6th International Conference on Particle Physics and Astrophysics

The estimation of the background induced by the misidentification of a jet as a photon by slice method at *pp* collider experiment

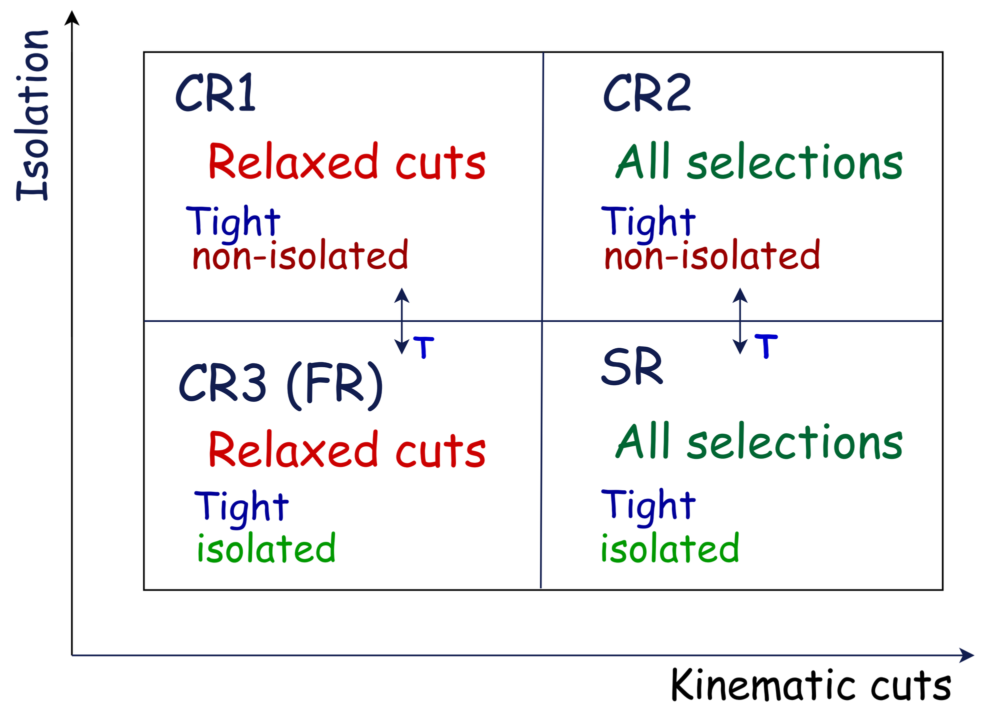
K. Kazakova, E. Soldatov, D. Pyatiizbyantseva
National Research Nuclear University MEPhI

Introduction

Background processes induced by object misidentification cannot be modelled well using the Monte Carlo (MC) simulation. Many analyses at the Large Hadron Collider (LHC) experiments use data-driven methods to solve this issue. This work presents an approach to the estimation of the background induced by the misidentification of a jet as a photon ($jet \to \gamma$). The method considers the shapes of data, signal and all other backgrounds. Taking into account the dependence between the isolation and the estimate leads to an increase in accuracy, which is important for differential cross-section measurements.

Methodology

The proposed method splits the phase-space into 4 regions based on kinematic cuts and photon *isolation*[a]. The signal region (SR) and control region 2 (CR2) are regions that pass all kinematic selections and where events have isolated and non-isolated photon candidates, respectively. The fit region (FR) and the control region 1 (CR1) are extended regions of the SR and the CR2 with relaxed cuts on several kinematic variables. Photons in all 4 regions pass the *tight*[a] identification criterion. The figure below represents the schematic illustration of the proposed method.



