IL NUOVO CIMENTO **44 C** (2021) 53
DOI 10.1393/ncc/i2021-21053-2

Communications: SIF Congress 2020

Beam-induced background identification with image vision techniques at ATLAS

I. Longarini

Sapienza Università di Roma and INFN Roma - Roma, Italy

received 14 January 2021

Summary. — At the LHC the interaction between the proton beam and the residual gas in the beam pipes originates secondary particles which, in the case of high-energy muons, are able to cross longitudinally the ATLAS detector and leave jet-like signatures in the calorimeter. Many new physics searches are affected by this beam-induced background. This study shows new ideas, based on image vision techniques, to suppress this source of background.

1. – Introduction

Proton beams at colliders like the LHC are affected by inevitable losses, due to the scattering with residual gas in the beam pipes and the interaction of protons with the machine components.

These two effects are sources of showers of secondary particles, which in most of the cases are unable to cross the heavy-material-shielding experiments like ATLAS [1]. High-energy muons are rather unaffected by this shielding material, but are able to cause significant energy deposits in the detector by radiative energy losses, originating the so-called Beam-Induced Background (BIB).

BIB muons crossing the detector can leave signatures reconstructed as hadronic jets, contaminating the signal region of many different analyses. For example, events with large missing transverse momentum and one hadronic jet are exploitable to search for new physics signatures, which may correspond to events in which an initial state radiation jet is recoiling against invisible beyond-Standard Model particles [2]. Other unconventional signatures can be mimicked by BIB events, as in the case of long-lived particles (LLP), produced in the LHC interaction point and decaying in the outermost regions of the hadronic calorimeter in pairs of light leptons or hadrons; in this case, one or more energy deposits by BIB muons can give rise to jets with low energy deposits in the electromagnetic calorimeter and no corresponding tracks in the inner detector, matching the expected signature of a LLP scenario [3].

2. – BIB rejection methods in ATLAS

BIB muons entering ATLAS are moving almost parallel to each of the two proton beams, crossing the detector from the outside. ATLAS has a right-handed coordinate system, where the origin is at the interaction ponit, the x-axis points towards the LHC centre, the y-axis points upwards and z is parallel to the proton beams. A cylindrical coordinate system is also used (r, ϕ, θ), where the azimuthal angle ϕ is on the x-y plane and θ is the angle on the r-z plane. The pseudorapidity η is defined as $\eta = -\ln \tan(\theta/2)$. An extensive description of the ATLAS experiment and its subdetectors is given in [1].

Since the ATLAS magnetic field does not bend the path of these muons in the ϕ coordinate, it is possible to identify an event containing BIB muons via a matching, in the ϕ and r coordinates, between eventual energy deposits left in one of the calorimeters and corresponding hits in the muon spectrometer on one or both the two end-caps of the detector.

Fake jets by BIB can have a very little fraction of energy released in the electromagnetic calorimeter (EMF) and are usually reconstructed without matching tracks in the inner detector. A tight cut on EMF or the jet f_{ch}, defined as the ratio between the scalar sum of the transverse momentum (p_T) of tracks associated with the jet and the jet p_T, can be a powerful discriminant between jets from collisions and BIB.

Objects reconstructed in ATLAS are also associated with a timing information (t), corrected by the expected time of flight of particles originating at the interaction point. Hence collisions products are expected to leave hits with timing $t \simeq 0$, while BIB muons will leave hits corrected by the factor $t_{\text{tof}} = \pm\sqrt{r^2 + z^2}/c$, where z and r are, respectively, the longitudinal and the radial coordinates of the hit and the sign is determined by the BIB direction. This opens the possibility to reject some of the BIB jets with a timing cut.

These criteria, described with more details in [4], are sufficient to discriminate between fake jets from BIB and collision jets, while are not efficient against long-lived particles, which can be produced with lifetimes large enough to decay in the hadronic calorimeter, originating jets with low EMF and $f_{ch} = 0$. While a jet timing cut is still efficient to remove BIB, those jets, which after the time-of-flight correction have $t = 0$, are unaffected. Collision products in the end-cap may also originate calorimetric deposits with matching tracks in the muon spectrometer, reducing the purity of a BIB rejection method based on this matching. In the following section, an alternative BIB rejection method, based on a convolutional neural network, is described.

3. – BIB rejection with convolutional neural networks

Since BIB muons are moving almost parallel to the beam pipes, it is expected that their energy deposits in the calorimeters will have different geometrical shapes and distribution, when compared to the ones originated by collision products. The idea under this work is to train a Convolutional Neural Network (CNN) to exploit the geometry of the calorimetric deposits associated to a jet, in order to discriminate between two classes, providing as output the probability that an input jet is BIB-like.

CNNs are widely used for image-based classification tasks, as a way to extract translation-invariant features from images and learn discriminant variables for a classification problem [5].

