PROCEEDINGS OF SCIENCE

# Heavy flavor jet tagging algorithm developments at CMS for HL-LHC

**Emanuel Pfeffer,**[a,*] **Daniel Bloch,**[b] **Ulrich Husemann,**[a] **Soureek Mitra,**[a] **Thomas Muller,**[a] **Max Neukum,**[a]
**On behalf of the CMS collaboration**

[a] *Karlsruhe Institute of Technology,*
*Kaiserstr. 12, 76131 Karlsruhe, Germany*
[b] *Institut pluridisciplinaire Hubert Curien,*
*23 rue du loess, BP 28, 67037 Strasbourg Cedex 2, France*

*E-mail:* emanuel.pfeffer@kit.edu, max.neukum@kit.edu

The rich physics program at the high luminosity LHC (HL-LHC) requires all final state particles to be reconstructed with good accuracy. However, it also poses formidable challenge of dealing with very high pileup. Different identification algorithms need to be upgraded along with the detectors to improve the overall event reconstruction in such a hostile collision environment. The new timing device in the proposed CMS detector at the HL-LHC allows for the construction of timing observables at the track-level as well as at the jet-level. This information when given as inputs to the deep neural networks, have a potential to improve the existing algorithms used for heavy flavor (HF) jet tagging. In this paper, the latest developments on the studies for HF jet tagging performance at the HL-LHC are presented.

*The Tenth Annual Conference on Large Hadron Collider Physics - LHCP2022*
*16-20 May 2022*
*online*

*Speaker




PoS(LHCP2022)270

## 1. CMS upgrade for HL-LHC

While the Large Hadron Collider (LHC) is currently operating in Run-3, preparations are in full swing for the subsequent operation that is planned for the year 2026. In this new phase, referred to as the high luminosity era (HL-LHC), the colliding beams will be even more narrowly focused. This will result in an increased collision rate and therefore higher activity around the collision point that the detectors have to cope with. With the increased collision rates of the HL-LHC, up to 200 simultaneous proton-proton collisions in addition to the collision of interest (pileup) are expected. To handle the challenges caused by the pileup, three major improvements are planned for an upgraded CMS detector: A tracker coverage up to $|\eta| \leqslant 4.0$ [1], a silicon-based calorimeter with high granularity and better resolution [2], and a MIP Timing Detector (MTD) [3] providing timing information to mitigate spurious tracks arising from pileup.

All three components are particularly interesting for jet flavor tagging, i.e. to distinguish udsg- (light-flavor), c- and b-jets. Chapter 2 explores the implications that these detector improvements have for a jet tagging algorithm at CMS.

## 2. Tagging performance

The DeepCSV algorithm [4] is a feed forward neural network, that is able to distinguish between b-, c- and light-flavor-jets. The network is composed of an input and an output layer with 5 hidden layers in between, each consisting of 100 nodes. Track-based lifetime information and secondary vertex information obtained with the Inclusive Vertex Finder (IVF) algorithm are fed as inputs. The output layer consists of a total of 4 nodes, namely, b, bb, c and light-flavor; whose values indicate the probabilities to be identified in the respective category for a jet.

In the following, the tagging performance in four different scenarios is examined. The receiver-operator-characteristics (ROC) curve is used as a metric, which shows the misidentification probability over the b-jet efficiency. This reveals the quality of the separability with which a b-jet can be distinguished from a c-jet or a light-flavor-jet. The four scenarios studied differ as follows:

- no pileup and no MTD
- high pileup of 200 but without MTD
- high pileup of 200 and a MTD with a time resolution of 30 ps
- high pileup of 200 and a MTD with a time resolution of 60 ps.

In these scenarios, “and a MTD” means that track and secondary vertices are reconstructed with corresponding information from the MTD and are thus refined from spurious tracks originating from pileup vertices. This would be a CMS detector built as of now without using the MTD at the new expected high pileup of 200. The reason for the two different time resolutions with the proposed MTD is having a time resolution of 30 ps at the beginning of the detector operation, which is expected to degrade to 60 ps over the run time.

