does not reproduce the data.

input variables, has been chosen as a discriminant between signal and background.

Since the signal events cannot be identified on an event by event basis their fraction (purity) P is determined for a given p_T bin statistically. The photon purity is defined as the ratio

$$P = \frac{N^{\gamma}}{N^{\gamma} + N^{jet}} \tag{10}$$

where N^{γ} (N^{jet}) is the number of single photons (em-jets) that passed selection criteria (section 4.2.4).

To determine purity we have used a statistical/probabilistic method. The ANN output in data is fitted by ANN outputs from MC photon and em-jet samples using HMCMLL routine [39] (from HBOOK package). This routine correctly incorporates statistical errors in MC and data histograms and was specially written for fitting MC fractions to data histogram [17]. With this technique the purities for the 17 considered p_T^{γ} bins were found.

Explanatory plots to the determination of photon fractions (purities) for all considered intervals of are presented in Figs. 35–37. The errors shown on the plots are statistical. Here MC histograms are corrected according to the their fraction found from the HMCMLL fit. To check a compatibility of data histograms and total (i.e. sum of photon and em-jet) MC histograms for ANN output χ^2 test was done. The found values of χ^2/ndf are shown in Table 2 [18].

The photon purities found for all p_T^{γ} intervals are shown in Table 3 (see also Figs. 38 and 39). The presented errors correspond to 68% confidence level for the two parameter fit (these parameters are the fitted values of signal and background fractions in the data) [39].

Table 2: χ^2 test on compatibility of total (photon+em-jet) MC and data histograms (Figs. 35–37) for ANN output.

p_T^{γ}	23–25	25–30	30–34	34–39	39–44	44–50	50–60	60–70	70–80
χ^2/ndf	1.61	0.76	0.84	0.37	1.07	0.27	1.01	1.11	1.89

p_T^{γ}	80–90	90–110	110–130	130–150	150–170	170–200	200–230	230–300
χ^2/ndf	1.62	1.22	0.86	1.38	0.53	0.33	0.62	1.30

At high p_T intervals the uncertainty of the found purity points is mostly caused by data statistics while for lower p_T it is mostly caused by the statistics of the em-jets sample which remains after the selection cuts (section 4.2.4).

It is interesting to look at the distributions of the four variables used as the neural network input vector (see sections 3, 4.2.4) in data and in MC signal and background sets after application of

[17] It was also applied for finding photon purities in Run I.

[18] Just statistical errors in MC photon, em-jet and data samples are taken into account in the calculation of χ^2 (i.e. errors caused by accuracy of purity determination are ignored in this table).

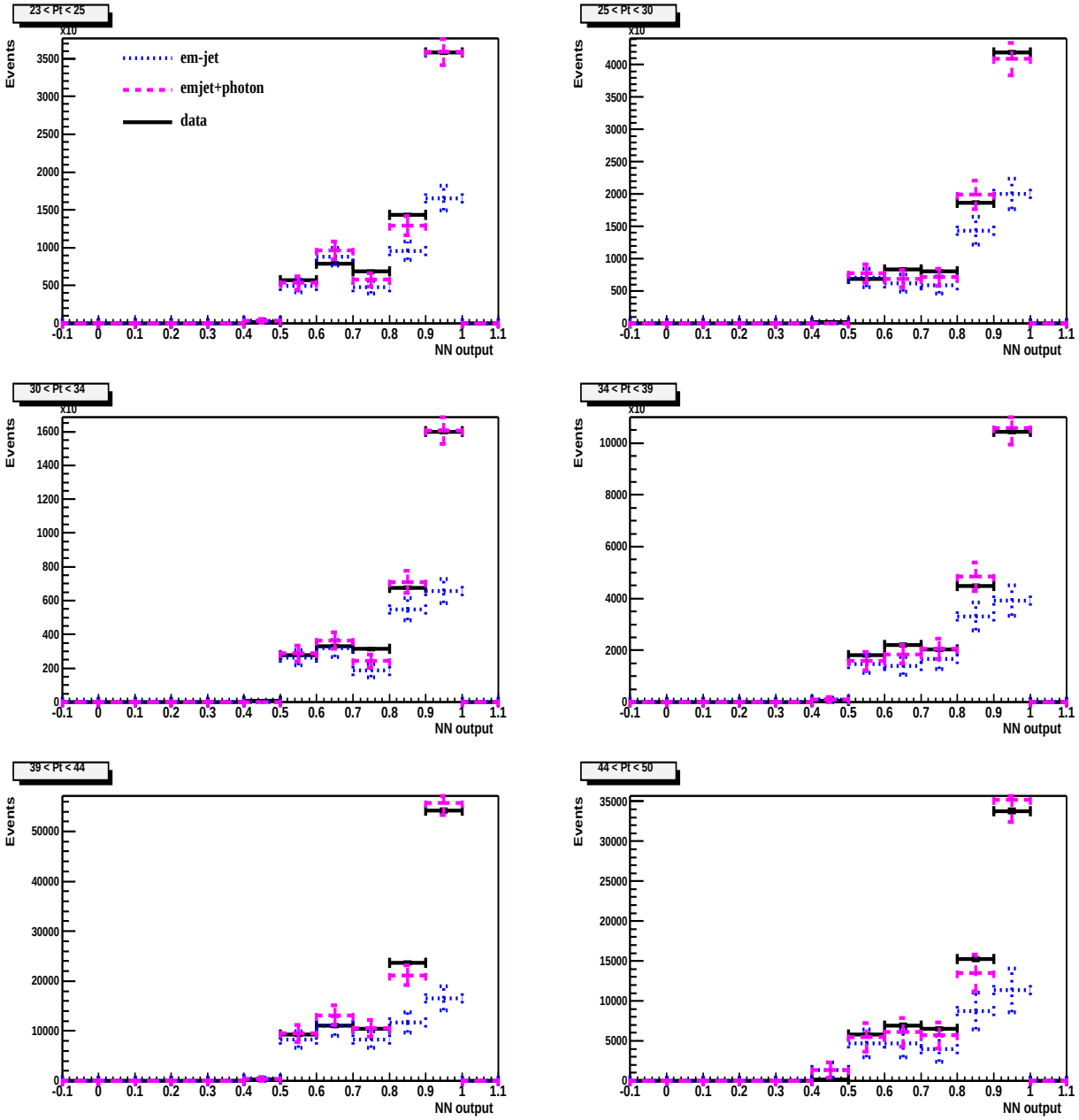

Figure 35: Distribution of the number of events in data over NN output after the cut "NN output > 0.5" for six p_T^γ intervals from 23 to 50 GeV. The fitted (to the data) distributions of the MC photons and jets are also shown. They are weighted with account of their fractions found from the HMCMLL fit. The found purity for these intervals are shown in Table 3.

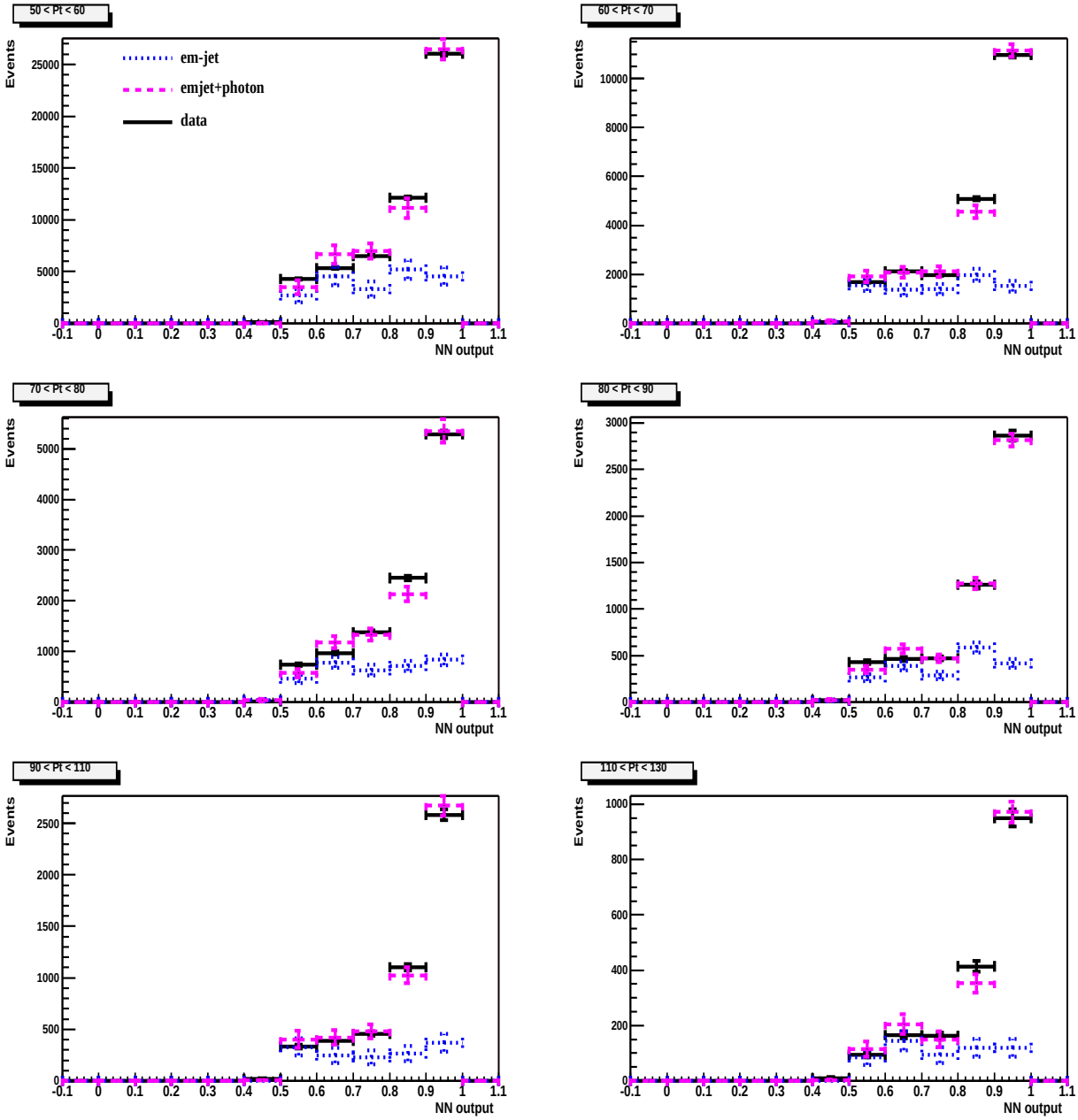

Figure 36: Distribution of the number of events in data over NN output after the cut "NN output > 0.5" for six p_T^γ intervals from 50 to 130 GeV. The fitted (to the data) distributions of the MC photons and jets are also shown. They are weighted with account of their fractions found from the HMCMLL fit. The found purity for these interval are shown in Table 3.

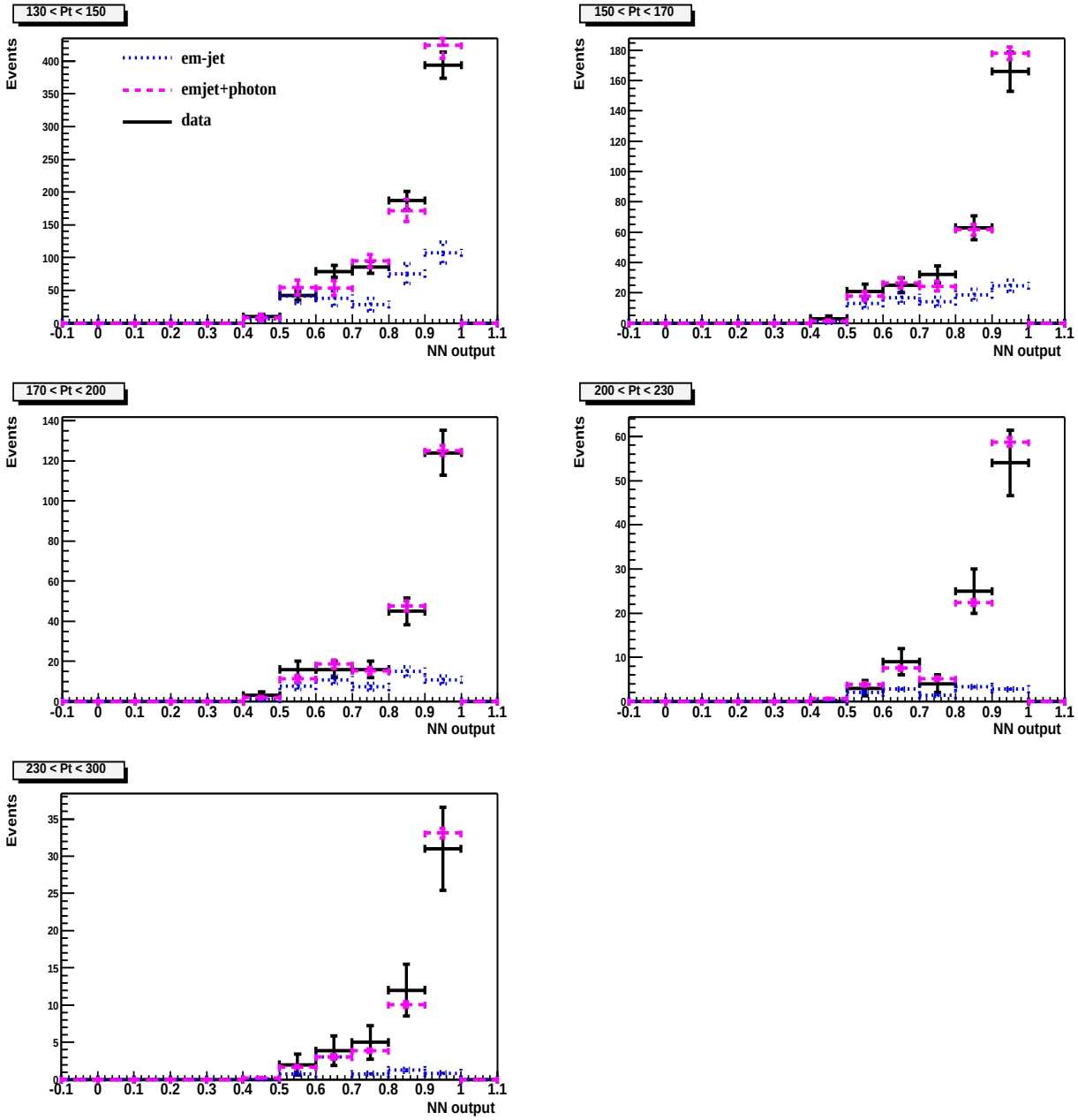

Figure 37: Distribution of the number of events in data over NN output after the cut ”NN output > 0.5” for five p_T^γ intervals from 130 to 300 GeV. The fitted (to the data) distributions of the MC photons and jets are also shown. They are weighted with account of their fractions found from the HMCMLL fit. The found purity for these interval are shown in Table 3.

Table 3: Photon purities (P) found from HMCMLL fit.

Δp_T^γ	23–25	25–30	30–34	34–39	39–44	44–50
P	0.357±0.062	0.346±0.067	0.384±0.064	0.436±0.076	0.484±0.058	0.512±0.053

Δp_T^γ	50–60	60–70	70–80	80–90	90–110	110–130
P	0.607±0.049	0.640±0.043	0.647±0.058	0.645±0.040	0.703±0.056	0.678±0.068

Δp_T^γ	130–150	150–170	170–200	200–230	230–300
P	0.685±0.089	0.727±0.101	0.761±0.110	0.847±0.172	0.901±0.228

cut on the ANN output $NN_{output} > 0.5$. They are presented in Appendix A for most of p_T^γ intervals. As in Figs. 35–37 MC distributions are weighted to account for their fractions found from the HMCMLL fit. The MC errors here include also errors on purity determination. The found purity for a given interval is shown in Table 3.

The determined from HMCMLL photon fractions were fitted [8] by the function P_f with two free parameters:

$$P_f = \frac{1}{1 + a_1\ (p_T^\gamma)^{a_2}} \tag{11}$$

We have chosen this form because we expect the data to be a sum of two falling cross sections (photons and jets) with their ratio having roughly the form $a_1(p_T)^{a_2}$ (compare with formula (10)). The fit, shown in Fig. 38 with its statistical uncertainty, assures that purity is a smooth function of p_T^γ. Here statistical uncertainties were found from:

$$(\delta P_f^{stat})^2 = \sum_{i,j=1}^{2} \left(\frac{\partial P_f}{\partial a_i} \frac{\partial P_f}{\partial a_j} \right) V_{ij} \tag{12}$$

where V_{ij} is an element of error (covariance) matrix. We see from Fig. 38 that the fitted function is smooth and does not exhibit the statistical fluctuations inherent in the data.

In order to estimate the systematic uncertainty from the choice of fitting function, we use two alternative functions. They are

$$P_f = 1 - e^{-(a_1 + a_2\ p_T^\gamma)} \tag{13}$$

and

$$P_f = a_1 + a_2\ log(p_T^\gamma) \tag{14}$$

The parameters of all fits as well as corresponding χ^2/ndf are accumulated in Table 4.

The systematic uncertainty is estimated by calculating the rms value of the different fits from the default choice (11). The complete error matrix for this systematic uncertainty is calculated using the definition of covariance, so that each element of the matrix is:

$$V'_{ij} = \frac{1}{N-1} \sum_{k=1}^{N} (x_k(i) - \mu(i))\ (x_k(j) - \mu(j)) \tag{15}$$

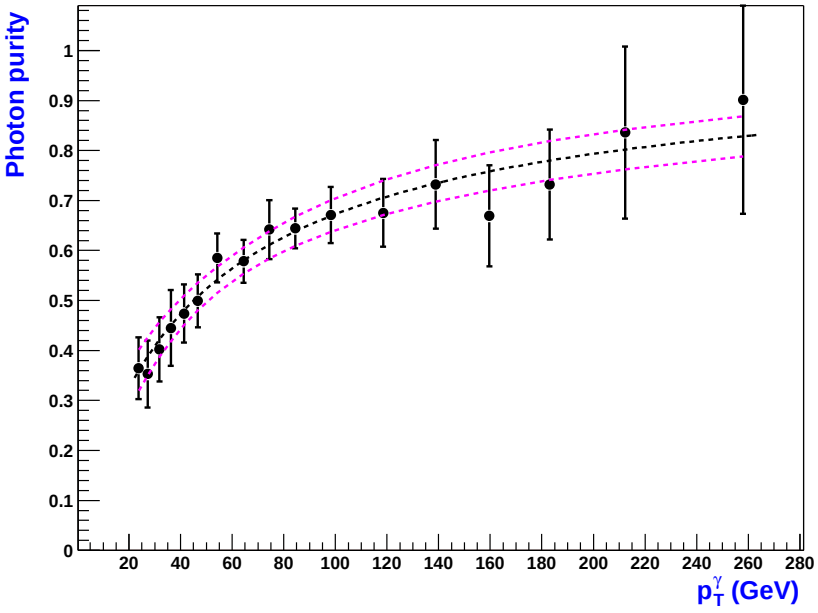

Figure 38: The photon purity vs. p_T^γ with the fitting curve (full line) and statistical uncertainty of the fit (dashed lines).

Table 4: Results of photon purity fit using different functions.

Parameters	Fitting functions		
	$1/(1+a_1\ (p_T^\gamma)^{a_2})$	$1-exp[-(a_1+a_2\ p_T^\gamma)]$	$a_1+a_2\ log(p_T^\gamma)$
a_1	38.86±16.71	0.285±0.074	-0.331±0.984
a_2	-0.971±0.112	0.009±0.006	0.221±0.028
χ^2/ndf	0.42	0.76	0.51

where i and j are different p_T^γ bins, the sum is over the various fitting functions, N is the number of different functions used to fit the data, and μ is the value of the default fit. The diagonal elements of this matrix is just the rms, i.e. we have $(\delta P_f^{syst1})^2 = V'_{ii}$.

Another source of systematic error is due to the choice of the number of bins in HMCMLL fit (see Figs. 35–37 for ANN output). It was varied from 6 to 14. Additional error due to such a variation was found to be $\delta P_f^{syst2} = 3\%$.

