Performance of heavy flavour identification algorithms in proton-proton collisions at 13 TeV at the CMS experiment

The CMS Collaboration

Abstract

Many measurements as well as searches for new physics beyond the standard model at the LHC rely on the efficient identification of heavy flavour jets, i.e. jets containing b or c hadrons. In this Detector Performance Summary the performance of these algorithms is presented, based on proton-proton collision data recorded by the CMS experiment at 13 TeV in 2016, corresponding to an integrated luminosity of 36 fb\(^{-1}\) and compared with expectations based on simulation. Correction factors for a different heavy flavour identification performance in data and simulation are evaluated. As the excluded mass regions for new physics continue to increase, searches often focus on boosted final states characterized by particles with large transverse momenta. The CMS Collaboration has developed dedicated b jet identification tools for this challenging environment. The performance of these algorithms is also presented.
Performance of heavy flavour identification algorithms in proton-proton collisions at 13 TeV at the CMS experiment

The CMS Collaboration

CMS-POG-CONVENERS-BTAG@CERN.CH
Abstract

Many measurements as well as searches for new physics beyond the standard model at the LHC rely on the efficient identification of heavy flavour jets, i.e. jets containing b or c hadrons. In this Detector Performance Summary the performance of these algorithms is presented, based on proton–proton collision data recorded by the CMS experiment at 13 TeV in 2016, corresponding to an integrated luminosity of 36 fb$^{-1}$ and compared with expectations based on simulation. Correction factors for a different heavy flavour identification performance in data and simulation are evaluated. As the excluded mass regions for new physics continue to increase, searches often focus on boosted final states characterized by particles with large transverse momenta. The CMS Collaboration has developed dedicated b jet identification tools for this challenging environment. The performance of these algorithms is also presented.
CSVv2: Combined Secondary Vertex version 2 algorithm, based on secondary vertex and track-based lifetime informations, it is an updated version of the CSV algorithm used in Run 1 combining the variables with a neural network instead of a likelihood ratio and the secondary vertex information is obtained with the Inclusive Vertex Finder algorithm.

CSVv2L, CSVv2M, CSVv2T: CSVv2 algorithm at the loose, medium, tight operating points, defined as the values of the discriminator cut for which the rate for misidentifying a light jet as a b jet is 10%, 1%, and 0.1%, respectively.

DeepCSV: a new algorithm based on the same set of observables used by the CSVv2 b-tagger, with a simple extension to use more charged particle tracks. This algorithm is based on a deep neural network training, with four hidden layer (i.e. six layers altogether) of a width of 100 nodes each.

DeepCSVL, DeepCSVM, DeepCSVT: DeepCSV algorithm at the loose, medium, tight operating points, with the same definition as for CSVv2 (see above)

c-tagger: a c jet identification algorithm exploiting properties related to displaced tracks, secondary vertices, and soft leptons inside the jets. The training of the classifiers is performed using a Gradient Boosting Classifier (GBC). Two separate GBCs are provided, one for discriminating c jets from light jets (CvsL) and one for discriminating c jets from b jets (CvsB).

JP: Jet Probability algorithm, based on the likelihood of tracks to come from the primary vertex (using the impact parameter significance values).

Softdrop subjet CSVv2: application of the CSVv2 algorithm in boosted jet topologies, where the tagging requirement is applies to the reconstructed softdrop subjets of AK8 jets.

Double-b tagger: a dedicated algorithm for identification of the decay of a boosted object to a b quark pair, for example for H → b̅b̅0 decays.
Glossary: Flavour tagger performance measurements in data

PtRel: Method for the measurement of the b-tagging efficiency in multijet events based on the transverse momenta of muons w.r.t. the jet axis.

System8: Method for the measurement of the b-tagging efficiency in multijet events with a muon, solving a system of 8 equations.

LT: Lifetime Tagging method for the measurement of the b-tagging efficiency in multijet events, based on template fits to the JP distributions.

mu+jets: Measured b-tagging efficiency in multijet events with a muon, based on the combination of the results from different measurements, obtained using the PtRel, the LT and the System8 methods.

Kin: Method for the measurement of the b-tagging efficiency in ttbar events in the dileptonic channel, based on a template fit to an MVA discriminator combining kinematic variables.

TnP: Method for the measurement of the b-tagging efficiency in ttbar events in the semileptonic channel. The b-tagging efficiency is measured with a tag and probe method (TnP). As a tagging requirement, the CSVv2M requirement is applied to either the b-jet on the hadronic or leptonic side, while the b-jet from the other side is used as probe.

TagCount: Method for the measurement of the b-tagging efficiency in ttbar events in the dileptonic channel. The b-tagging efficiency is obtained by counting the number of events with two b-tagged jets in the selected sample of events.