The fit is performed in the FR for those variables on which restrictions are removed, where the $jet \to \gamma$ process used for the fit is derived from the CR1. To study the dependence of the result on the isolation criteria, the non-isolated CR1 and CR2 are split into consecutive intervals over the isolation variable. In this way, the number of $jet \to \gamma$ background events for a given isolation slice i in the CR1 can be estimated as follows:

$$N^{jet\to\gamma}_{\mathrm{CR1(i)}} = N^{\mathrm{data}}_{\mathrm{CR1(i)}} - N^{\mathrm{sig}}_{\mathrm{CR1(i)}} - N^{\mathrm{bkg}}_{\mathrm{CR1(i)}}.$$

Thus, the total number of events in the FR estimated from the non-isolated slice of the CR1 is given by:

$$N^{\mathrm{data}}_{\mathrm{FR(i)}} = N^{\mathrm{sig}}_{\mathrm{FR(i)}} + N^{\mathrm{bkg}}_{\mathrm{FR(i)}} + N^{jet\to\gamma}_{\mathrm{FR(i)}}.$$

The fitting parameter $T_{(i)}$ allows to estimate the number of $jet \to \gamma$ events in the FR: $N^{jet\to\gamma}_{\mathrm{FR(i)}} \approx T_{(i)} \cdot N^{jet\to\gamma}_{\mathrm{CR1(i)}}$. Finally, the fitted $jet \to \gamma$ yield is extrapolated to the SR. The estimate for each slice and kinematic variable is determined by the equation:

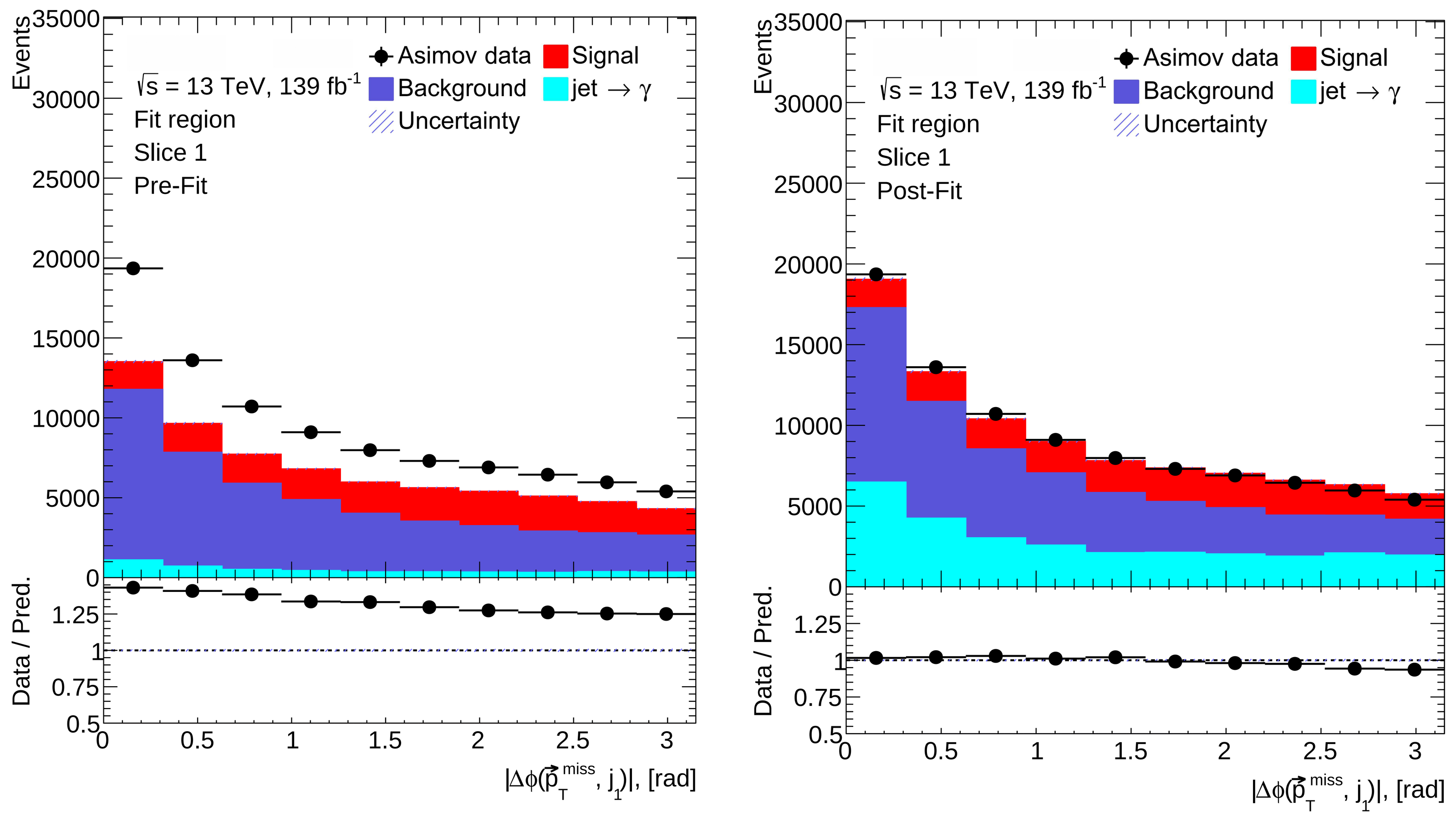
$$N^{jet\to\gamma}_{\mathrm{SR(i)}} = T_{(i)} \cdot (N^{\mathrm{data}}_{\mathrm{CR2(i)}} - N^{\mathrm{sig}}_{\mathrm{CR2(i)}} - N^{\mathrm{bkg}}_{\mathrm{CR2(i)}}).$$