In Fig.39 we plot the systematic band in uncertainty caused by usage of alternative fitting functions and variation of the number of bins in HMCMLL fit as well as the default fit and its statistical error.

The total uncertainty caused by purity determination (shown by dash-dotted line in Fig. 39) is presented in column “Purity” of Table 7.

An additional systematic uncertainty was assigned due to the fragmentation model used in PYTHIA [23]. Firstly, one needs to note that PYTHIA generally well describe production of such parti-

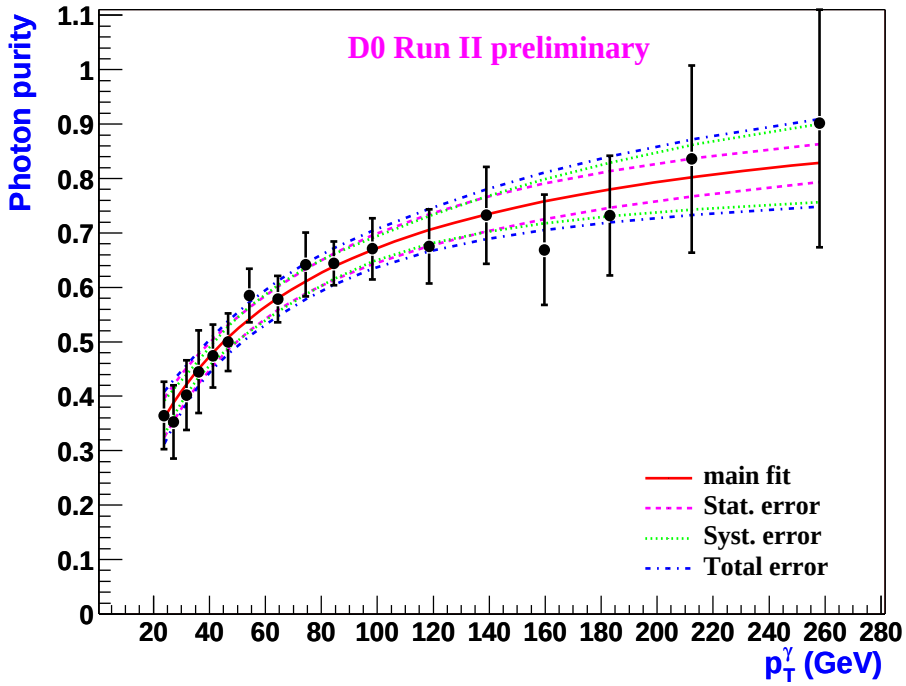

Figure 39: Statistical error from the default fit (dashed line), band in systematic uncertainty from three possible fitting functions and variation of the number of bins in HMCMLL fit (dotted line) and total error (dash-dotted line).

cles as π^0, η mesons in jets. Fig. 40 demonstrates the η/π^0 ratios measured in different experiments compared with PYTHIA predictions vs. x_T. χ^2 test on the point compatibility gives $\chi^2/ndf = 0.77$.

But the problem here is that the fragmentation functions $D_{j\to\pi^0}(z, M_f)$ (or $j \to \eta$) [19] currently used in MC event generators are fitted to LEP data ($e^+e^- \to$ hadrons) corresponding essentially to $z \leq 0.7$ (i.e. to the non-isolated particles) while for the tight isolation cuts, used here, mostly contributes the behavior of fragmentation functions near the end point, $z \to 1$. In addition, the gluon-to-pion fragmentation function is poorly constrained at moderate and large z by e^+e^- annihilation process because the subprocesses involving gluons in the e^+e^- annihilation appear at higher orders in α_s, while subprocesses involving outgoing gluons in hadronic collisions contribute at lowest order. Although the contributions from gluon fragmentation are not expected to dominate, large uncertainties on their magnitude may still affect quantitative predictions [41]. According to [40] and also taking into account noticeable discrepancy between [42] and [43] sets of parton-to-pion fragmentation functions at $z \gtrsim 0.85-0.90$ [41] the uncertainty in π^0 (as well as in η) production was taken $\simeq 50\%$ for this z region.

Thus, an additional systematic uncertainty to the photon purity was estimated by varying the ratios between π^0 and $\{\eta, \omega, K_s^0\}$ (containing γ/π^0 in the decay channel) and between η and

[19] Here z is a fraction of the final state parton momentum carried out by π^0 (η).

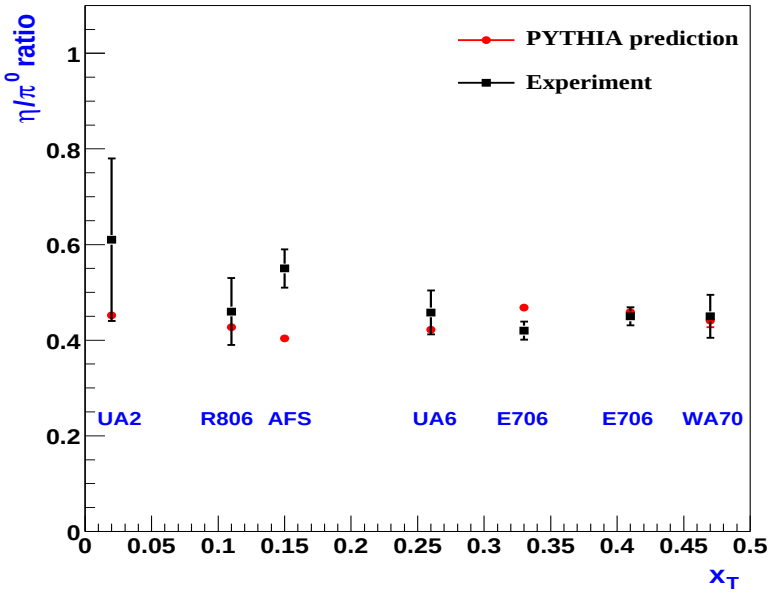

Figure 40: Ratios of η to π^0 production rates (mesons are produced in jets) vs. x_T measured in seven experiments compared with PYTHIA 6.2 predictions. χ^2 test on the point compatibility gives $\chi^2/ndf = 0.77$.

$\{\pi^0, \omega, K_s^0\}$ by $\pm 50\%$ [20]. The ratios of the nominal default purity to the purities obtained after such variations in the first and second cases are shown in Figs. 41 and 42. The relative errors shown here are caused by purity determination procedure (see Fig. 39).

We have defined final uncertainty (as a function of p_T^γ) as the average value of absolute variations of the purity caused by $\pm 50\%$ variations of the two ratios. This uncertainty can be parametrized by $p_0\ exp(p_1 \cdot p_T^\gamma)$ with $p_0 = 0.201$ and $p_0 = -0.0428$ and is shown for all p_T^γ bins in separate column "Frag" of Table 7. As we see that those variations mostly influence p_T^γ region up to ~ 50 GeV and leads to the additional $\sim 7\%$ systematic error in purity at $p_T^\gamma \simeq 25$ GeV, $\sim 4.5\%$ at $p_T^\gamma \simeq 35$ GeV, $\sim 2\%$ at $p_T^\gamma \simeq 50$ GeV and $\sim 1\%$ at $p_T^\gamma \gtrsim 70$ GeV (not included in Fig. 39).

Test on the stability of the photon cross section with respect to the final cut on ANN output can be found in Appendix C. From the ratios at Figure 62 we can estimate the size of cross section variation due to choice of ANN cut within 0.3–0.7. It was found to be equal 6.1% and included to the total systematic error (see column NN_{cut} of Table 7).

4.5 Unsmearing correction.

Unsmearing is correcting cross section due to finite resolution of the calorimeter. It is especially important for the case of steeply falling spectrum. In Run I DØ calorimeter had good enough energy resolution for EM objects in order to neglect the unsmearing effect. In Run II energy

[20] Such a variation became possible due to separate generations of π^0 and η, ω, K_s^0 em-jets (see section 4.1 and [29]).

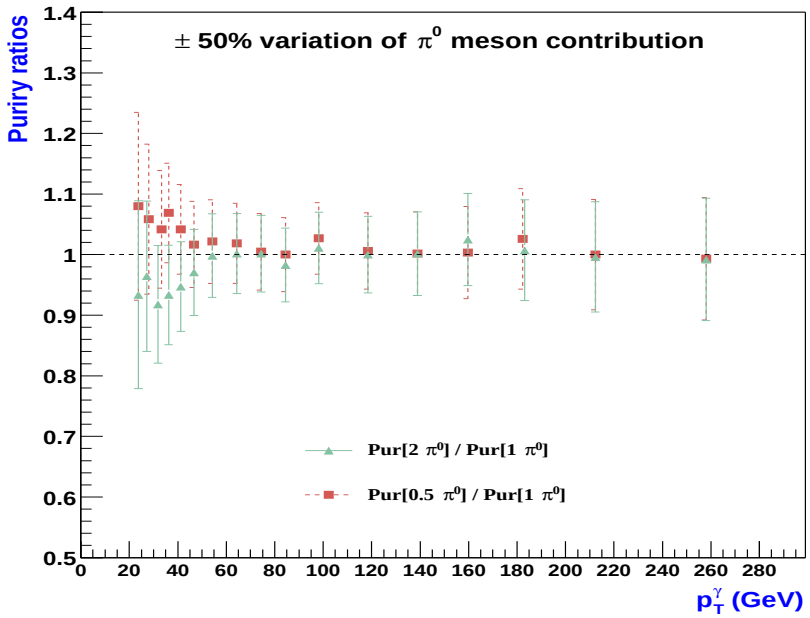

Figure 41: Ratio of the nominal default photon purity to the purity obtained after $\pm 50\%$ variation of ratios between π^0 and η, ω, K_s^0 mesons as a function of p_T^γ.

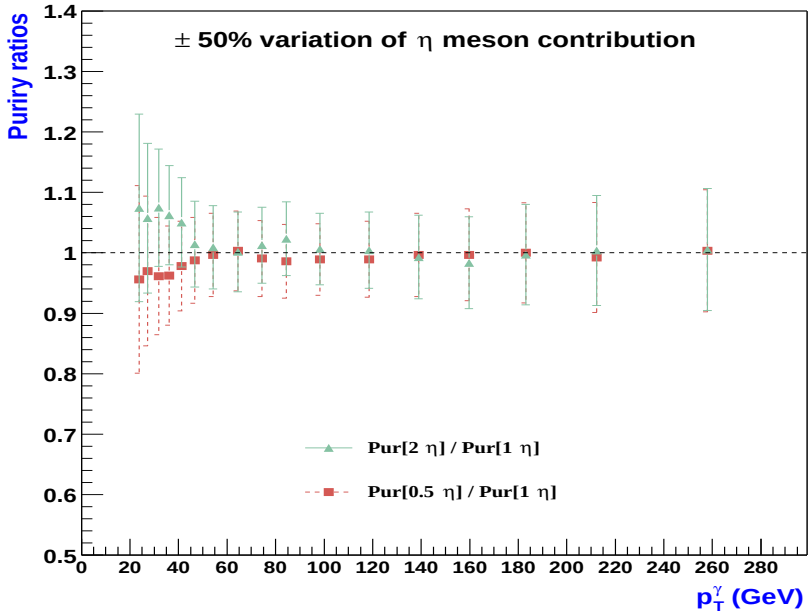

Figure 42: Ratio of the nominal default photon purity to the purity obtained after $\pm 50\%$ variation of ratios between η and π^0, ω, K_s^0 mesons as a function of p_T^γ.

resolution for EM objects has degraded and one needs to unfold the observed spectrum to recover initial (unsmeared) one.The details on the determination of unsmearing factor f_{unsm} from equation (1) can be found in Appendix B and are applied in section 5.

5 Presentation of results and comparison with theory.

Here we combine acceptance, trigger, selection efficiencies, purities and unsmearing factor found in sections 4.2.1, 4.2.2, 4.2.4, 4.4 and 4.5 and evaluate the cross section using formula (1). In Table 5 we show number of photon candidates remaining after step-by-step application of the selection criteria. The "Initial" criteria shows the total number of EM clusters in the events preselected by Common Sample Group that satisfy the primary vertex requirements and have $p_T > 23$ GeV (see section 3).

Table 5: Number of γ candidates (in thousands) after each cut.

Cut	Number of γ candidates	% from Initial
Initial (CSG+vertex cuts +$p_T > 23$ GeV	79 908	100
$\|\eta\| < 0.9$	41 837	52
EM fiducial	36 699	46
EMfrac> 0.95	21 461	27
$Iso(\Delta R02) < 0.10$	13 247	17
No matched tracks	5 710	7
$E_T^{miss}/p_T^{\gamma} < 0.70$	5 430	7
NN$_{output} > 0.5$	2 732	3.4

We see from Table 5 that after the final selection on the ANN output about 2.7 million photon candidates remained. These events are used to calculate the cross section in 17 p_T bins with average values varied from 23.9 to 258.0 GeV.

The results of the measurement are shown in Fig. 43 as a function of p_T^{γ} for the central pseudo-rapidity region $|\eta^{\gamma}| < 0.9$ with the full experimental (systematic $\oplus$ statistical) errors. The data are plotted at the p_T^{γ}-weighted average of the fit function for each bin.

One can see that in the presented range $23.9 < p_T^{\gamma} < 258.0$ GeV the central photon cross section falls by about 5 orders of magnitude. Statistical errors vary from 0.1% in the first p_T^{γ} bin to 13.2% in the last bin while systematic errors are within $12 - 23$%.

The superimposed theoretical curve corresponds to the QCD NLO predictions based on the JETPHOX program [19, 47] (red full line) with CTEQ6.1M set of parton distribution functions (PDF). The obtained results were also compared with the QCD NLO predictions [20] based on the small-cone approximation and with the same PDF. The two predictions are based on different sets of fragmentation functions, which are [52] in the first case and [53] in the second. It is worth

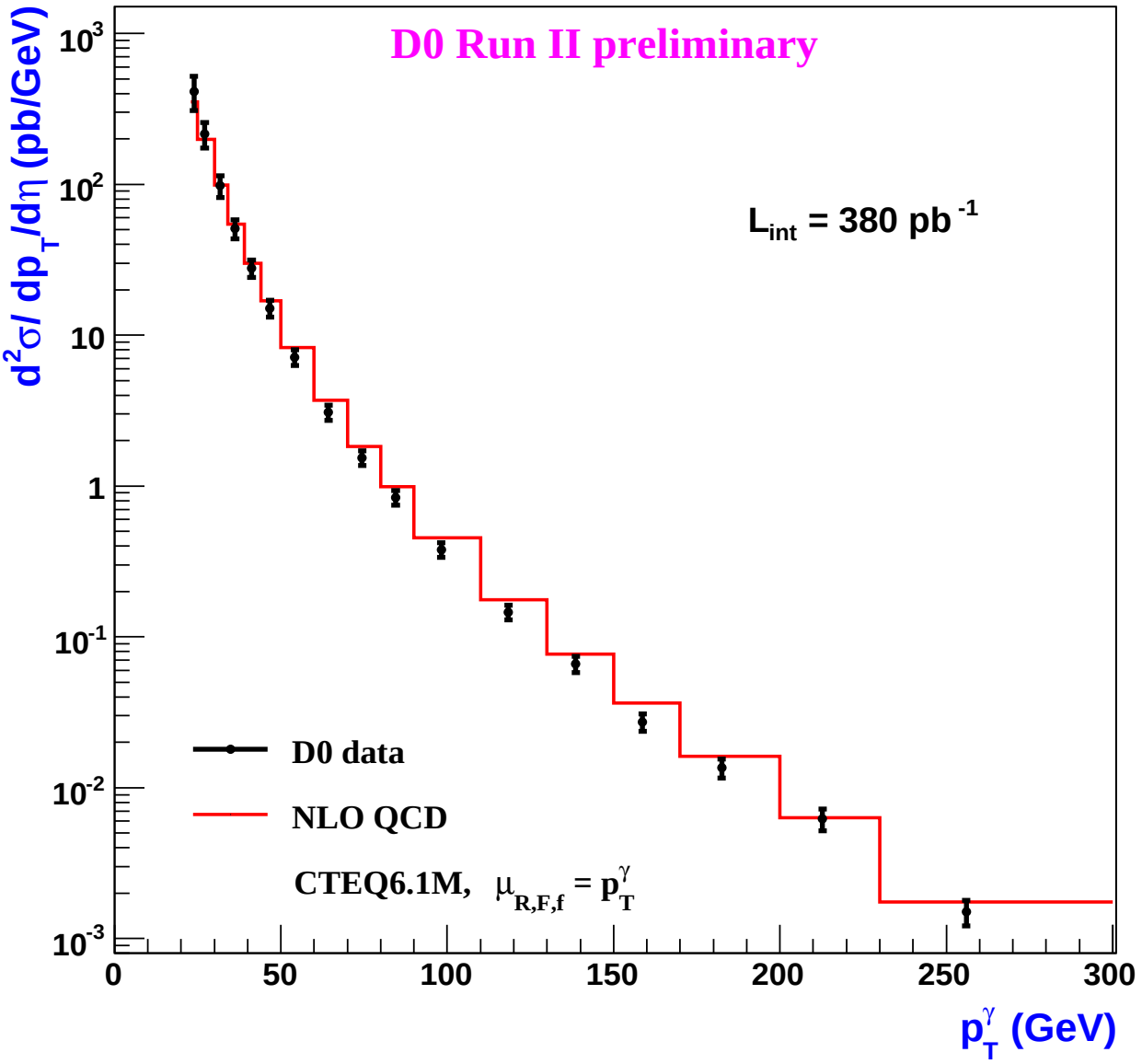

Figure 43: The isolated photon cross section $d\sigma/dp_T^\gamma d\eta$ vs. p_T^γ for the central ($|\eta| < 0.9$) rapidity region. The full (systematic $\oplus$ statistical) errors are shown. The red curve is theoretical NLO QCD predictions [19, 47] done for $\mu_R = \mu_F = p_T^\gamma$ with CTEQ6.1M PDF set. The data are plotted at the p_T^γ-weighted average of the fit function for each bin.

Table 6: Differential cross section for the central region ($|\eta| < 0.9$) with statistical and systematic uncertainties.

p_T^γ bin (GeV)	$\langle p_T^\gamma \rangle$ (GeV)	$d^2\sigma/dp_T^\gamma d\eta$ (pb/GeV)			$\delta\sigma_{stat}$ (%)	$\delta\sigma_{syst}$ (%)	$\delta\sigma_{tot}$ (%)
		theor. (CFGP)	theor. (GV)	measured			
23 – 25	24.1	354	353	419	0.1	23.4	23.4
25 – 30	27.2	200	198	222	0.1	18.9	18.9
30 – 34	31.8	102	98.6	100	0.2	16.2	16.2
34 – 39	36.1	53.6	54.2	53.0	0.2	14.8	14.8
39 – 44	41.2	31	30	28.5	0.3	13.8	13.8
44 – 50	46.7	17.4	16.9	15.1	0.4	13.1	13.1
50 – 60	54.2	8.5	8.24	7.38	0.4	12.6	12.6
60 – 70	64.3	3.75	3.69	3.14	0.6	12.2	12.2
70 – 80	74.4	1.85	1.83	1.54	0.9	11.9	12.0
80 – 90	84.4	0.996	0.992	0.837	1.3	11.8	11.9
90 – 110	98.2	0.454	0.455	0.391	1.4	11.7	11.8
110 – 130	118.4	0.179	0.176	0.148	2.3	11.9	12.1
130 – 150	138.9	0.0764	0.0769	0.0676	3.5	12.3	12.8
150 – 170	158.7	0.0374	0.0366	0.028	5.6	12.7	13.9
170 – 200	183.1	0.0157	0.0161	0.0143	6.5	13.6	15.1
200 – 230	212.3	0.00637	0.00634	0.00627	9.8	13.9	17.0
230 – 300	255.3	0.00188	0.00176	0.00154	13.2	14.7	19.7

emphasizing that in spite of such a discrepancy both predictions are in agreement within 7% what is seen from Table 6 and Fig. 44 which shows ratio of the predictions done by P. Aurenche *et al* to the predictions done by W. Vogelsang *et al* for three scales. Sensitivity of the theoretical predictions to the isolation requirements in the ring of $\Delta\mathcal{R} = 0.2$ and to the EM fraction in the cone of $\mathcal{R} = 0.2$ was tested [51]. It is worth mentioning that the isolation conditions for the theoretical predictions based on the code [19] slightly differs from those based on [20]. Just 10% isolation in the cone of $R = 0.4$ was required in first case while for theoretical results based on [20] both 10% isolation in the ring of $R = 0.2$ and 5% veto on hadronic energy in the cone of $R = 0.2$ around a photon were required.