IterativeFit: Method for the measurement of the b-tagging efficiency in ttbar events in the dileptonic channel. This method is based on the calibration of the full b-tagging discriminator shape.
Commissioning studies
commissioning on AK4 jets: Different event topologies

- **Inclusive multijet sample**: Events with online selection requiring at least one AK4 jet with a $p_T > 40$ GeV. Data to simulation comparison is performed for jets with a $p_T$ between 50 and 250 GeV. This topology is dominated by light-flavour jets and has a relatively large contribution form pile-up jets.

- **Muon enriched jet sample**: Events with online selection requiring at least two AK4 jets with a $p_T > 40$ GeV of which at least one contains a muon with a $p_T > 5$ GeV. Data to simulation comparison is performed for jets with a $p_T$ between 50 and 250 GeV and containing a muon. This topology is enriched by heavy-flavour jets.

- **Dilepton $t\bar{t}$ sample**: Events with online selection requiring at least one isolated electron and at least one isolated muon. Events are further selected by requiring at least two AK4 jets with a $p_T > 20$ GeV and a muon and electron with $p_T > 25$ GeV. This topology is dominated by $b$-jets from the top quark decays.

- **Single lepton $t\bar{t}$ sample**: Events with online selection requiring at least one isolated lepton, and requiring offline exactly one isolated electron ($p_T > 40$ GeV) or muon ($p_T > 30$ GeV). Events are further selected by requiring at least four AK4 jets with a $p_T > 25$ GeV and well separated from the isolated lepton. This topology is expected to have a higher fraction of $c$ jets (from the hadronic $W$ decay) compared to the other topologies.

→ In all data-to-simulation comparisons, the simulated distributions are normalized to the data.
The impact parameter significance of the tracks in jets from the dilepton $t\bar{t}$ sample. The IP significance is defined as the value of the impact parameter divided by its uncertainty. Different colors show the contributions in simulations from different jet flavours. Some small discrepancies are observed due to the sensitivity of this variable to the tracker alignment and the uncertainty on the track parameters.

The corrected secondary vertex mass for the leading secondary vertex (sorted according to increasing uncertainty on the 3D flight distance), for jets in an inclusive multijet sample. The secondary vertex mass is corrected for the observed difference between its flight direction and its momentum obtained by summing the momentum of its associated tracks. Different colors show the contributions in simulations from different jet flavours.
The flight distance significance for the leading secondary vertex (sorted according to increasing uncertainty on the 3D flight distance) for jets in a muon enriched multijet sample. The flight distance significance is defined as the value of the flight distance divided by its uncertainty. Some small discrepancies are observed due to the sensitivity of this variable to the tracker alignment and the uncertainty on the track parameters. Different colors show the contributions in simulations from different jet flavours.

The “MassVertexEnergyFraction” provides information on the ratio of the energy contained in the secondary vertex with respect to the total jet energy (multiplied by the ratio of the secondary vertex mass and the average charged B hadron mass [1]). The distributions are shown for jets in the single lepton $t\bar{t}$ sample. Different colors show the contributions in simulations from different jet flavours.

The charm tagger consists of two binary classification algorithms, one to discriminate charm jets from light jets (CvsL) and one to discriminate charm jets from bottom jets (CvsB). The top panel shows the discriminator distributions for the CvsL discrimination whereas the lower panel shows the discriminator distributions for the CvsB discrimination, both for jets in the inclusive multijet sample. The peaked structures in the distributions are caused by jets for which no tracks pass the track selection criteria and therefore all discriminating jet properties are missing (put to default values). Different colors show the contributions in simulations from different jet flavours.
commissioning on AK8 jets:
Different boosted event topologies

- **Boosted muon enriched subjet sample**: Events from a combination of single jet (AK4 and AK8) triggers requiring a muon inside the jet. Data to simulation comparison is performed for soft drop subjets of AK8 jets with a $p_T > 350$ GeV. The subjets are required to have at least one muon ($p_T > 7$ GeV) close to the subjet axis ($\Delta R < 0.4$) and which carries less than half of the total subjet momentum.

- **Boosted double-muon tagged jet sample**: Events are obtained by combining the triggers mentioned above with dijet triggers requiring a muon in at least one of the two jets. This allows for a lower $p_T$ threshold on the AK8 jets ($p_T > 250$ GeV). The subjets are required to have at least one muon ($p_T > 7$ GeV) close to the subjet axis ($\Delta R < 0.4$). The sum of the transverse momenta of the two muons with respect to the transverse momentum of the AK8 jet is required to be less than 0.6.
The CSVv2 discriminator distribution for soft drop subjets of muon-tagged AK8 jets with $p_T > 350$ GeV in muon enriched multijet samples. Different colors show the contributions in simulations from different jet flavours.