For this BIB identification problem, the CNN inputs are built starting from calorimetric clusters associated to jets reconstructed in ATLAS. Each jet cluster is a point in the pseudorapidity-azimuthal (η-ϕ) plane and corresponds to a certain amount of energy released in a given layer of the detector's calorimeter. The two coordinates η, ϕ and the calorimeter layer are used to create three 3D grids around the jet axis: η and ϕ are divided in a 15×15 grid, corresponding to an $\eta \times \phi = 0.45 \times 0.45$ region around the jet axis, while the third dimension is different for each of the 3D grids, corresponding to the layers of the barrel, tile-extended barrel and end-cap regions of the calorimeter. In each of these grids, a cell is set to 1 if there is energy released in the corresponding region, 0 otherwise. Jet coordinates in the η, ϕ plane are also given as input to the CNN.

In order to test the CNN performance on the BIB jet identification problem, the following three samples have been produced:

- *BIB jets* are obtained by data collected by ATLAS, exploiting a trigger which selects events in which at least one jet with EMF < 0.06 is present [6]. Only the triggering jets are used, when satisfying $p_T > 20$ GeV and $|\eta| \leq 2.5$.
- *QCD jets*, obtained by the Monte Carlo simulation of QCD mulitijet events, satisfying $EMF < 0.4$, $p_T > 20$ GeV and $|\eta| \leq 2.5$.
- *LLP jets* from simulated Monte Carlo events with light invisible long-lived particles of mass $m = 400$ MeV, selected by a similar trigger to the one used in the BIB jets case, which includes additional requirements to reduce BIB contamination. Reconstructed jets are requested to satisfy $EMF < 0.4$, $p_T > 20$ GeV and $|\eta| \leq 2.5$.

The CNN has been trained to discriminate between two classes at a time: BIB jets *vs.* QCD jets and BIB jets *vs.* LLP jets. In both cases, the network is trained to output a score defined between 0 and 1, assigning 0 to BIB jets and 1 to QCD (or LLP) jets. Figures 1(a) and (b) show the normalised distributions of the CNN output variable, for jets not used during the training step, in the case of a CNN discriminating QCD and BIB jets, or LLP and BIB jets, respectively.

Since jets produced by the decay of LLPs in the outer region of the calorimeter can have similar features of BIB jets, it is expected that the CNN separation power will be worse when classifying LLP and BIB jets, with respect to the case in which the network is trained to distinguish between QCD and BIB jets. In fig. 1(a), the QCD jets CNN output distribution is showing that more events are correctly classified than in the case of LLP jets, visible in fig. 1(b). By assuming a desired efficiency on the QCD (LLP) jet sample at 0.95, the BIB jets rejection factors are 0.882 in the case of BIB *vs.* QCD discrimination and 0.612 when discriminating BIB and LLP.

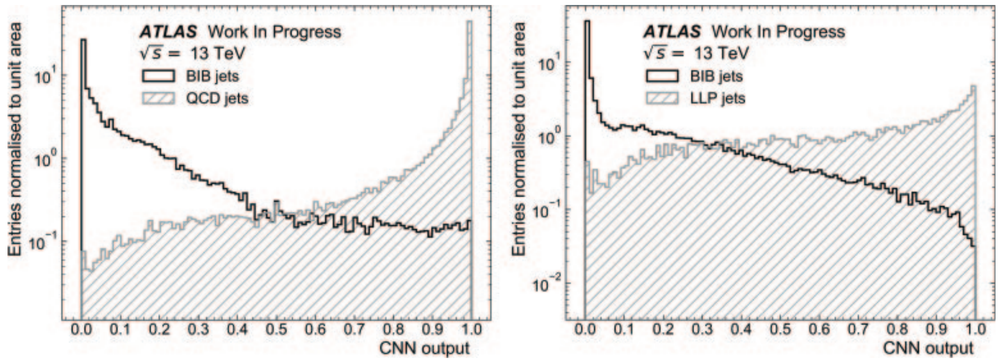

Fig. 1. – Normalised CNN output distributions for two different network trainings, only jets not used at training steps are shown. Panel (a) shows the two distributions of the CNN output for BIB jets (black) and QCD jets (hatched grey). Panel (b) shows the two distributions for the same network, trained to discriminate BIB jets (black) and LLP jets (hatched grey).

4. – Conclusions

In conclusion, the task of BIB rejection can be accomplished by various approaches based on the differences between jets from collision processes and jets from BIB, briefly summarised in sect. **2**. A CNN-based tagger relying on jet constituent information can be an additional powerful option to remove events with fake jets, alongside other BIB identification methods.

REFERENCES

[1] ATLAS Collaboration, *J. Instrum.*, **3** (2008) S08003.
[2] ATLAS Collaboration, *J. High Energy Phys.*, **2018** (2018) 126.
[3] ATLAS Collaboration, *Eur. Phys. J. C*, **80** (2020) 450.
[4] ATLAS Collaboration, *J. Instrum.*, **8** (2013) P07004.
[5] de Oliveira L. *et al.*, *J. High Energy Phys.*, **2016** (2016) 69.
[6] ATLAS Collaboration, *J. Instrum.*, **8** (2013) P07015.