Sensitivity of the theoretical predictions to the isolation requirements in the ring of $\Delta\mathcal{R} = 0.2$ and to the EM fraction in the cone of $\mathcal{R} = 0.2$ was tested [51]. Variation of the $Iso(\Delta\mathcal{R}02)$ cut from 10 to 5% and from 10 to 15% leads to changes in the predicted cross section of less than 2%. Variation of allowable hadronic energy in the cone of $\mathcal{R} = 0.2$ (i.e. inside of EM cluster) from 4 to 6% also leads to relative changes in the cross section $\leq$ 2% [51].

The theoretical predictions presented in Fig. 43 correspond to the choice of renormalization, factorization and fragmentation scales as $\mu_R = \mu_F = \mu_f = p_T^\gamma$. If all scales are varied to $\mu_R = \mu_F = \mu_f = 0.5p_T^\gamma$ or to $2p_T^\gamma$ the cross sections change within $\pm$12–13% (see Fig. 5).

Table 7: Systematic uncertainties (in %) due to purity, fragmentation (Frag), ANN cut choice (NN_{cut}), trigger ϵ_t and selection ϵ_s efficiencies, energy scale (ES), p_T^γ correction $C_{p_T^\gamma}$, PV choice (PV), acceptance (Acpt), luminosity (Lum) and unsmearing (Unsm).

p_T^γ bin (GeV)	Purity	Frag	NN_{cut}	ϵ_t	ϵ_s	ES	$C_{p_T^\gamma}$	PV	Acpt	Lum	Unsm	**Total syst.**
23 – 25	13.3	7.3	6.1	11.6	4.7	6.3	7.8	3.6	1.5	6.5	1.0	23.4
25 – 30	10.7	6.3	6.1	5.3	4.6	6.1	7.3	3.6	1.5	6.5	1.0	18.9
30 – 34	8.3	5.4	6.1	2.2	4.3	5.9	6.5	3.6	1.5	6.5	1.0	16.2
34 – 39	7.0	4.4	6.1	1.1	4.0	5.7	5.9	3.7	1.5	6.5	1.0	14.8
39 – 44	6.3	3.5	6.1	1.0	3.7	5.5	5.2	3.7	1.5	6.5	1.0	13.8
44 – 50	6.0	2.8	6.1	1.0	3.5	5.4	4.6	3.7	1.5	6.5	1.0	13.1
50 – 60	5.8	2.1	6.1	1.0	3.4	5.3	3.9	3.8	1.5	6.5	1.0	12.6
60 – 70	5.6	1.4	6.1	1.0	3.2	5.1	3.3	3.9	1.5	6.5	1.0	12.2
70 – 80	5.5	1.1	6.1	1.0	3.1	5.1	2.9	3.9	1.5	6.5	1.0	11.9
80 – 90	5.3	1.1	6.1	1.0	3.1	5.0	2.7	4.0	1.5	6.5	1.0	11.8
90 – 110	5.2	1.0	6.1	1.0	3.1	4.9	2.6	4.1	1.5	6.5	1.0	11.7
110 – 130	5.6	1.0	6.1	1.0	3.1	4.8	2.8	4.2	1.5	6.5	1.0	11.9
130 – 150	6.2	1.0	6.1	1.0	3.1	4.8	3.0	4.4	1.5	6.5	1.0	12.3
150 – 170	6.9	1.0	6.1	1.0	3.1	4.7	3.2	4.5	1.5	6.5	1.0	12.7
170 – 200	7.7	1.0	6.1	1.0	3.1	4.7	3.3	4.6	1.5	6.5	1.0	13.6
200 – 230	8.7	1.0	6.1	1.0	3.1	4.6	3.5	4.8	1.5	6.5	1.0	13.9
230 – 300	9.7	1.0	6.1	1.0	3.1	4.6	3.7	5.0	1.5	6.5	1.0	14.7

We have also varied separately fragmentation scale in the range of $0.1p_T^\gamma < \mu_f < p_T^\gamma$ [17, 54] leaving two other scales μ_R and μ_F equal to p_T^γ. In this case the cross section has changed by 5% at $p_T^\gamma = 23.9$ and just by 1% at $p_T^\gamma = 98.2$ GeV.

The calculated experimental cross section together with systematic and statistical errors is presented in Table 6. For comparison the theoretical numbers predicted by [19] (CFGP) and [20] (GV) are also given (columns marked by 'CFGP' and 'GV' respectively). The sources of experimental systematic uncertainties (in %) are shown in Table 7. One can see that the largest uncertainty is caused by the purity estimation (columns 2–4 of Table 7) [21)].

The ratio of theoretical predictions done with MRST2004 to ones with CTEQ6.1M are shown in Fig. 45. One can see that variations are within $6-7$% and maximal for $60 < p_T^\gamma < 140$ GeV.

The ratio of the measured cross section to the NLO QCD predictions [47] calculated with the CTEQ6.1M PDF set is presented in Fig. 5. Ratios of the nominal theory predictions [47] (with the PDF set corresponding to the best fit and all scales chosen as $\mu_{R,F,f} = p_T^\gamma$) to the predictions

[21)] Note, that the syst. errors due to the fragmentation and choice of ANN cut (see below) should be also associated with an uncertainty of the purity determination.

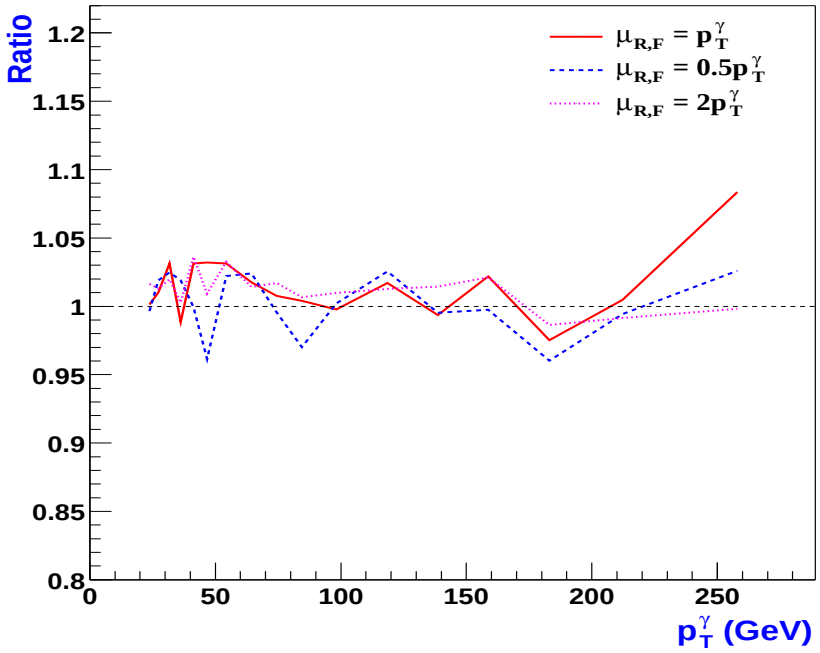

Figure 44: Ratios of theoretical predictions done by P. Aurenche *et al* to the predictions done by W. Vogelsang *et al* are shown for three scales.

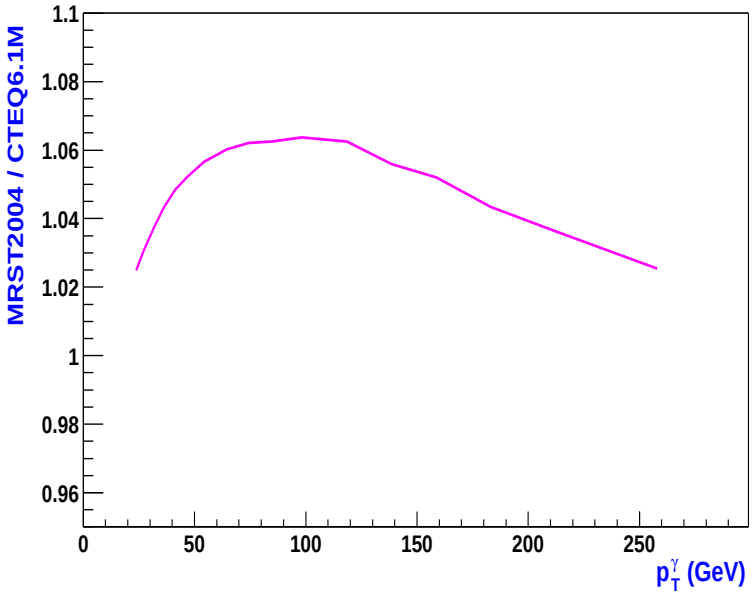

Figure 45: Ratio of theoretical predictions with MRST2004 to CTEQ6.1M.

with $\mu_{R,F,f} = 0.5p_T^\gamma$ and $2p_T^\gamma$ are shown on the plot by dotted blue curves. The uncertainties of the theoretical prediction caused by the CTEQ6.1M PDF uncertainties have been found using the 40 PDF sets [46]. These uncertainties are shown in Fig. 5 by the dash-and-dot red curves [48]. They grow from 4-5% at $p_T^\gamma \simeq 25$ GeV to 9% at $p_T^\gamma \simeq 250$ GeV (with a tendency to grow at higher p_T^γ). A similar treatment using the approach of [20] yields comparable results, specifically a 1-2% variation, with similar p_T^γ dependence.

6 Conclusion.

We have measured the inclusive cross section for isolated photons in the central pseudorapidity region ($|\eta| < 0.9$) at $\sqrt{s} = 1.96$ TeV. It was compared with two NLO QCD theoretical predictions done by [19, 47] and [20]. using as CTEQ6.1M as MRST2004 PDF sets.

The normalization of the whole data set agrees (within uncertainties) with the NLO QCD predictions but the shape vs. p_T^γ does not. The disagreement at small p_T's favours soft gluon resummation [56] and enhanced k_T-effect [9, 57] or/and modification of partonic PDF. At medium and high p_T's we observe, vice versa, by about 15-17% larger theoretical cross sections than the data ones what may require to re-consider as the QCD μ-scales or/and some PDFs at high x, Q^2. Our measurements confirm the analogous disagreement with theory observed earlier in UA2 [4] and CDF [5] experiments.

Acknowledgments

We are greatly thankful to John Womersley, Andrew Alton, Markus Wobisch for encouragements, permanent support, discussions and fruitful suggestions as well as to the members of QCD and EM ID groups, especially to Heidi Schellman, Michael Strauss, Marek Zelinski, Alexander Kupco, Yuri Maravin and Jan Stark for help and many useful discussions.

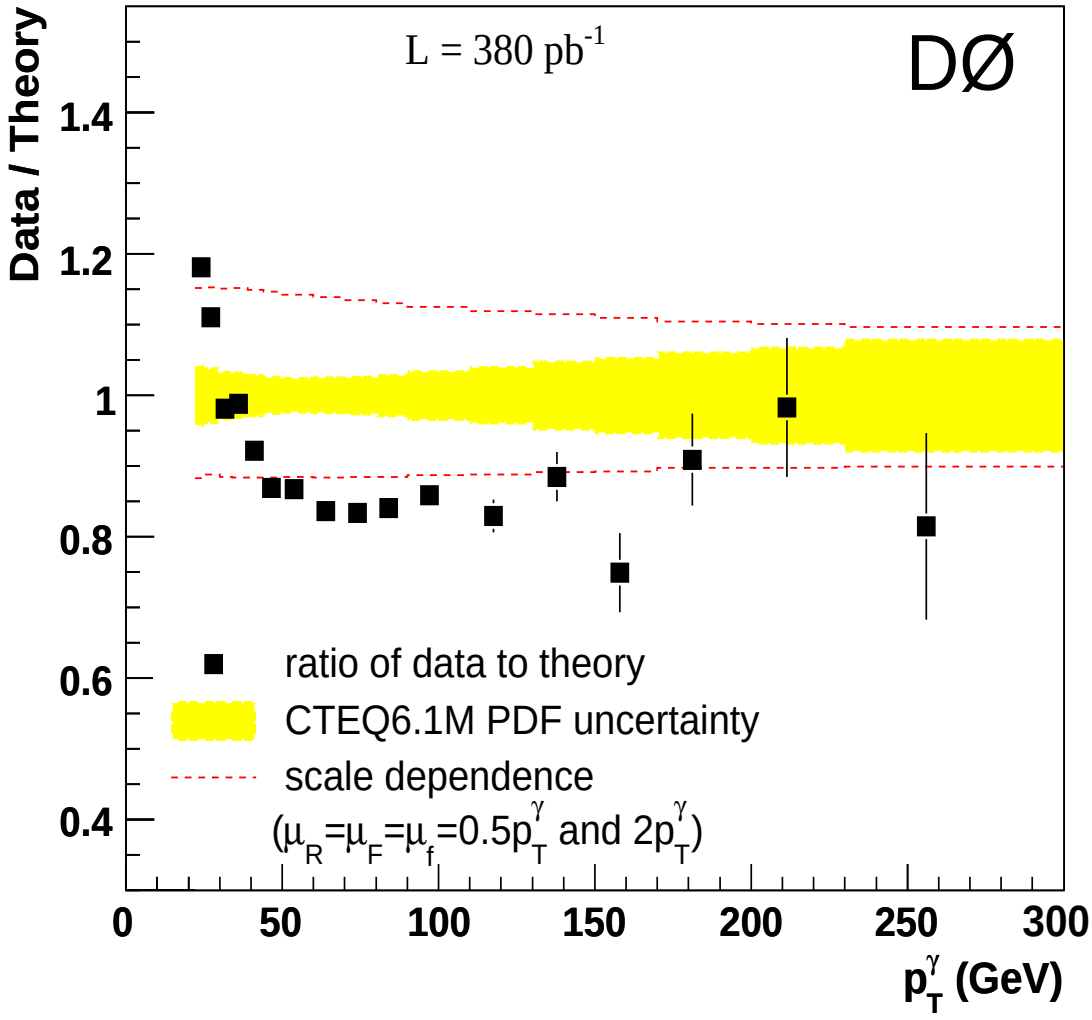

Figure 46: The ratio of the measured cross section to the NLO QCD predictions done with [47] (theor. predictions based on the code [20] are within 7%), with the CTEQ6.1M PDF set is presented. Ratios of the nominal theory predictions (with $\mu_{R,F,f} = p_T^\gamma$ and CTEQ6.1M PDF set corresponding to the best fit) to the predictions with $\mu_{R,F,f} = 0.5p_T^\gamma$ (upper dotted blue curve) and $2p_T^\gamma$ (lower curve) as well as to the predictions with the largest upper and lower CTEQ6.1M PDF errors (dash-and-dot red curves) are also shown.

References

[1] UA1 Collaboration, C. Albajar *et al.*, Phys.Lett **209B** (1998)385.

[2] UA2 Collaboration, R. Ansari *et al.*, Phys.Lett. **176B** (1986)239.

[3] CDF Collaboration. F. Abe *et al.*, Phys.Rev.Lett. **68** (1992)2734; F. Abe *et al.*, Phys.Rev. **D48** (1993)2998; F. Abe *et al.*, Phys.Rev.Lett. **73** (1994)2662; F. Abe *et al.*, Phys.Rev. **D57** (1998)1359 (see also p.67).

[4] UA2 Collaboration, J. Alitti, *et al.*, Phys.Lett. **263B** (1991)544.

[5] CDF Collaboration, D. Acosta, *et al.*, Phys.Rev. **D65**(2002)112003.

[6] DØ Collaboration, F. Abachi *et al.*, Phys.Rev.Lett. **77**(1996)5011; DØ Collaboration, B. Abbott *et al.*, Phys.Rev.Lett. **84**(2000)2786; DØ Collaboration, V. Abazov *et al.*, Phys.Rev.Lett. **87**(2001)251805

[7] S. Grinstein, R. Piegaia, D0 Note 3350; S.L. Linn, D0 Note 3564, M. Strauss D0 Note 3868.

[8] M. Strauss, M. Price, "A measurement of the isolated direct photon cross section at 630 GeV", D0 Note 3699.

[9] E706 Collaboration, L. Apanasevich *et al.*, Phys.Rev. **D70**(2004)092009, hep-ex/0407011.

[10] P. Aurenche, J. Lindfors, Nucl.Phys.**B168**(1980)296.

[11] P. Aurenche, A. Douiri, R. Baier, M. Fontannaz, D. Schiff, Phys.Lett.**B140**(1984)87.

[12] T. Ferbel and W.R. Molzon, Rev.Mod.Phys. **56** (1984)181.

[13] J.F. Owens, Rev.Mod.Phys. **59** (1987)465.

[14] E.N. Argyres, A.P. Contogouris, N. Mebarki and S. D.P. Vlassopulos, Phys.Rev. **D35**, (1987)1584.

[15] P. Aurenche, R. Baier, M. Fontannaz, J.F. Owens and M. Werlen, Phys.Rev. **D39** (1989)3275.

[16] E.L. Berger and J. Qiu, Phys.Rev. **D44** (1991)2002.

[17] W. Vogelsang and A. Vogt, Nucl.Phys. **B453** (1995)334.

[18] S. Frixione and W. Vogelsang, CERN-TH/99-247, hep-ph/9908387.

[19] P. Aurenche, R. Baier, M. Fontannaz, D. Schiff, Nucl.Phys. **B297** (1988)661;
F. Aversa, P. Chiappetta, M. Greco, J.P. Guillet, Nucl.Phys. **B327** (1989)105;
S. Catani, M. Fontannaz, J.Ph. Guillet and E. Pilon, Preprint LAPTH-907/02, JHEP 0205 (2002), 028, hep-ph/0204023.

[20] L. E. Gordon and W. Vogelsang, Phys.Rev. **D48**, 3136 (1993); L. E. Gordon and W. Vogelsang, Phys.Rev. **D50**, 1901 (1994).

[21] CDF Collaboration, Measurement of the Cross Section for Prompt Diphoton Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96TeV$, Submitted to Phys.Rev.Lett., hep-ex/0412050.

[22] T. Binoth, J.P. Guillet, E. Pilon, M. Werlen, "A full next to leading order study of direct photon pair production in hadronic collisions", Eur.Phys.J. **C16** (2000)311-330, hep-ph/9911340.

[23] T. Sjostrand, Comp.Phys.Comm. **82** (1995)74.

[24] F. Fleuret, "The D0 Electron/Photon Analysis Package EMAnalyze", D0 Note 3888. http://www-d0.fnal.gov/phys_id/emid/d0_private/emid.html.

[25] The code written by Marco Verzocchi was used for calculation of the luminosity.

[26] D0 Collaboration, S. Abachi *et al.*, Nucl.Instrum.Methods, Phys.Res., A**338** (1994)185.

[27] D. Chapin, H. Fox, J. Gardner, R. Illingworth, A. Lyon,J. Zhu, "Measurement of $Z \to e^+e^-$ and $W \to e^{\pm}\nu$ production cross section with $|\eta| < 1.05$", D0 Note (version 3.1).

[28] A. Kupco, C. Royon, M. Voulitainen, M. Wobisch "Measurement of inclusive jet p_T spectra and dijet mass spectra in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV".

[29] M. Alexeev, D. Bandurin, D. Kirillov, N. Skachkov, "On a reduction of background to the direct photon using calorimeter information", D0 Note 4305.

[30] A. Alton, A. Askew, O. Atramentov, and Y. Maravin, "Photon Identification for D Run II data", D0 Note 4487.

[31] John Gardner's program has been used for this aim.

[32] Y. Fisyak, J. Womersley, "D0 GEANT Simulation of the Total Apparatus Response." Geant-3 based simulation package of D0 detector D0GSTAR.

[33] A. Alton, A. Askew, Y. Maravin, "Analysis of $Z\gamma$ events in D0 Run II Data", D0 Note xxxx.

[34] Proc. of CERN School of Computing, 1991, Ystad, Sweden, CERN 92-02, p.113 – 170.