The double–$b$ discriminator distribution for double-muon tagged AK8 jets with $p_T > 250$ GeV in muon enriched multijet samples. The requirement of having two muons in the jet makes this topology sensitive to the production of two bottom or charmed hadrons from gluon splitting. Different colors show the contributions in simulations from different jet flavours.
Efficiency and scale factor measurements
Efficiency and scale factor measurements: DeepCSV misidentification probability

The misidentification probability measurement for the medium working point of the DeepCSV algorithm. This measurement uses the negative tag method performed on a sample of inclusive multijet events. In the top panel the misidentification probability in data and simulation is presented as a function of the jet $p_T$. The bottom panels show the scale factor for light-flavour jets, where the solid curve is the result of a fit to the resulting scale factor values as a function of the jet $p_T$ and the dashed lines represent the overall statistical and systematic uncertainty on the measurement. The scale factors are typically larger than one in a broad jet $p_T$ range. The scale factors shown in this and in the following slides show deviations from unity, as the quantities of relevance for heavy flavor identification are not perfectly modeled by simulations.
Comparison of the b jet scale factor (SFb) for the tight working point of the CSVv2 tagger. The combination of the measurements performed with methods based on muon enriched multijet events (LT, ptre1, system8) is presented in blue squares, while the combination of SF from ttbar events (with the Kin and TnP methods) is displayed with red bullets. Also shown is the full combination of muon-enriched and ttbar measurements including an additional 1% uncertainty to cover any residual sample dependence. The combined SFb value with its overall uncertainty is displayed as a green area. To increase the visibility of the measurements, the resulting scale factors on ttbar and muon enriched multijet events are slightly displaced with respect to the bin center for which the measurement was performed. The last bin includes the overflow.
The fitted Kin distribution for jets passing (left) and failing (right) the CSVv2M tagging requirement and with jet $p_T$ between 70 and 100 GeV. The Kin Method is performed on a very pure dileptonic $t\bar{t}$ topology with exactly one isolated muon and one isolated electron with an invariant mass larger than 90 GeV and transverse missing energy larger than 40 GeV. A kinematic discriminator is constructed that can separate $b$ jets from non-$b$ jets purely from kinematic properties (invariant masses and angular separations) of the event. The kinematic discriminator distribution is unrolled in bins of jet multiplicity with the discriminator output transformed from $[-1, 1]$ to $[-1, 1]+2 \cdot (N_{\text{jets}} - 2)$. A binned likelihood fit is performed simultaneously on the kinematic discriminator shapes for jets passing and failing the $b$-tagging requirements for different working points and taggers.
The fitted $-\text{Log}(\lambda)$ distribution of the Tag and Probe (TnP) method for jets from the leptonic top decay with $p_T$ between 70 and 100 GeV passing (left) and failing (right) the medium working point of the CSVv2 tagger.

The TnP method is performed on a semi leptonic $t\bar{t}$ topology with exactly four jets with $p_T>30$ GeV. These four jets are assigned to the quarks (a b from the leptonic top decay, a b from the hadronic top decay and two light quarks from the hadronic W decay) based on a likelihood discriminant ($-\text{Log}(\lambda)$). One of the two (hadronic or leptonic) b jets should be tagged by the CSVv2 medium working point, while the other serves as a probe to measure the tagging efficiencies of a certain working point from a given tagger by performing a template fit.
The upper panels show the scale factor for c jets ($S_{F_{c}}$) for the loose operating points of the c tagger, measured in a semi leptonic $t\bar{t}$ (blue) and $W+$charm (red) topology. Due to the limited statistics the $t\bar{t}$ based method is not binned in $p_{T}$. The thick error bars represent the statistical error and the narrow error bars the combined statistical and systematic uncertainties. The combined scale factor value with its overall uncertainty is displayed as a hatched area.

The lower panels show the same combined scale factor value with the result of a constant fit function superimposed (solid curve). The combined statistical and systematic uncertainty is centered around the fit result, represented by the points with error bars. The last bin includes the overflow.
Efficiency and scale factor measurements: Boosted topologies

CSVv2 b jet scale factors ($SF_b$) for the loose working point, obtained with the Lifetime Tagging (LT) method for AK4 b jets (green) and softdrop subjets of AK8 jets (red) as a function of the (sub)jet transverse momentum. The hatched area shows the combined statistical and systematic uncertainty. The scale factors derived for the two different jet types agree well.

Scale factor ($SF_{\text{double } b}$) as a function of the jet $p_T$ for the loose operating point of the double-b tagger measured using a sample of double-muon tagged AK8 jets. The measurement is again performed using the LT method. The hatched area shows the combined statistical and systematic uncertainty. The measurement is done for jets with $p_T$ between 250 and 840 GeV, which is driven by the size of the available data sample. Jets with a larger $p_T$ are included in the last $p_T$ bin.