Method Application

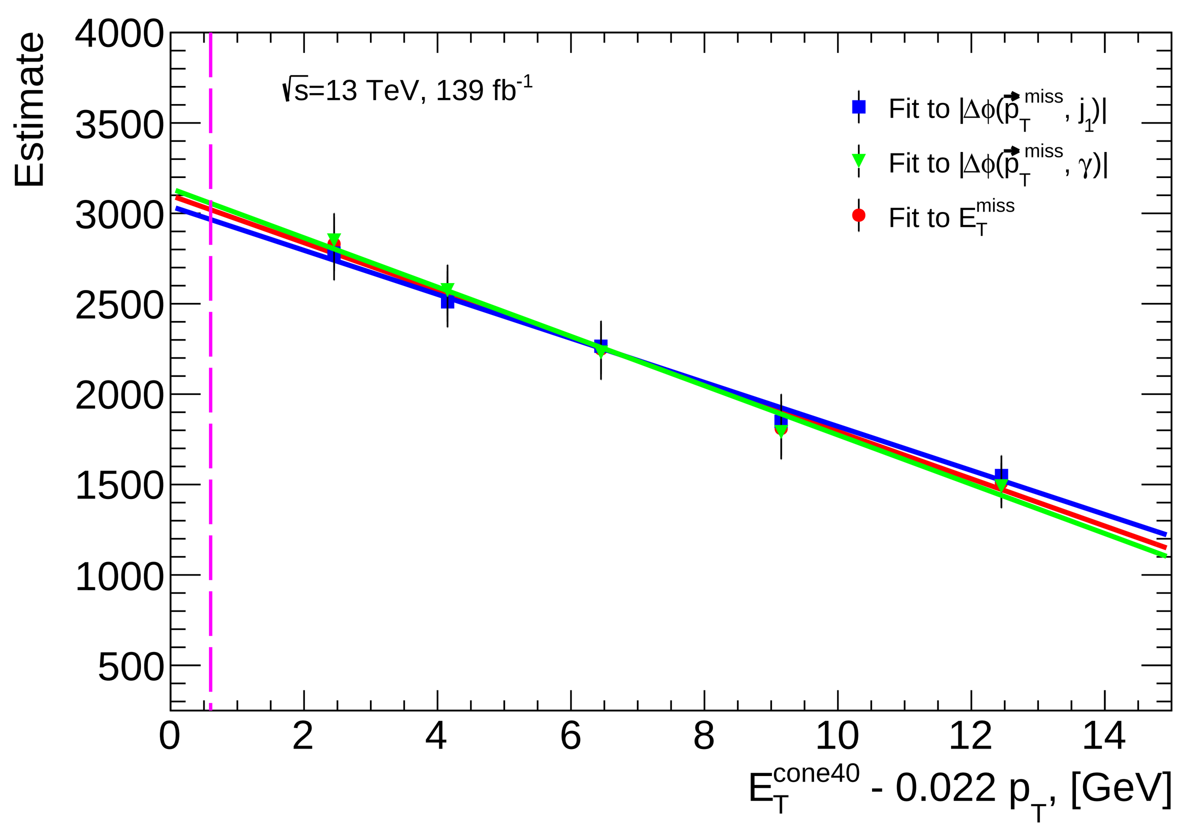
The proposed method is applied to the estimation of the $jet \to \gamma$ background to the associated production of $Z(\to \nu\bar{\nu})\gamma$. In addition to the signal and the estimated background ($\approx 12\%$ of all background events), the $\gamma + j$ process is taken into account as the main background ($\approx 33\%$ of all background events). The «Asimov» simulated dataset is used to evaluate the performance of the method. It is the sum of the MC backgrounds and the signal. The phase space of the SR used in the analysis is defined in the table below, where j_1 is the hadronic jet with the highest transverse momentum:

Selection	Cut value
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 130 GeV
E_{T}^{γ}	> 150 GeV
Number of tight photons	$N_\gamma = 1$
Lepton veto	$N_e = 0,\ N_\mu = 0$
$\lvert\Delta\phi(\overrightarrow{p}_{\mathrm{T}}^{\mathrm{miss}}, \gamma)\rvert$	> 0.7
$\lvert\Delta\phi(\overrightarrow{p}_{\mathrm{T}}^{\mathrm{miss}}, j_1)\rvert$	> 0.4

The processes considered in the study are generated using MadGraph 5 [1] MC event generator for pp-collisions with $\sqrt{s} = 13$ TeV and the integrated luminosity of 139 fb^{-1}. Pythia 8 [2] is used for parton showering and hadronization, Delphes [3] is used for detector simulation with the ATLAS experiment geometry [4]. Five isolation slices are chosen: [2.45, 4.15, 6.45, 9.15, 12.45]. The fit is performed for $E_{\mathrm{T}}^{\mathrm{miss}}$, $|\Delta\phi(\overrightarrow{p}_{\mathrm{T}}^{\mathrm{miss}}, \gamma)|$ and $|\Delta\phi(\overrightarrow{p}_{\mathrm{T}}^{\mathrm{miss}}, j_1)|$ variables. The pictures below present the pre-fit and post-fit for $|\Delta\phi(\overrightarrow{p}_{\mathrm{T}}^{\mathrm{miss}}, j_1)|$.



The liner extrapolation to the SR is shown below:



The extrapolation target is defined as the expected value of the $jet \to \gamma$ events in the SR. It is equal to 0.6, while the standard deviation $\sigma = 0.4$. Thus, the estimate of the $jet \to \gamma$ events in the SR is 3020 ± 80. Systematic uncertainties come from the choice of the extrapolation target and the set of variables. Finally, the estimate of the $jet \to \gamma$ events in the SR derived by the slice method is 3020 ± 80 (stat.) $\pm$ 120 (syst.). The MC prediction is $N_A^{jet\to\gamma} = 3090 \pm 180$ events.

Conclusion

The slice method for the estimation of $jet \to \gamma$ events is performed. The new approach considers the event distribution shapes in all regions and takes into account the dependency on isolation, which increases the estimation accuracy. The result obtained using simulated «Asimov» data coincides with the actual MC prediction within the uncertainty.

Acknowledgments

This work is supported by the Russian Science Foundation under the grant 21-72-10113.

References:
1. Allwall J., et. al. // JHEP. 2014. V.2014. №07. P.079.
2. Sjostrand T., et. al. // Comput. Phys. Commun. 2015. V.191. P.159-177.
3. de Favereau J., et. al. // JHEP. 2014. V.2014. №2. P.057.
4. Aad G., et. al. // JINST. 2008. V.3. №08. P. S08003.

katerina.kazakova@cern.ch

[a] Isolated photon satisfies $E_{\mathrm{T}}^{cone40} - 0.022p_{\mathrm{T}} < 2.45$ GeV, where E_{T}^{cone40} is the energy deposited in EM calorimeter inside Δ_R cone of 0.4 except the energy of photon itself. Photon candidates passing the tight ID criteria are required to satisfy all constraints on shower-shape variables in the electromagnetic calorimeter.