[35] C. Peterson, T. Rognvaldsson and L. Lonnblad, "JETNET 3.0. A versatile Artificial Neural Network Package", Lund University Preprint LU-TP 93-29. Version 3.5 is used here.

[36] C. Peterson and E. Hartman, Neural Network **2** (1989)475.

[37] Particle Data Group, D.E. Groom *et al*, The European Physical Journal C15 (2000) 1.

[38] D. Bandurin's report at D0 JES workshop on 1st of February, 2005.

[39] R. Barlow, C. Beeston, Comp.Phys.Comm. **77** (1993)219-228. See also p.120-131 of HBOOK manual.

[40] Private (email) communication with Torbjorn Sjostrand.

[41] T. Binoth, J.P. Guillet, E. Pilon, M. Werlen, Eur.Phys.J.direct **C4** (2002)7, hep-ph/0203064 (see Section 3).

[42] B.A. Kniehl, G. Kramer and B. Potter, Nucl.Phys. **B582** (2000)514.

[43] J. Bienewiess, B.A. Kniehl and G. Kramer, Phys.Rev. **D53** (1996)6110.

[44] The author is very thankful to Michael Begel for pointing to this fact as well as to Jan Stark for the following discussion on this subject.

[45] S. Jain, "Scale and oversmearing for high-p_T electrons", D0 Note 4402.

[46] J. Pumplin et al., JHEP 0207, 12 (2002); D. Stump et al., JHEP 0310, 046 (2003).

[47] P. Aurenche *et al.*, JETPHOX package, http://wwwlapp.in2p3.fr/lapth/PHOX_FAMILY/main.html. See also [19].

[48] The author is very grateful to Markus Wobisch for the help in calculation of CTEQ6.1M PDF errors.

[49] D.V. Bandurin and N.B. Skachkov, ”On the application of “direct photon+jet” process for setting the absolute scale of jet energy and for determining the gluon distribution at the Tevatron in Run II”, Phys.Part.Nucl. **35**(2004)66-106, hep-ex/0304010.

[50] D.V. Bandourin and N.B. Skachkov, ”Photon+jet event rate estimation for gluon distribution determination at the Tevatron RUN II”. Proceedings of XVI ISHEP ”Relativistic Nuclear Physics and Quantum Chromodynamics”, Dubna, June 10–15, 2002. Eds. A.N. Sissakian, V.V. Burov, A.I. Malakhov. Dubna, 2004, vol. I, p.65-75.

[51] Private (email) communication with Werner Vogelsang.

[52] L. Bourhis, M. Fontannaz, J.P. Guillet, “Quark and gluon fragmentation functions into photons”, Eur.Phys.J. **C2** (1998)529, hep-ph/9704447.

[53] M. Gluck, E. Reya and A. Vogt, “Parton fragmentation into photons beyond the leading order” Phys.Rev. **D48**, (1993)116, [Erratum-ibid. D **51**, (1995)1427].

[54] M. Gluck, L. Gordon, E. Reya and W. Vogelsang, Phys.Rev.Lett. **73**(1994)388.

[55] A.D. Martin *et al.*, Eur.Phys.J. **C4** (1998)463.

[56] E. Laenen, G. Oderda, and G. Sterman, Phys.Lett. **B438**, (1998)173.

[57] A. Lipatov and N. Zotov (2005), hep-ph/0507243.

[58] D. Bandurin, O. Atramentov, Y. Maravin, “Photon energy scale for jet energy scaling”, DØ Note 4974.

Appendix

A. Comparisons of MC-data for the input variables used in ANN.

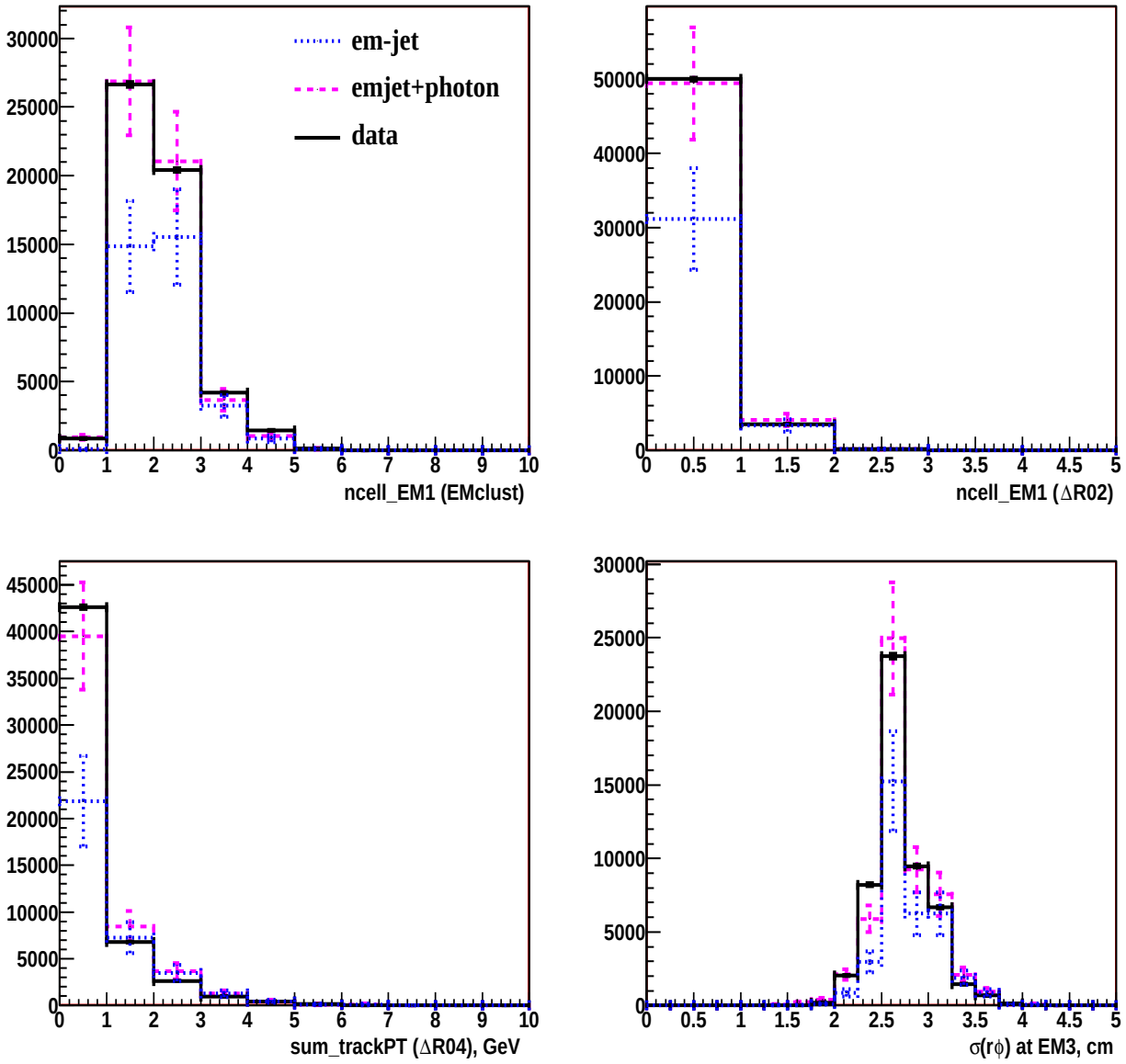

Figure 47: Distribution of the number of events in data and the fitted distributions for MC photons and jets over the four variables used as neural network input vector. MC distributions are weighted with account of their fractions found from the HMCMLL fit. Black histograms (full line) correspond to the data, blue ones (dotted line) to MC em-jets and purple ones (dashed line) to MC total 'em-jet+photons'. All distributions are built after the cut "$NN_{output} > 0.5$" for $\mathbf{23 < p_T^{\gamma} < 25}$ GeV.

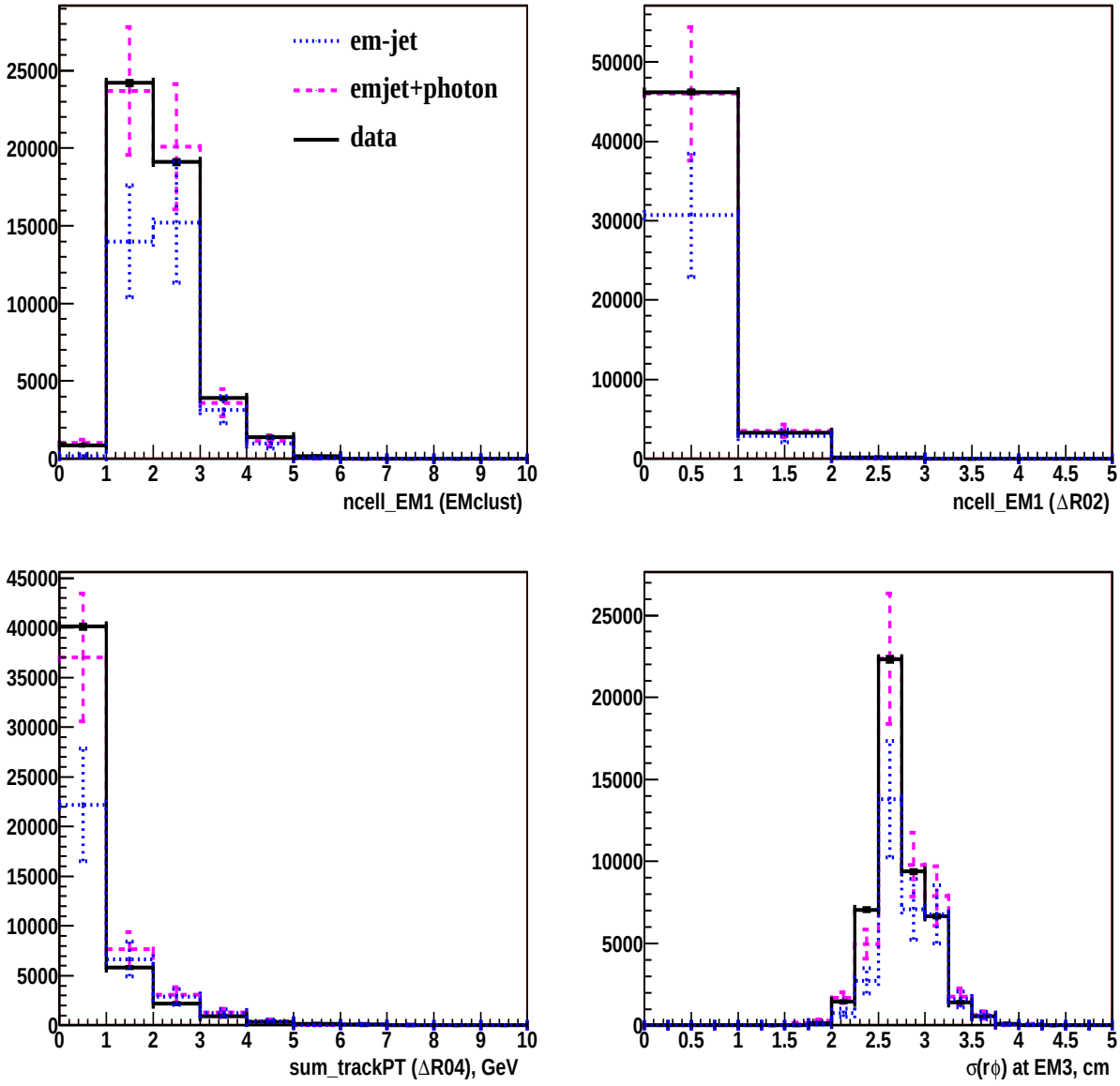

Figure 48: Same as in Fig. 47 but for $\mathbf{25 < p_T^{\gamma} < 30}$ GeV.

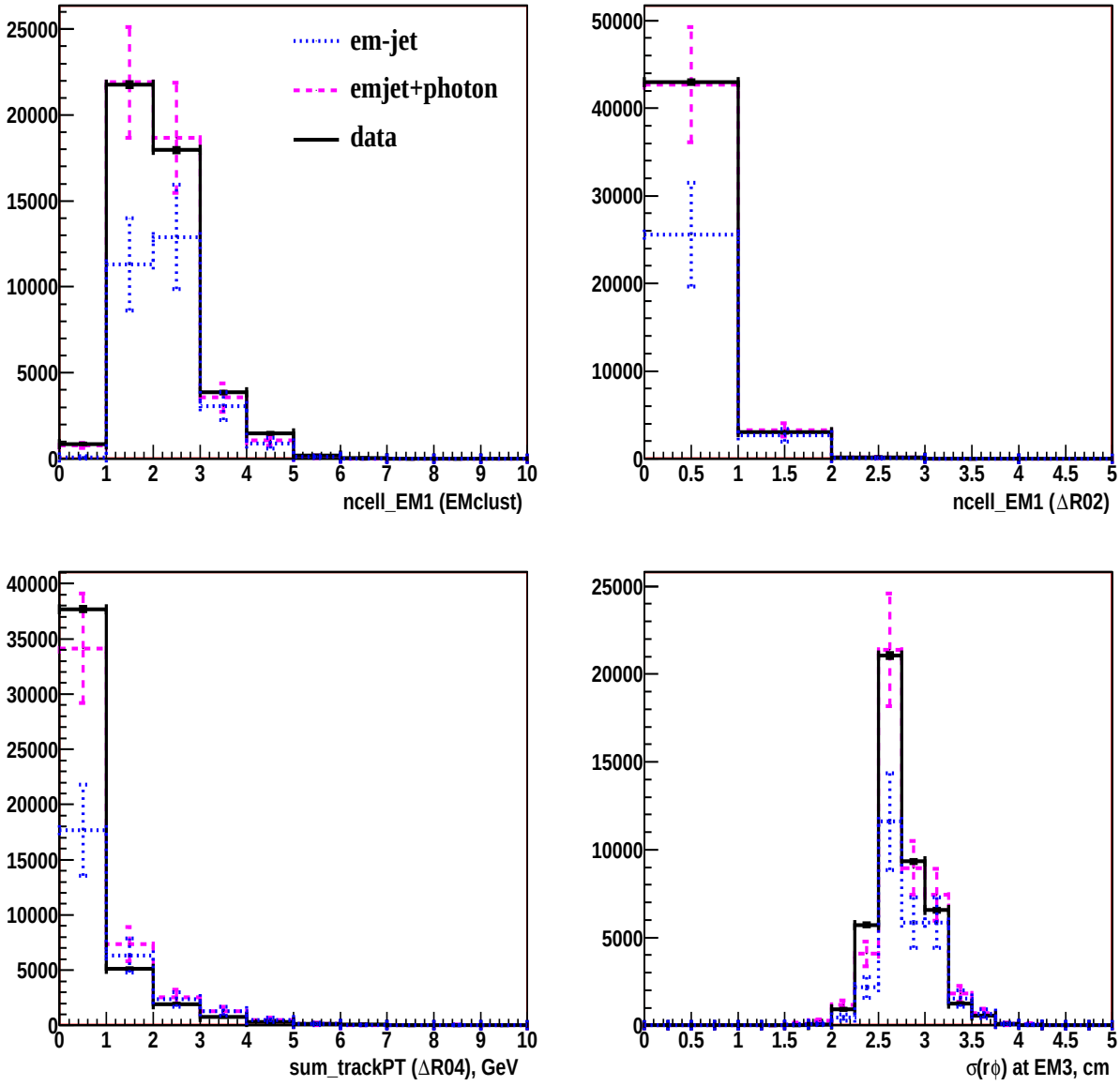

Figure 49: Same as in Fig. 47 but for $\mathbf{30 < p_T^{\gamma} < 34}$ GeV.

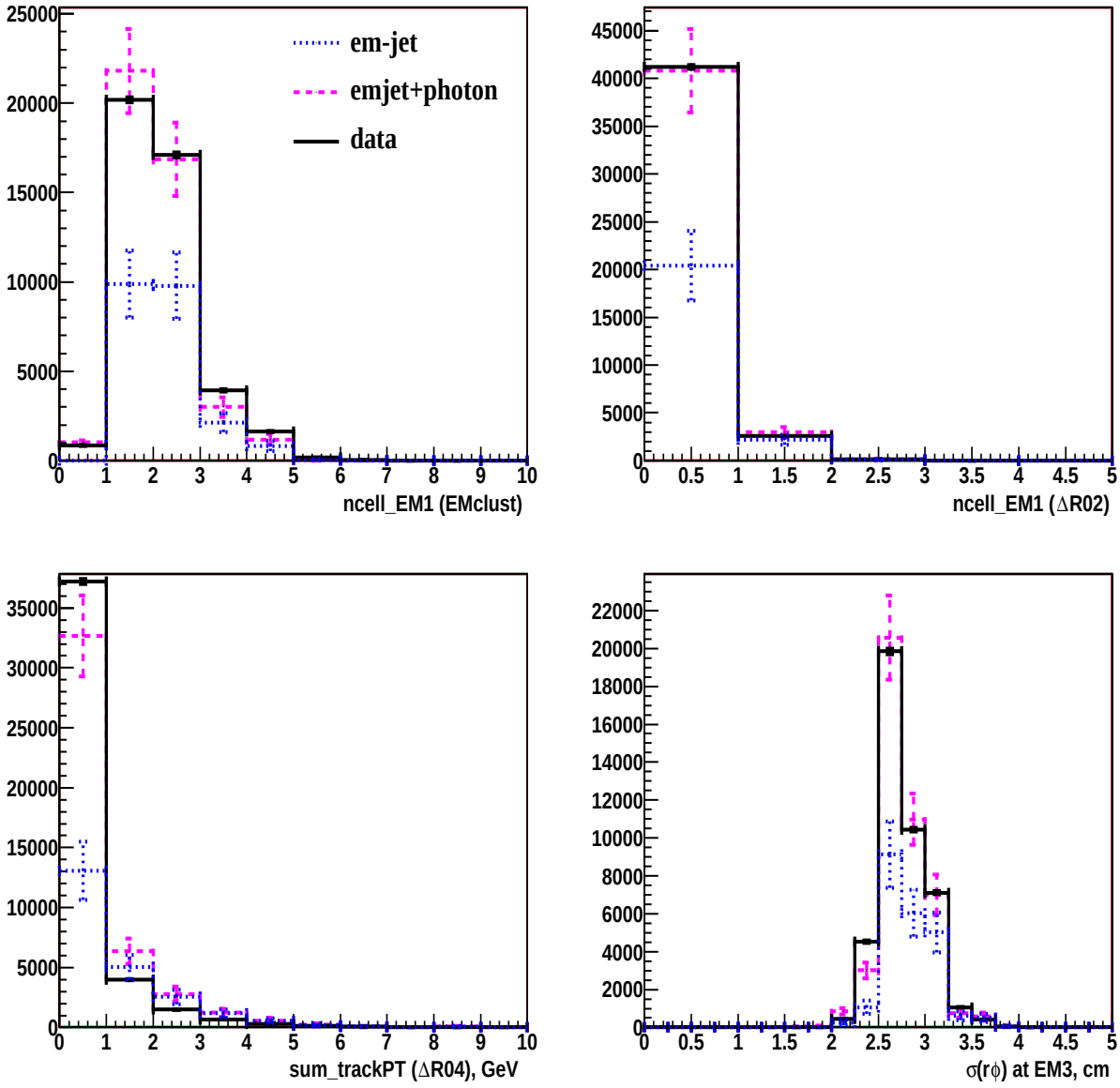

Figure 50: Same as in Fig. 47 but for $\mathbf{39 < p_T^{\gamma} < 44}$ GeV.

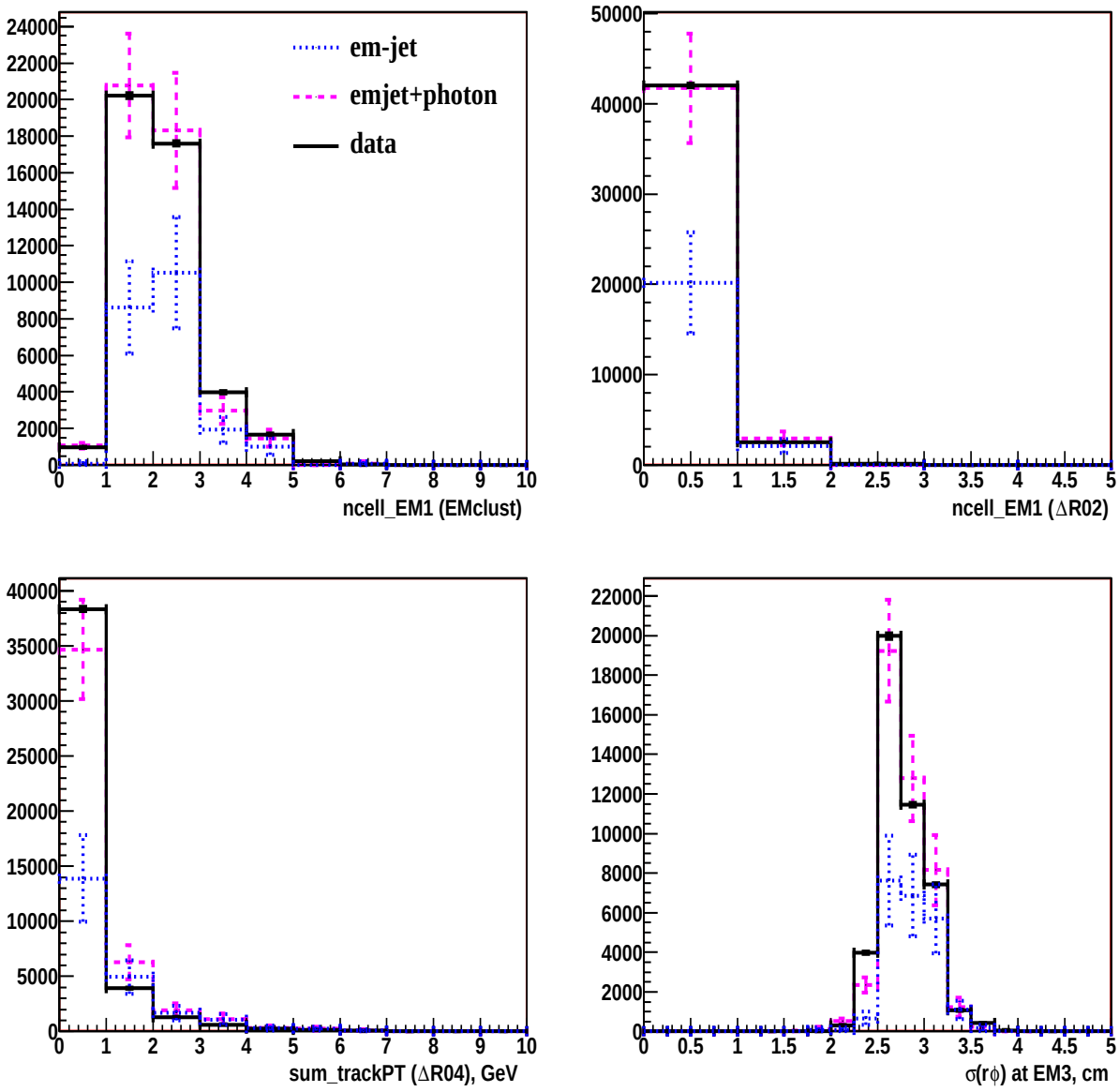

Figure 51: Same as in Fig. 47 but for $44 < \mathbf{p_T^\gamma} < 50$ GeV.

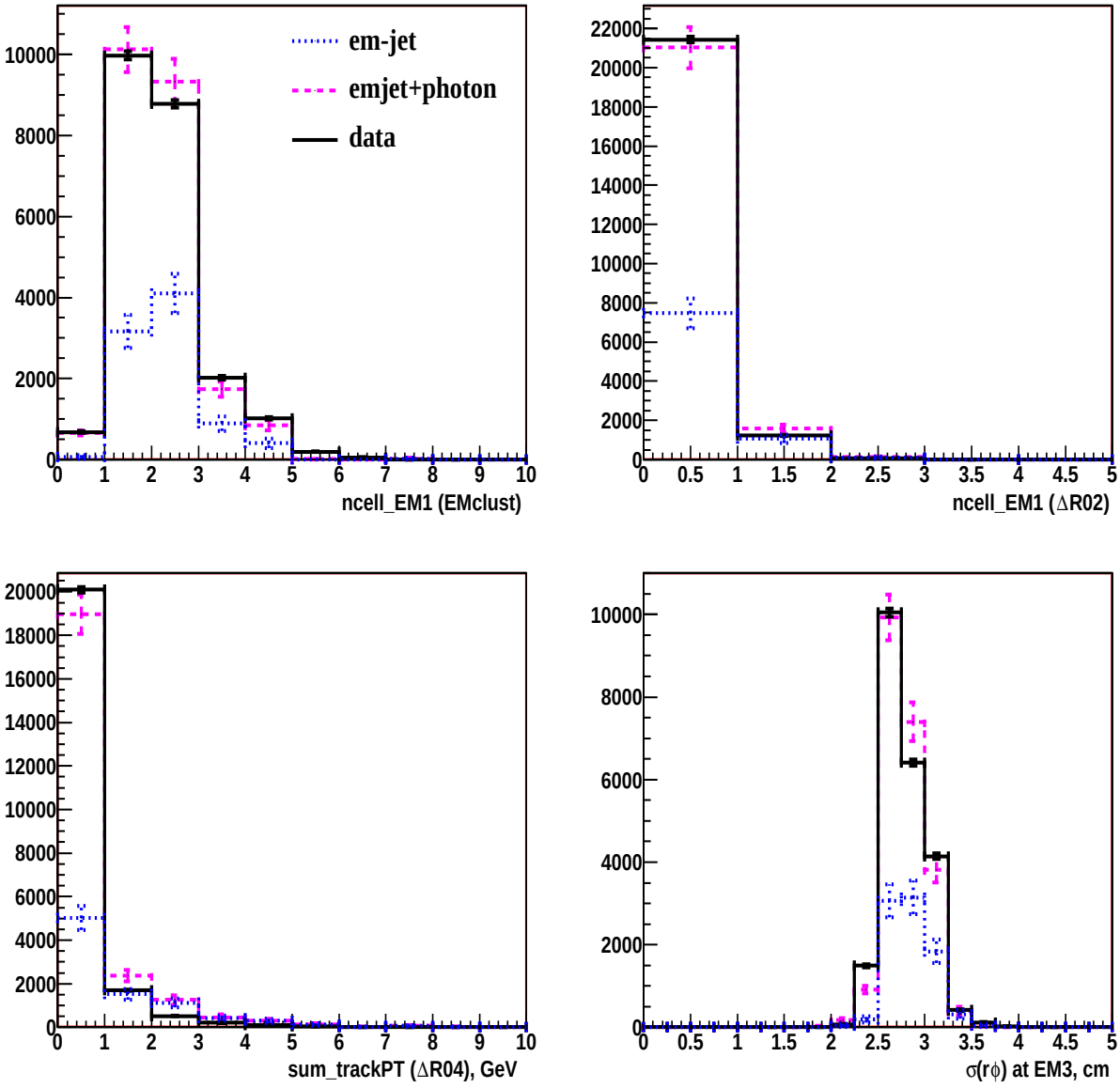

Figure 52: Same as in Fig. 47 but for $\mathbf{60 < p_T^{\gamma} < 70}$ GeV.

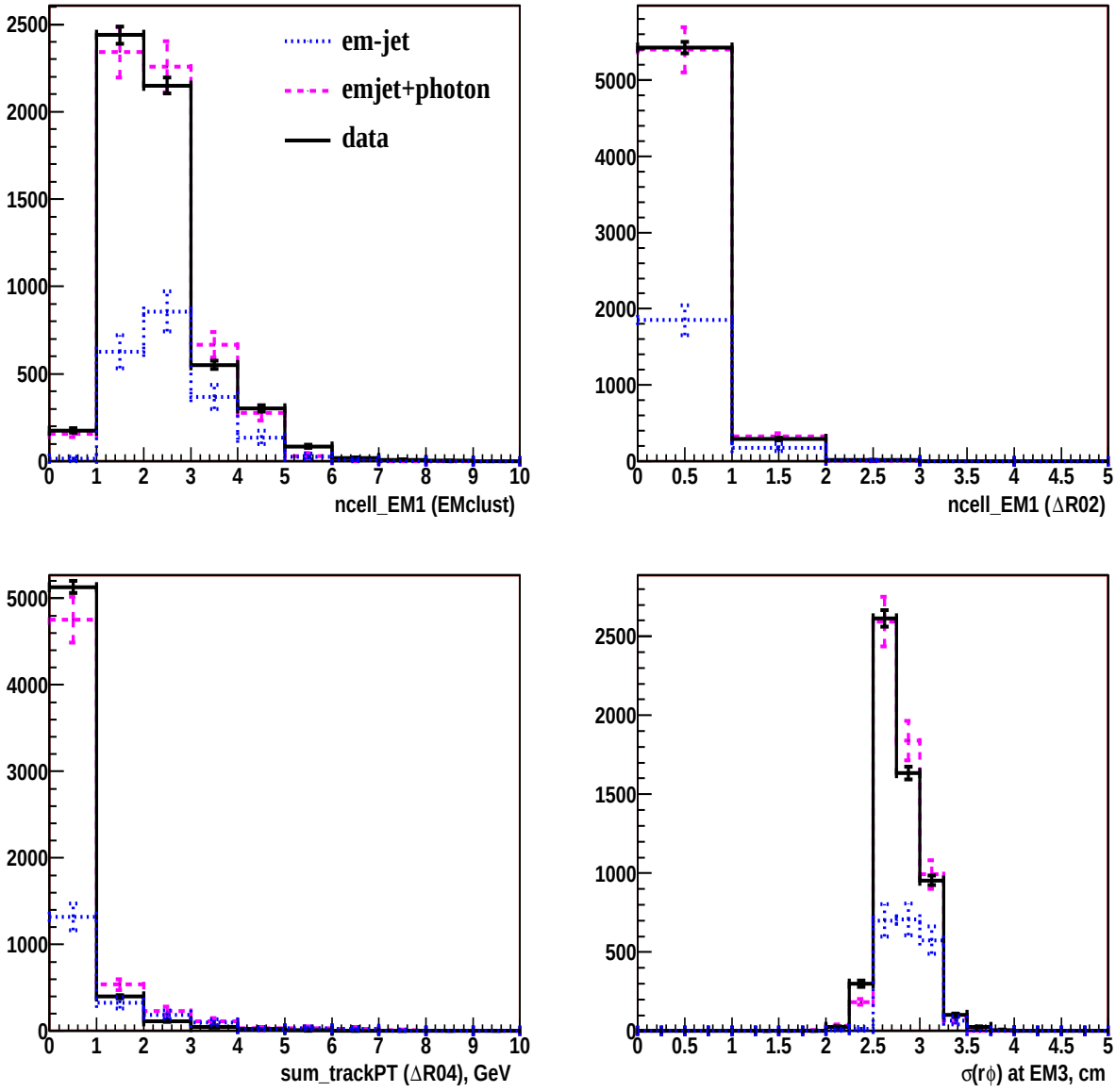

Figure 53: Same as in Fig. 47 but for $80 < p_T^{\gamma} < 90$ GeV.

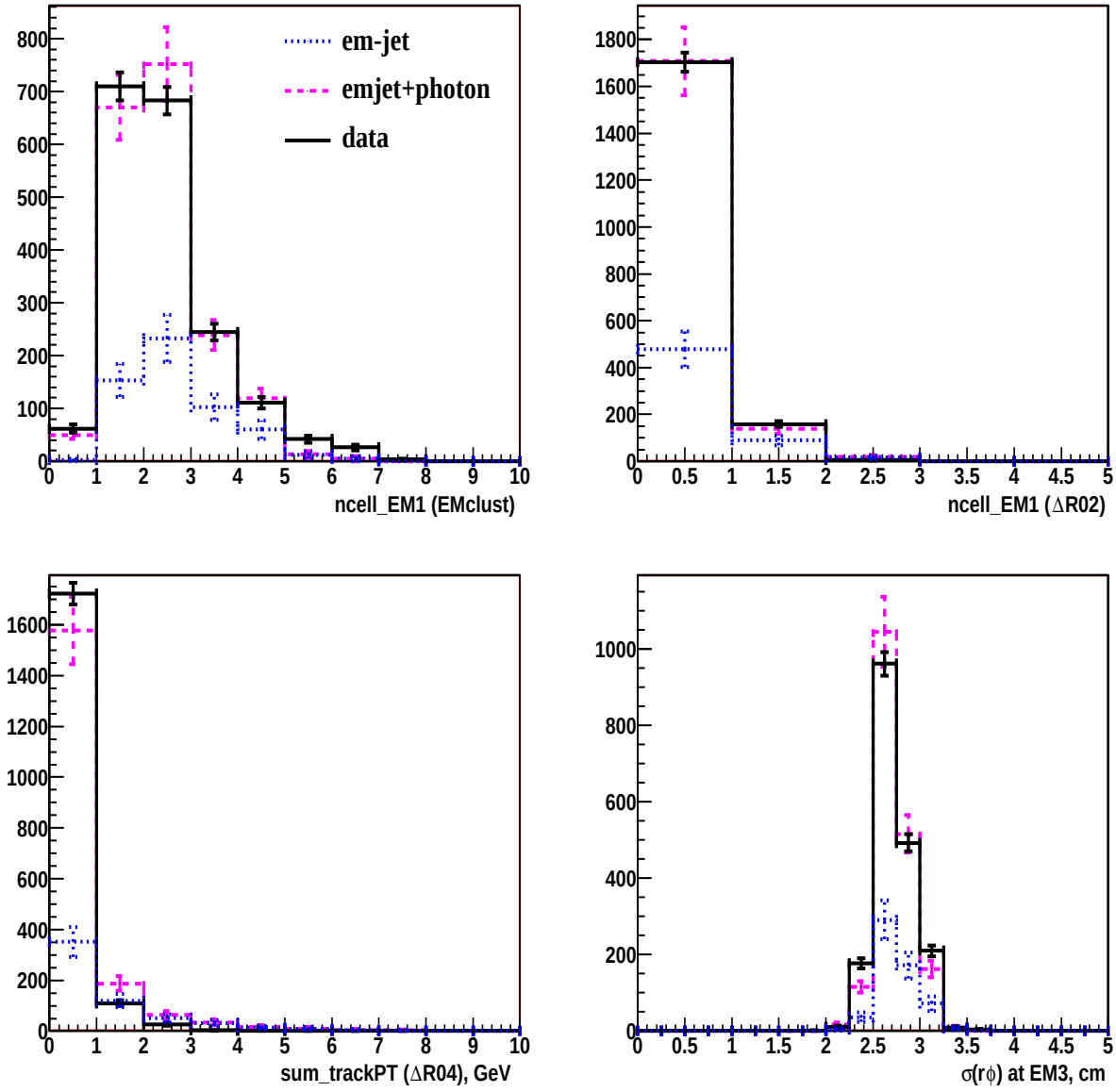

Figure 54: Same as in Fig. 47 but for $\mathbf{110 < p_T^{\gamma} < 130}$ GeV.

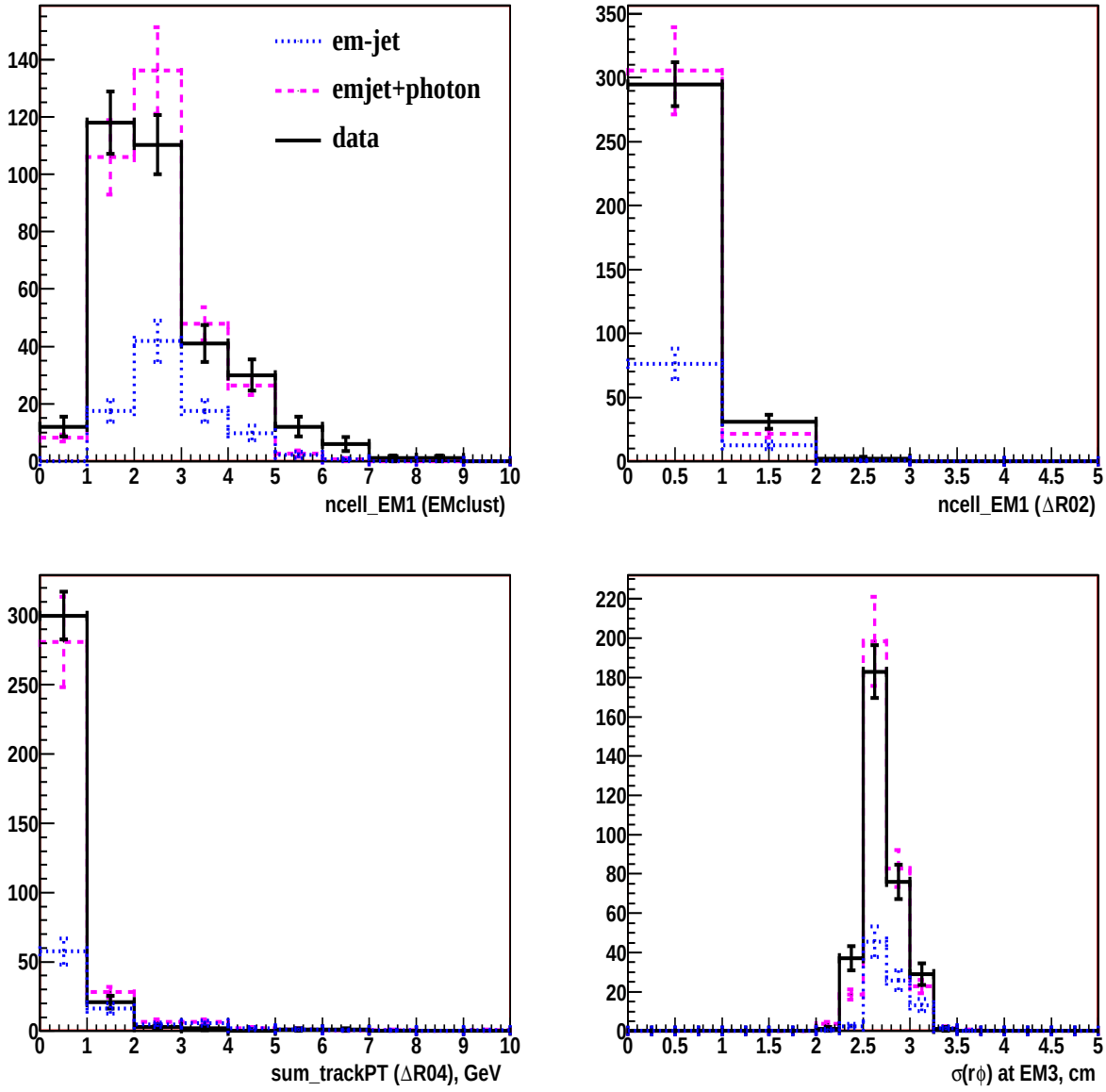

Figure 55: Same as in Fig. 47 but for $\mathbf{150 < p_T^{\gamma} < 170}$ GeV.

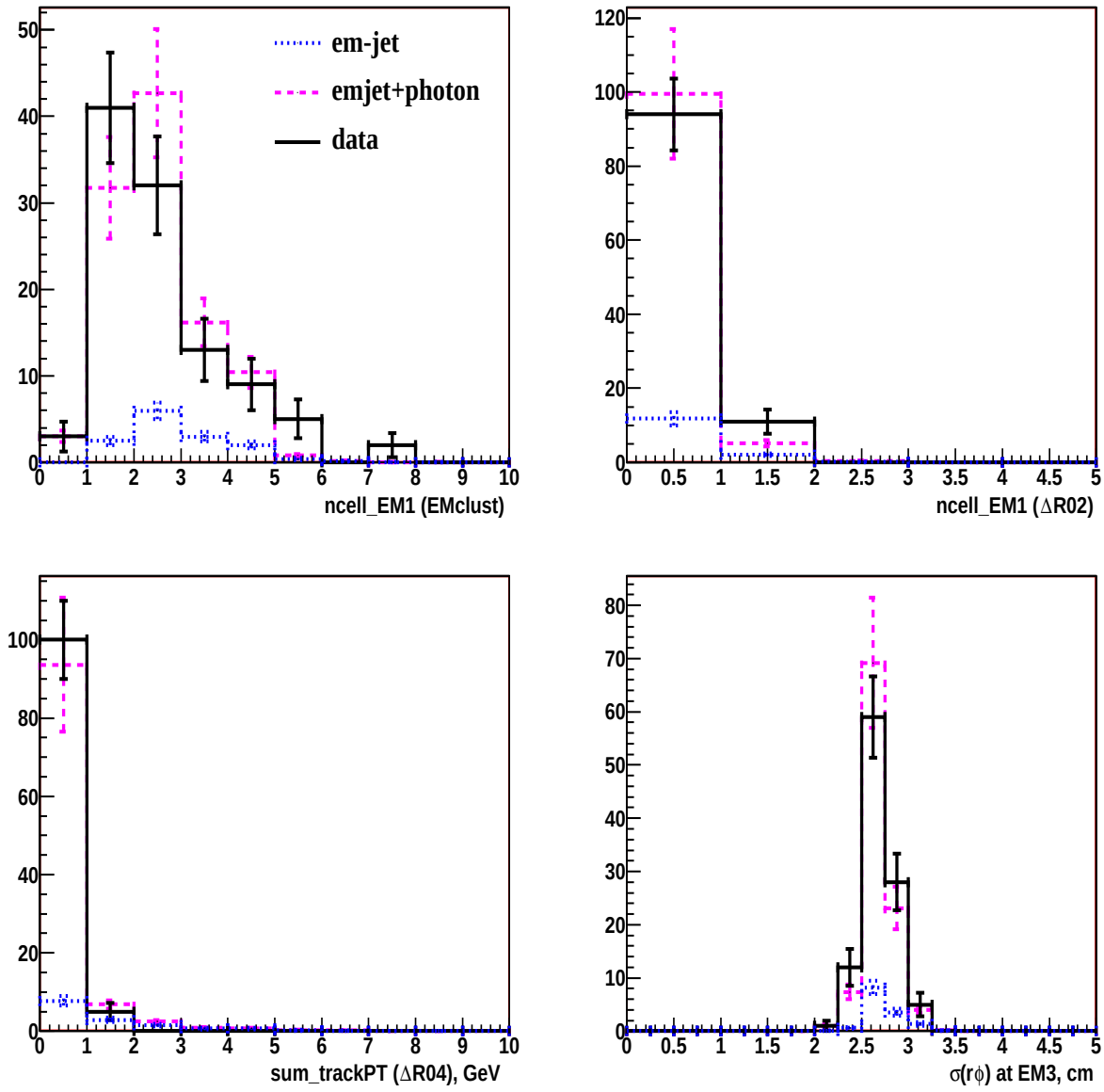

Figure 56: Same as in Fig. 47 but for $\mathbf{200 < p_T^{\gamma} < 230}$ GeV.

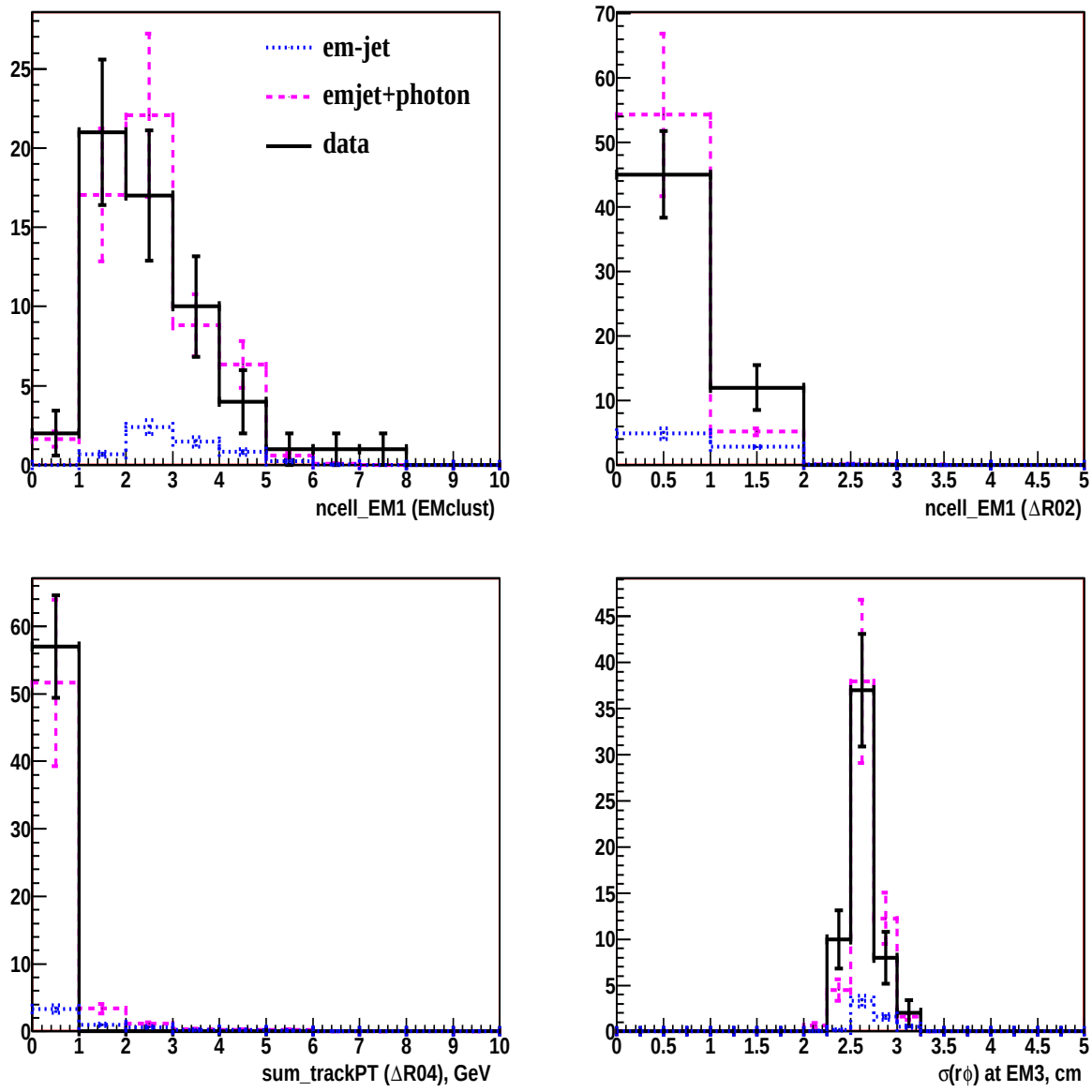

Figure 57: Same as in Fig. 47 but for $\mathbf{230 < p_T^{\gamma} < 300}$ GeV.

B. Unsmearing the cross section.

Unsmearing procedure includes the following steps:

- Make an ansatz to parametrize the initial p_T^γ spectrum. For this aim two formulas were tested:
$$\sigma(p_T) = par[0] * p_T^{-par[1]} * (1 - 2p_T/\sqrt{s})^{par[2]} \quad (16)$$
and
$$\sigma(p_T) = par[0] * p_T^{-par[1]} * exp^{-p_T/par[2]} * (1 - 2p_T/\sqrt{s})^{par[3]} \quad (17)$$
- Smear the ansatz analytically using the known energetic resolution for EM objects ($(\sigma_E/E)^2 = C^2 + S^2/E + N^2/E^2$). Here the following values were accepted [45]:
$$C = 0.0439, \;\; S = 0.224\ \sqrt{\mathrm{GeV}}, \;\; N = 0.29\ \mathrm{GeV}^2 \quad (18)$$
- Fit the smeared distribution to the data in order to adjust parameters $par[i]$ in (16), (17).
- Find the correction as the ratio of unsmeared to smeared curves in the points where cross section has been determined.

Since the resolution is defined for the energy E and we need to correct p_T spectrum, value of E for every p_T bin was defined from the average $|\eta^\gamma|$ presented in Fig. 58 as function of p_T^γ. This dependence is plotted after application of main cuts of subsection 4.2.4.

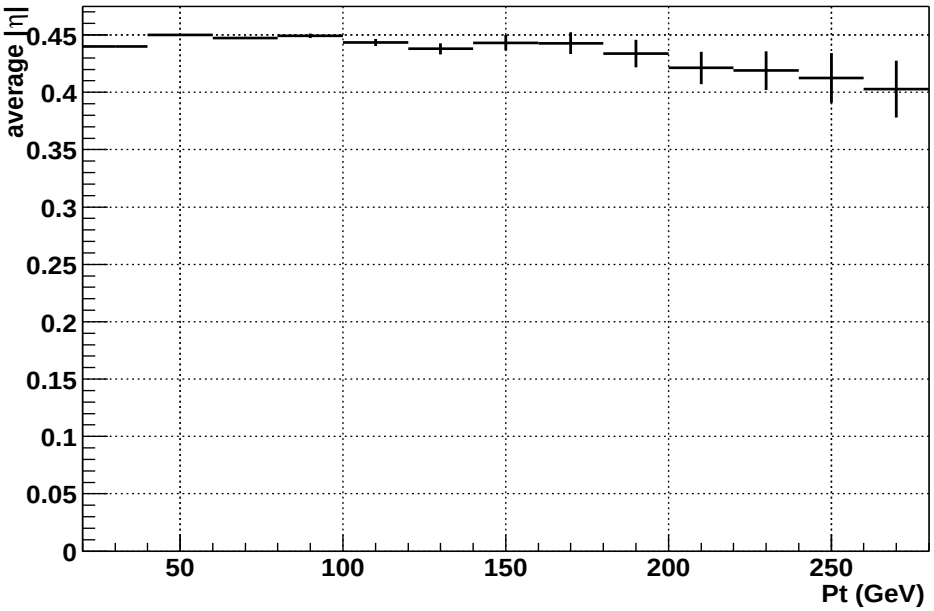

Figure 58: The dependence of average $|\eta^\gamma|$ on p_T^γ.

The ansatzes (16) and (17) smeared with EM resolution (18) were fitted to data points with

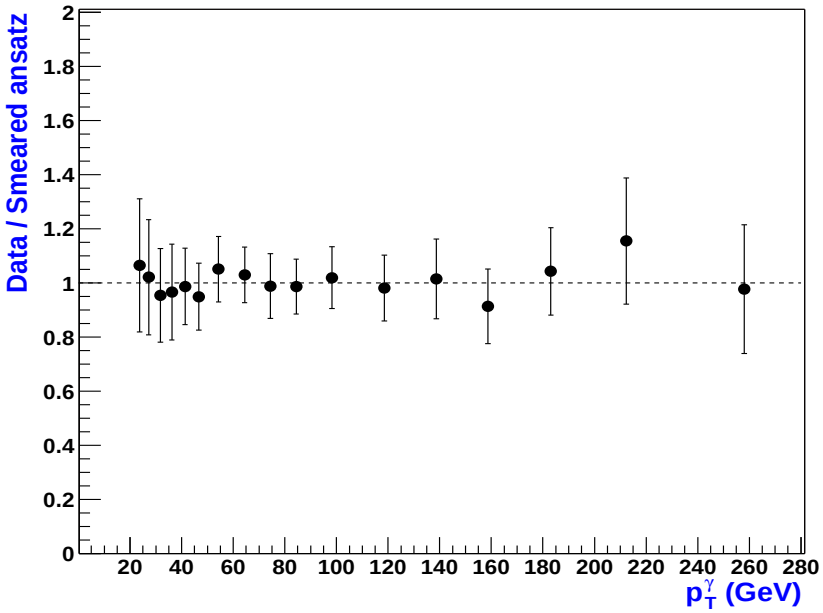

Figure 59: Ratio of the ansatz (17) smeared with EM resolution (18) to the data points. Here errors are caused by syst. error in the measured cross sesction (Table 7).

$\chi^2/ndf = 0.17$ and 0.12 respectively. Fig.59 shows ratio of the smeared ansatz (17) to the data points. The errors are caused by syst. error in the measured cross sesction (see Table 7). Differences between smeared ansatzes (16) and (17) are within 1%.

The parameters in expressions (16) and (17) obtained after the fit are:

$$par[0] = (4.68 \pm 1.10) \cdot 10^8, \;\; par[1] = 4.37 \pm 0.07, \;\; par[2] = 6.44 \pm 0.68 \qquad (19)$$

for ansatz (16) and

$$par[0] = (1.88 \pm 0.64) \cdot 10^8, par[1] = 4.05 \pm 0.17, par[2] = 30.90 \pm 9.68, par[3] = -18.44 \pm 7.37$$

for ansatz (17).

The final unsmearing correction is shown in Fig.60. The errors are defined as difference between average corrections obtained with ansatzes (16) and (17).

One can see that the found correction can parametrized by linear function $f_{unsm} = A - B \cdot p_T^\gamma$ with $A = 0.971 \pm 0.002$ and $B = (9.061 \pm 1.432) \times 10^{-5}$. The relative error caused by the fit is 0.2% at $p_T^\gamma = 23.7$ GeV and grows up to 0.5% at $p_T^\gamma = 258$ GeV. But due to a possible distinction of the true resolution from (18) accepted here we increase the relative errors of the fit to 1% for all points. This correction will be taken into account to get the final cross section [22)] from eq. (1).

[22)] Is is interesting to note that this procedure repeated for Run I resolution (with $C = 0.004$, $S = 0.15$) gives correction 0.993 ± 0.005.

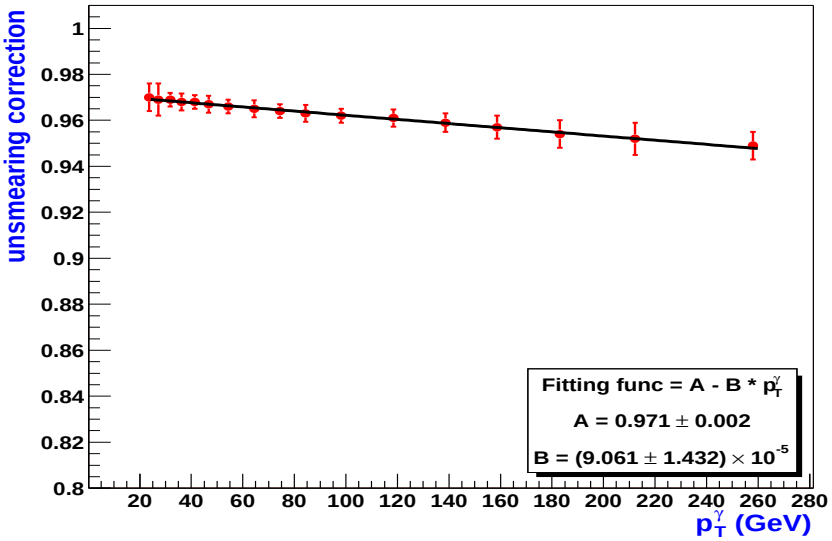

Figure 60: Dependence of the unsmearing correction on p_T^γ. Linear fitting function is shown on the plot with values of its parameters.

C. Variation of cross section with cut on NN_{output}.

Let us remind that the used ANN was trained on MC signal and background events and has to produce 0 in case of background and 1 in case of signal events. Thus, natural choice for a boundary between these types of events is middle point 0.5. That is why the photon purity estimation was done after application of cut '$NN_{output} > 0.5$'. Certainly, one can use a tighter cuts to achieve a higher purity. But on the other hand, error to the purity would also increase due to reduction of statistics (mainly in background and data).

In Fig. 61 we demonstrate ratios of photon purity found with $NN_{output} > 0.3$ and $NN_{output} > 0.65$ to the purity found with $NN_{output} > 0.5$ cut. Only largest errors (that correspond to tighter cut in the ratio) are shown. Ratios of the final cross sections found with $NN_{output} > 0.3$ and $NN_{output} > 0.65$ to one with 'middle-point' cut $NN_{output} > 0.5$ are presented in Fig. 62. As in Fig. 61, only largest errors are shown.

From Fig. 62 one can conclude that the cross section is very stable with respect to variation of the cut on ANN output.

D. Study of energy saturation at high p_T^γ.

In principle, there can be a danger that the calorimeter signal from high energetic photon will be saturated. To study this question, total energy depositions of the EM cluster as well as the energy of the hottest cell at five calorimeter compartments (EM1–EM4, FH1) caused by the photons with $p_T^\gamma \geq 170$ GeV have been considered. They are presenetd in Fig. 63 and 64. As we see, no energy saturation is observed.

E. Study of systematic error due to the choice of vertex position.

Since we are selecting signal EM clusters without track match, there is some probability of misidetification of the chosen primary vertex (PV) in an event. This fact can lead to an incorrect calculation of photon transverse momentum p_T^γ. In order to detrmine the systematic error due to the choice of PV, we have calculated p_T^γ w.r.t. the PV that gives minimal p_T^γ in an event, $p_{T,min}^\gamma$, in addition to $p_{T,best}^\gamma$, defined for the best PV in the same event. The distributions over $frac = (p_{T,best}^\gamma - p_{T,min}^\gamma)/p_{T,best}^\gamma$ for three p_T^γ intervals are shown in Fig. 65. Here all PV are considered. We see that the mean value of the sfift $\langle frac \rangle$ is within 0.6-0.8% and almost stable w.r.t. p_T^γ. Same distributions just for the events with $frac > 10^{-3}$ are shown in Fig. 66. The fraction of events that are underflow (i.e. with $frac < 0.1\%$) for those histograms is about 90% and is practically flat w.r.t. p_T^γ. It is worth to note that about $75 \pm 3\%$ of events have just one PV. The distributions over the number of PV for five p_T^γ intervals are presented in Fig. 67. The bottom right graph of this figure shows a fraction of events versus p_T^γ. Taking into account the stat. errors one can conclude that this fraction is flat (being of order 0.22 ± 0.02).

Thus, an average error due to the choice of a wrong vertex for the extreme case considered above is about $0.6 - 0.8\%$. For the following analysis the systematic uncertainty of $\Delta p_T^{vert} = 0.8\%$ has been taken. Such an uncertainty leads to the uncertainty in the photon cross section [23)] of 3.6-5.0% (it slightly grows with p_T^γ). But at the same time it is worth mentioning that some simple criteria on the choice of considered PV (EM cluster pointing, PV χ^2, number of tracks at PV) can even reduce it.

[23)] See equation (9) and the text above.

F. Photon energy scale correction.

As we wrote in Section 4.3, due to calibration on electrons the photon p_T is overestimated as compared with particle level: by about 2% at 20 GeV, 1% at 40 GeV. Additional influence on the p_T scale is caused by a presence of background events. These events, mainly caused by π^0 (and η) meson, lead to a change of the observable p_T^{reco} due to following two effects.

Firstly, results from a few, as colider $pp(\bar{p})$ as fixed target, experiments give evidences that π^0 cross section drops faster than direct photon cross section [2, 9]. It means that for the π^0 events an average p_T in a given bin of p_T^{reco} should be smaller than that for the photon events. Secondly, the produced π^0 (or η) mesons usually have some soft accompanying particles around them (i.e. have $z < 1$). Due to interaction with a matter upstream the calorimeter, magnetic field and the fact that EM cluster cone size is much smaller than the "fragmentation cone size" (with $R \simeq 0.5 - 1.0$), some fraction of the p_T carried by the soft particles are lost. One can expect that fraction of the lost energy (p_T) will be maximal at small p_T's and then gradually drops. Fig. 68 shows distributions of the number of events over the relative shift between transverse momenta of the photon candidate in the QCD events at the calorimeter $p_{T,calo}^{em.clust}$ and at the particle $p_{T,part}^{jet}$ levels after application of all photonic criteria. Thus, in spite of strict photon ID criteria we are not able to collect full pT of initial particle jet (parton) in EM cluster. We can observe that the EM jet, registered as a EM cluster – photon candidate, loses, for instance, about 8.5% of its initial particle level transverse momentum at $10 < \hat{p}_\perp^{min} < 20$ GeV and about 4% at $40 < \hat{p}_\perp^{min} < 60$ GeV. Fraction of the lost p_T (i.e. shift to a smaller reconstructed p_T) depends on a p_T scale as well as η and selection criteria [58].

The both effects, described above, work in the same direction of descreasing observable p_T.

What we measure in data is p_T of photon candidate (EM cluster), two sets of events contribute: single photons and $\pi^0(\eta)$- em-jets. Thus, the question we should answer is what is combined shift of the observed p_T w.r.t. true (particle level) photon p_T. Fig.69 shows distribution of EM cluster p_T for signals (1st column), for background (3rd column), of true photon p_T (middle column) for $14 < p_T^{reco} < 28$ GeV in MC signal and QCD events generated with $\hat{p}_\perp^{min} = 10$ GeV. We see that recon'd p_T of signal EM cluster is shifted, on the average, by "+"3% w.r.t. true pT while pT of EM cluster produced by em-jets is shifted by "-"7%! We also wee that background event rate drops much faster and have noticeble smaller RMS than for the signals. Analogous behaviour takes place for other p_T^{reco} intervals, too, as shown in Figs.70–72.

Since when we register "$\gamma + jet$" event we have a mixture of direct photon and QCD events, the total combined shift of the energy scale for a photon candidate (F_{comb}) should be estimated from the weighted sum of the negative shift caused by direct photons (F_γ) and the positive shift caused by electromagnetic jet (F_{emj}):

$$F_{comb} = F_\gamma \cdot P + F_{emj} \cdot (1 - P) \tag{20}$$

Here P is a fraction (purity) of the direct photon events in the accumulated data sample and serves here as a weight factor.

The total shift caused by the sum (in terms of equation (20)) of the shifts caused by photons and em-jets is shown in Fig. 73.

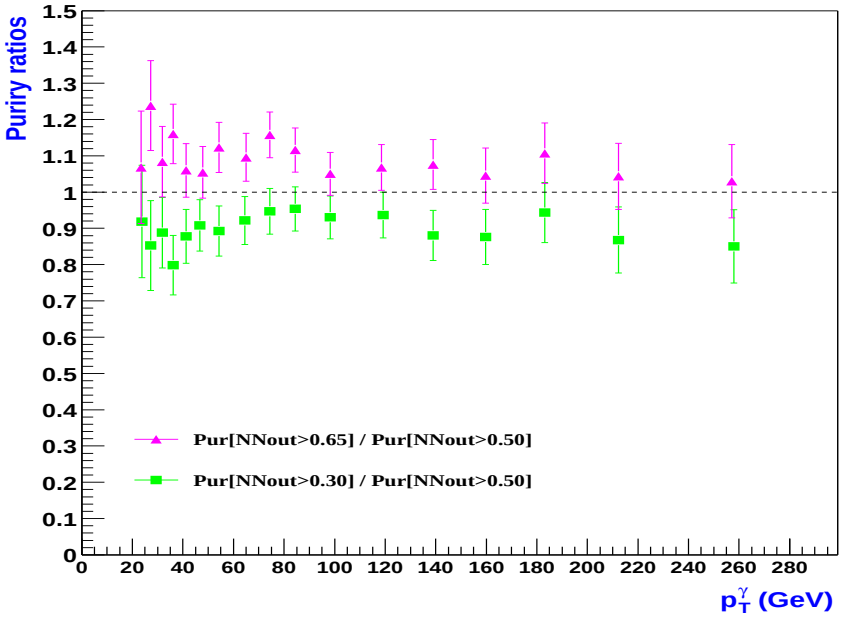

Figure 61: Ratios of photon purity found with $NN_{output} > 0.3$ and $NN_{output} > 0.65$ to the purity with $NN_{output} > 0.5$ cut.

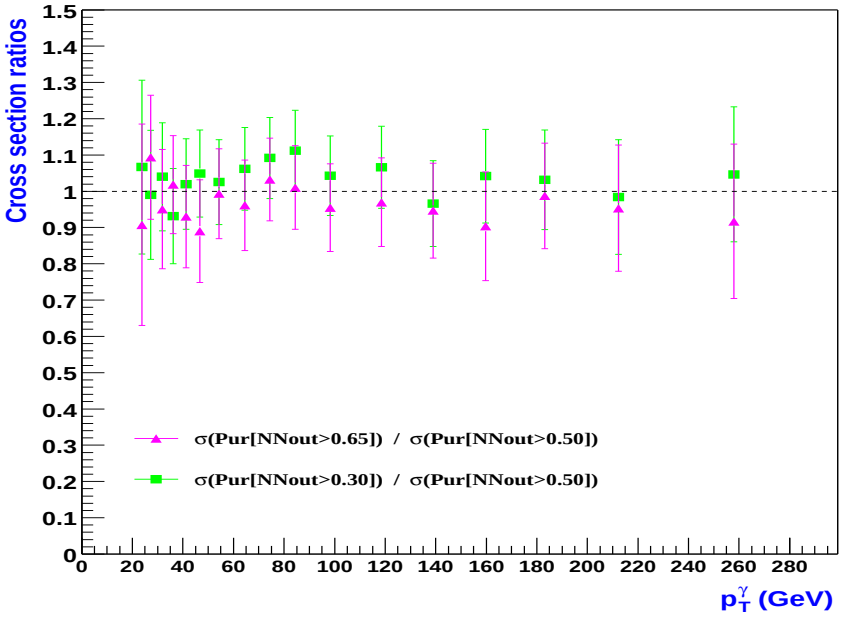

Figure 62: Ratios of cross sections found with $NN_{output} > 0.3$ and $NN_{output} > 0.65$ to the cross section with $NN_{output} > 0.5$ cut.

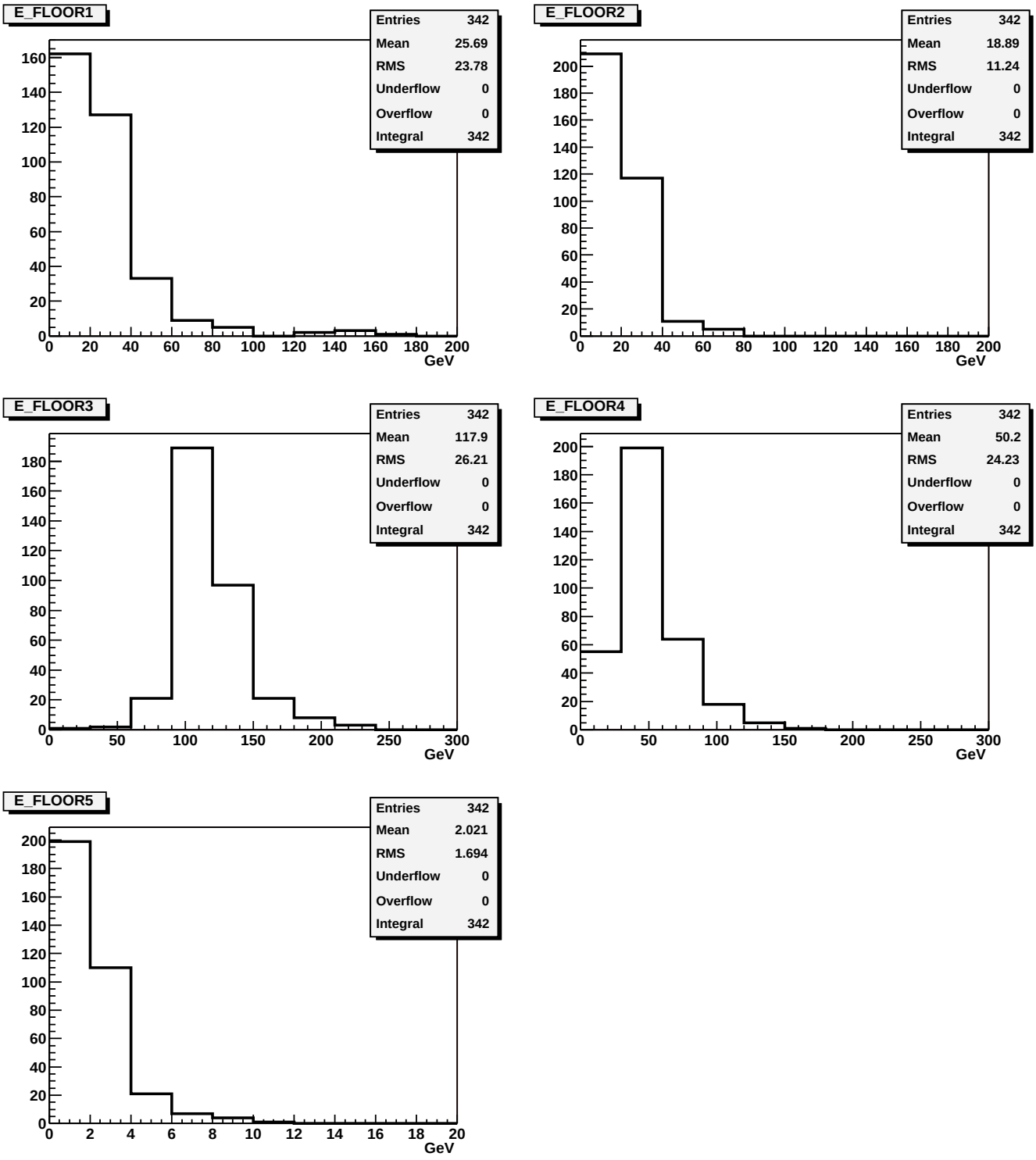

Figure 63: Distribution over the layer energy at EM1–EM4, FH1 of the photon EM cluster at $p_T^{\gamma} \geq 170$ GeV in *data*.

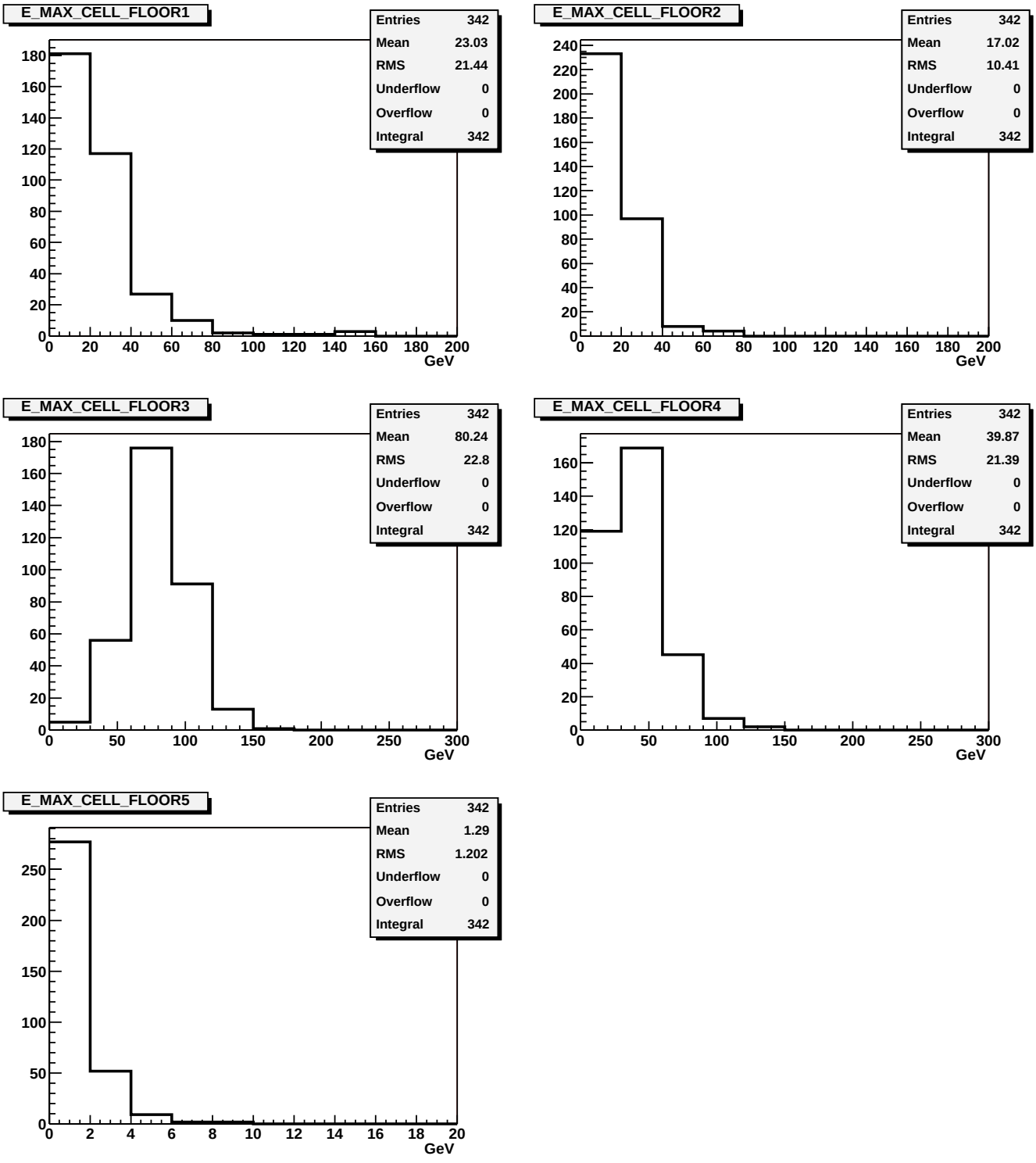

Figure 64: Distribution over the energy of hottest cell at EM1–EM4, FH1 of the photon EM cluster at $p_T^{\gamma} \geq 170$ GeV in *data*.

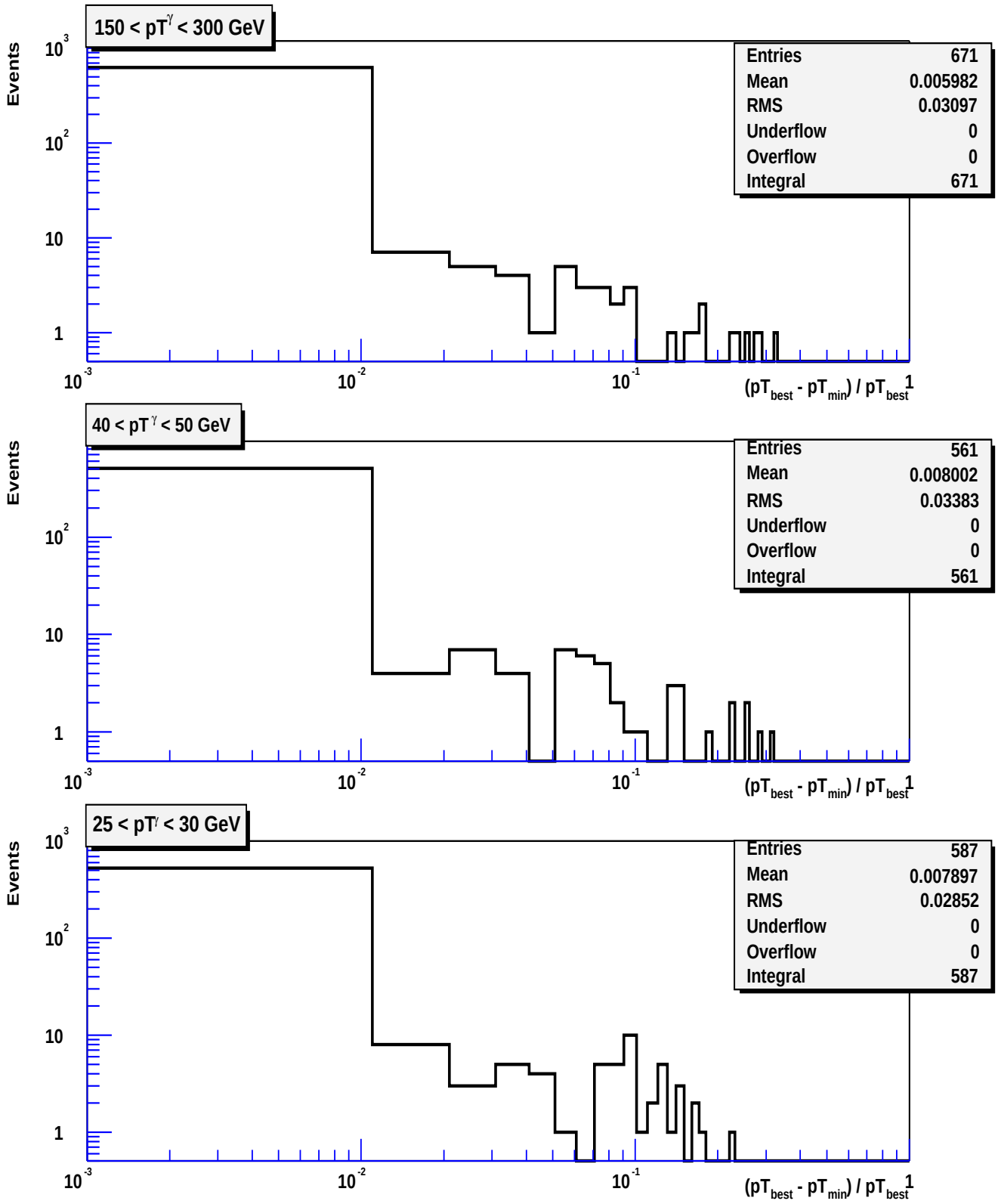

Figure 65: Distribution over $frac = (p^{\gamma}_{T,best} - p^{\gamma}_{T,min})/p^{\gamma}_{T,best}$ for three p^{γ}_{T} intervals in *data*. Here pT_{best} is the transverse momentum defi ned for the best primary vertex (PV), pT_{min} is pT for the PV that produces a minimal pT in a given event. **ALL PV are considered** ($0 \leq frac < 1$).

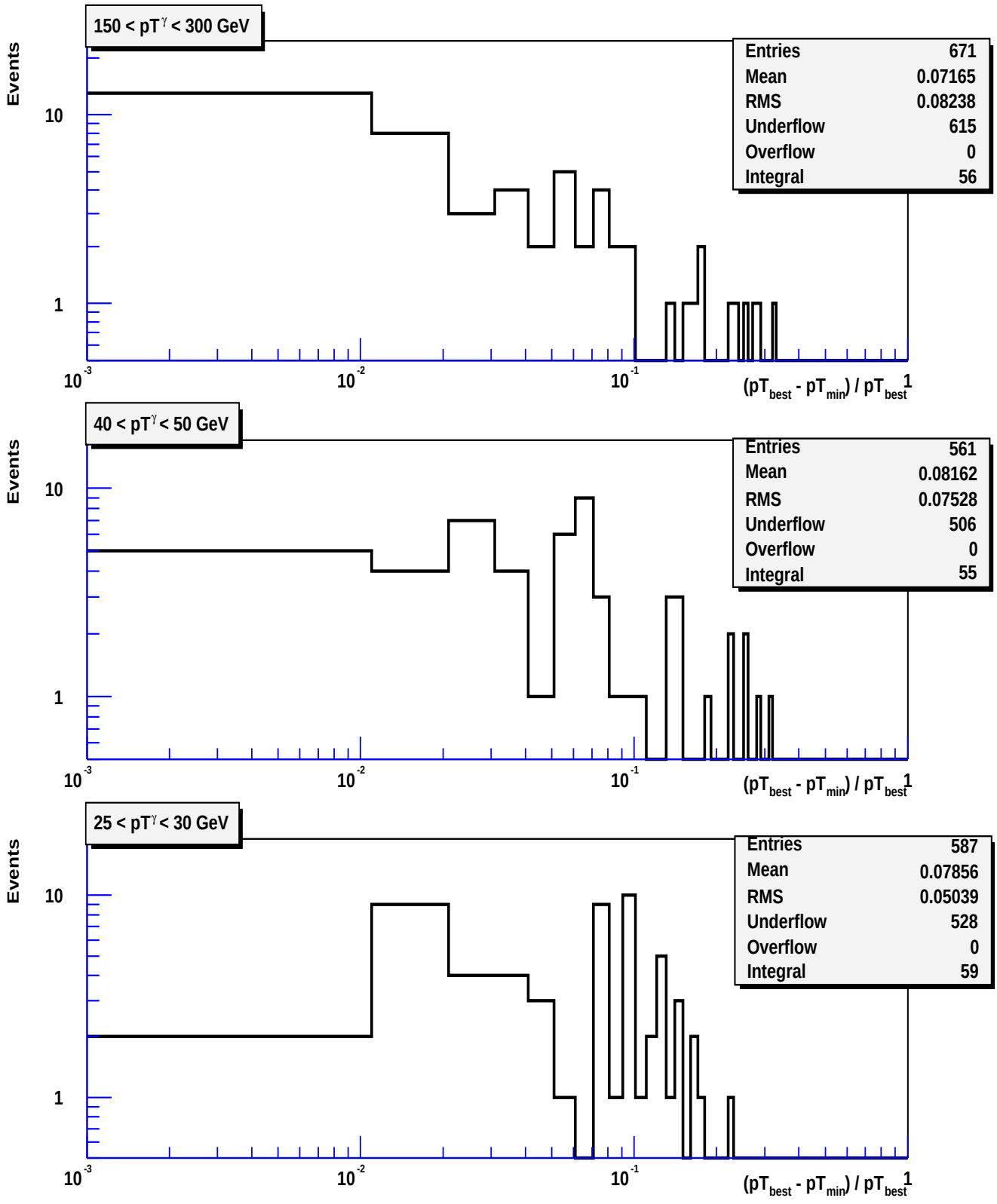

Figure 66: Distribution over $frac = (pT_{best} - pT_{min})/pT_{best}$ for three p_T^{γ} intervals in *data*. Here pT_{best} is the transverse momentum defi ned for the best primary vertex (PV), pT_{min} is pT for the PV that produces a minimal pT in a given event. **Just the events with $frac > 10^{-3}$ are considered**.

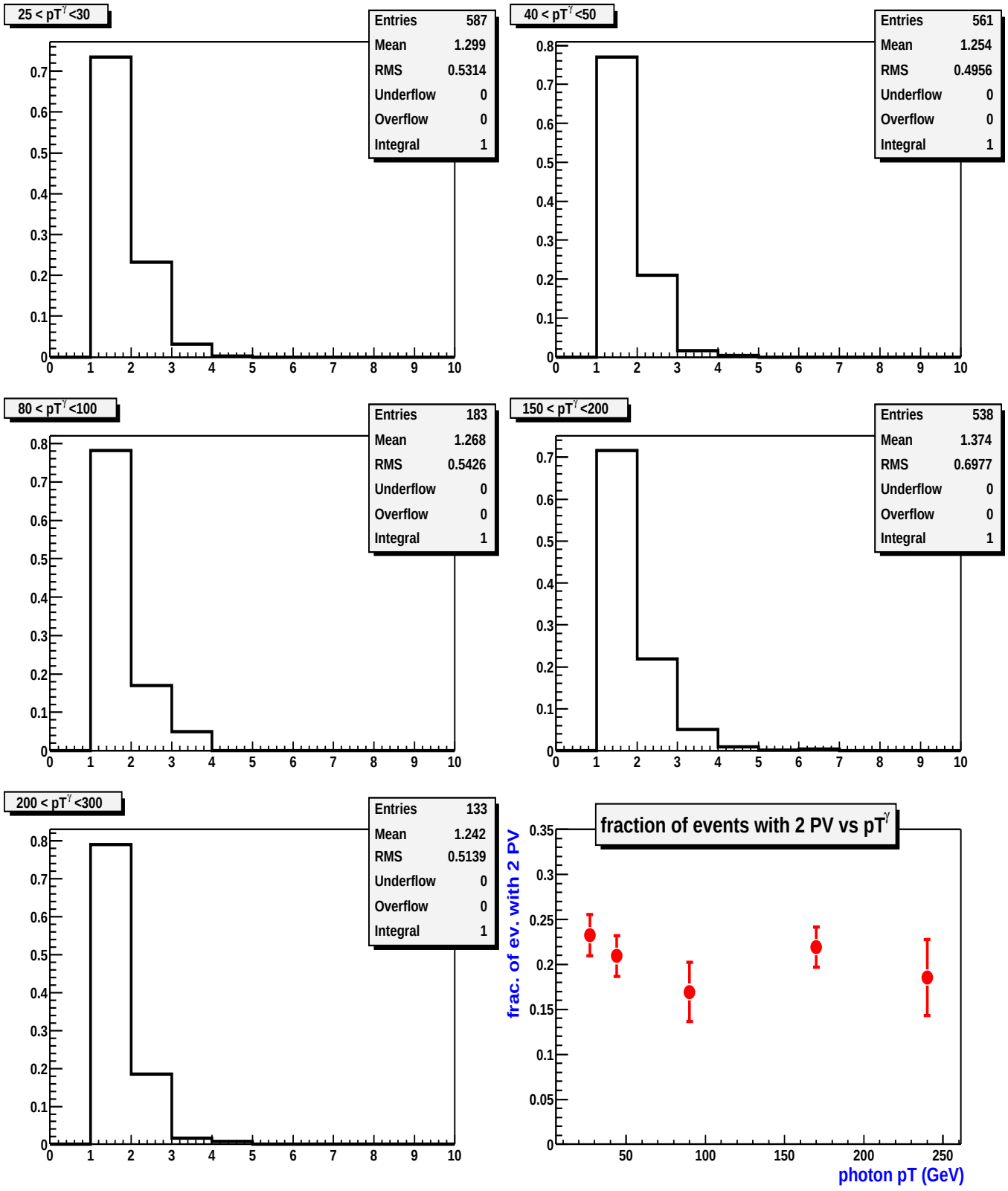

Figure 67: Distribution over the number of PV for 5 p_T^γ intervals. The bottom right graph shows a fraction of events (with stat. error) vs. p_T^γ.

$$(\mathbf{p_T^{part.jet}} - \mathbf{p_T^{em.clust}})/\mathbf{p_T^{part.,jet}}$$

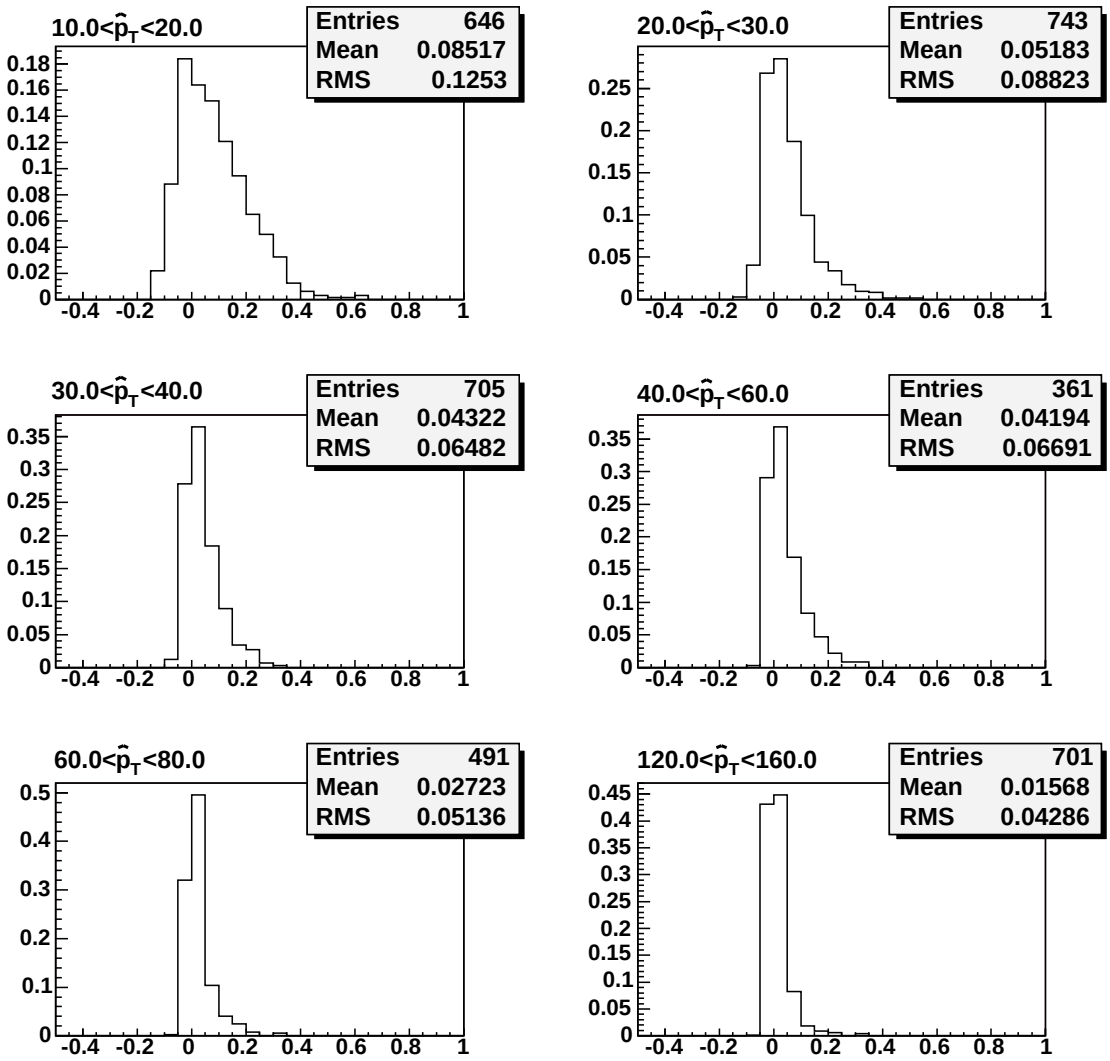

Figure 68: Distributions of the number of events over the relative shift between transverse momenta of the photon candidate in the QCD event at the calorimeter $p_{T,calo}^{em.clust}$ and at the particle $p_{T,part}^{jet}$ levels.

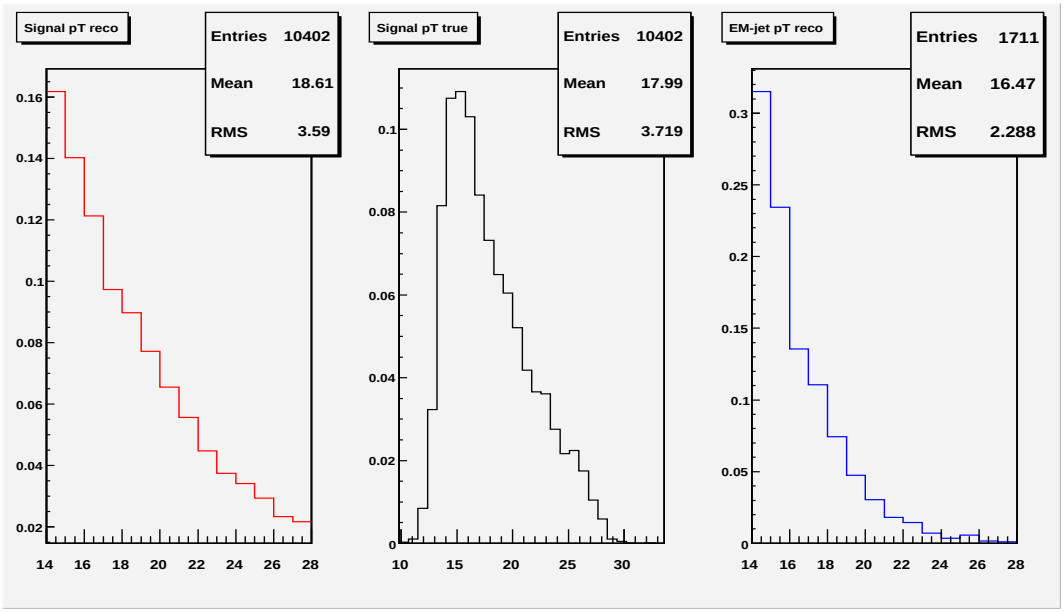

Figure 69: Distribution of EMcluster p_T for signals (1st column), for background (3rd column), of true photon p_T (middle column) for $14 < p_T^{reco} < 28$ GeV in MC signal and QCD events generated with $\hat{p}_\perp^{min} = 10$ GeV.

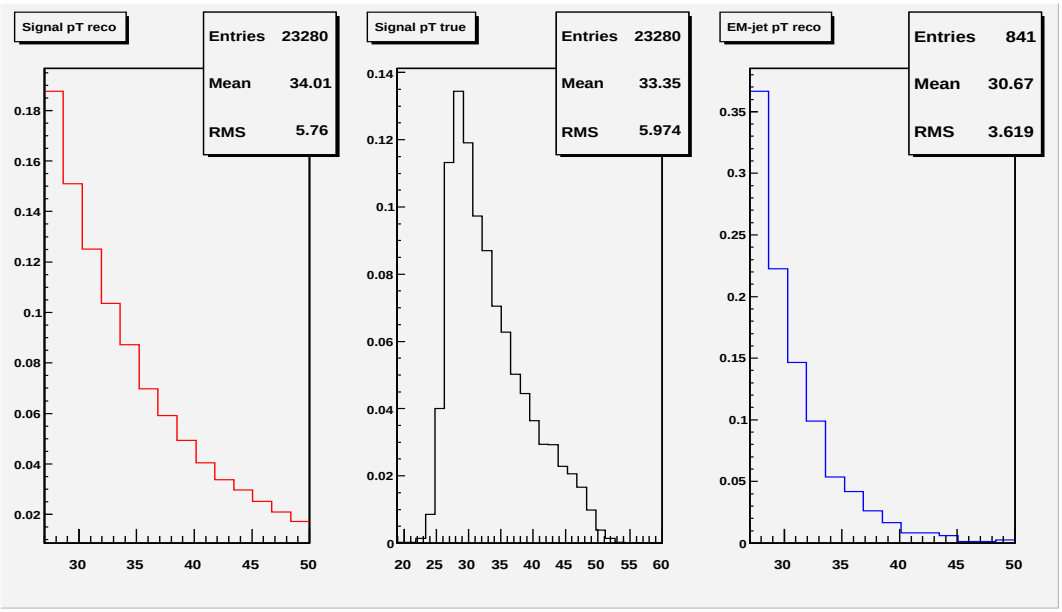

Figure 70: Same as in Figure 69 but for $27 < p_T^{reco} < 50$.

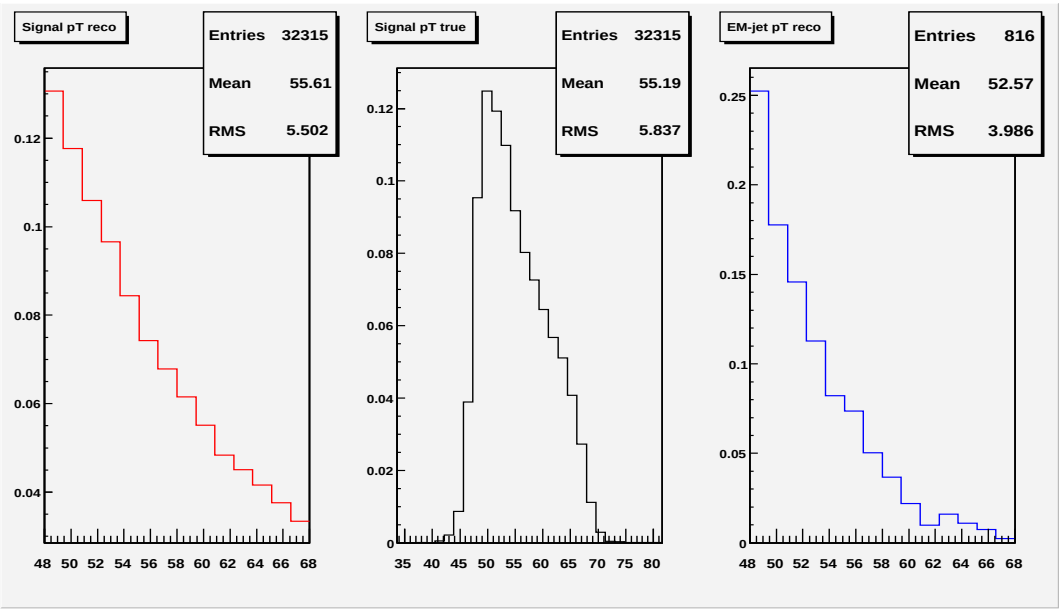

Figure 71: Same as in Figure 69 but for $48 < p_T^{reco} < 68$.

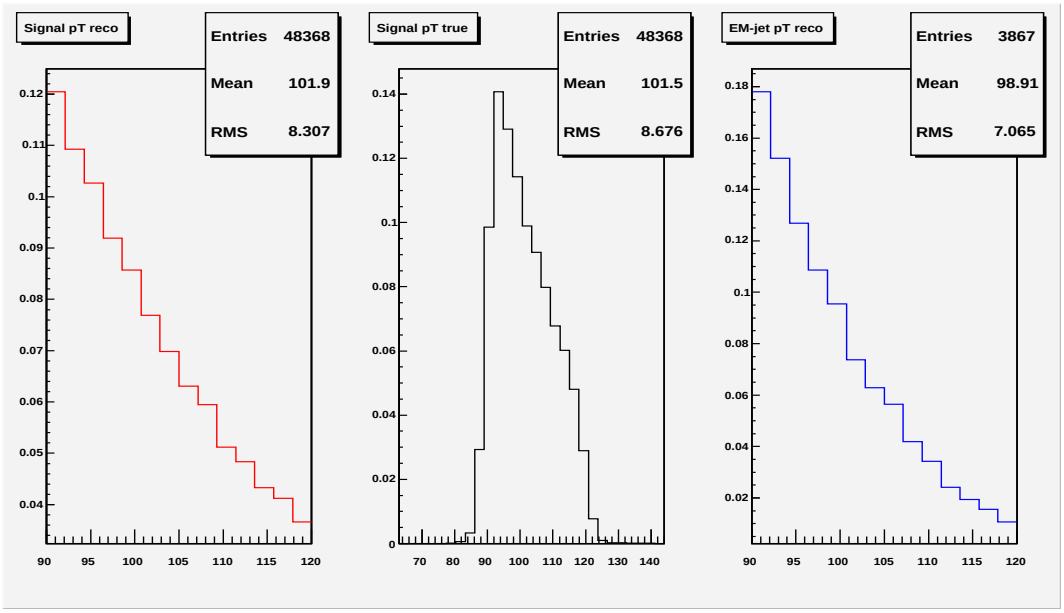

Figure 72: Same as in Figure 69 but for $90 < p_T^{reco} < 120$.

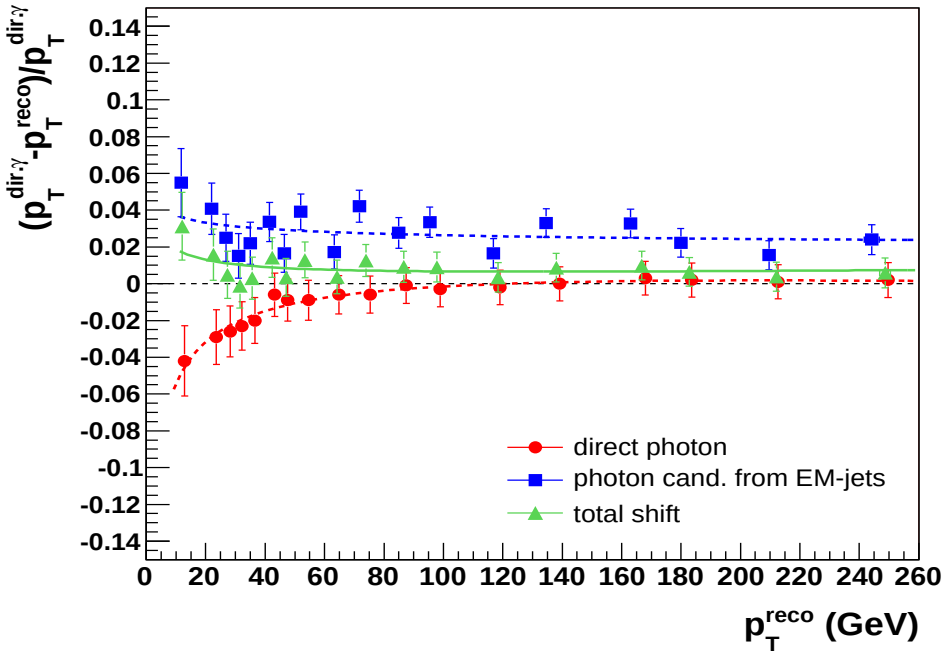

Figure 73: Average values of relative p_T shifts $F \equiv (p_T^{part,\gamma} - p_T^{calo,\gamma})/p_T^{part,\gamma}$ in signal and background events as well as the total shift (see equation (20)).