The performance of the DeepCSV algorithm is evaluated using simulated $t\bar{t}$ events with jets reconstructed using the anti-$k_T$ algorithm with cone size 0.4 [5]. All jets are required to have at

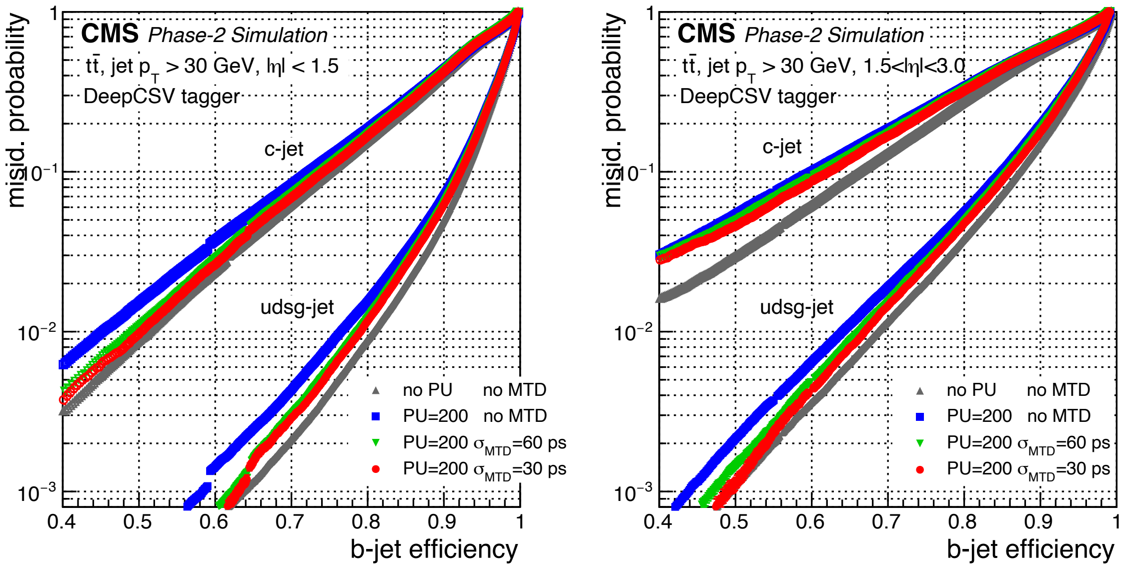


**Figure 1:** b-jet tagging efficiency for DeepCSV for $|\eta| < 1.5$ (left) and $1.5 < |\eta| < 3.0$ (right). A hypothesis without pileup (gray) is compared to cases with pileup but without timing from MTD (blue), and pileup with additional timing with a resolution of 30 (red) and 60 ps (green), respectively [3].

least a transverse momentum ($p_{\mathrm{T}}$) of 30 GeV. The performance evaluation is done separately for barrel ($|\eta| \leq 1.5$) and endcap ($1.5 < |\eta| < 3.0$).

The resulting ROC curves are shown in Figure 1. It is prominent that for both b-jet vs light-flavor-jets and c- vs light-flavor-jets, the tagging efficiency increases when MTD information is used for jet tagging at all misidentification probabilities. Especially at low misidentification probabilities and consequently e.g. the tight working point (0.001), this effect can be noted well. Naturally, no performance can match the no pileup scenario, which is merely presented as an ideal case, but does not represent realistic physics events in the detector. The expected decrease of the tagging performance due to aging of the MTD can be read from the difference between the red and the blue curve. This degradation depends on the average pileup and pileup density. The b-jet tagging efficiency decreases by about 10 % at high pileup of 200 compared to the no-pileup case at 0.001 light misdentification rate. To address the effects of degradation, the secondary vertexing algorithms were modified to include timing information. This requires tracks to be within a specific time resolution of $3.5\sigma_t$ with $\sigma_t$ = 30 ps – 40 ps. As a result, the number of spurious reconstructed secondary vertices is reduced by 30 % [3].

## 3. Performance vs pileup

It is worth studying the dependence of the quality of the b-jet identification over the pileup density (events / mm). This allows, for example, to conclude whether b-jet tagging deteriorates (potentially significantly) for a high pileup density compared to events with moderate pileup. This study is presented in Figure 2. In this figure, for a fixed light-flavor-jet misidentification probability of 0.01, the b-tagging efficiency is displayed as a function of the pileup density. It is observed that the DeepCSV tagger shows a significantly lower and almost no dependence on the pileup density if

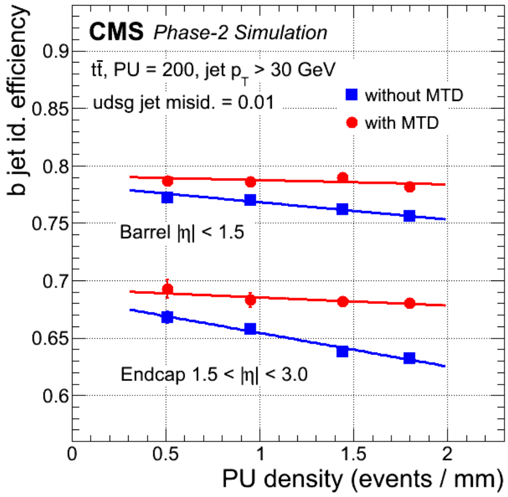


**Figure 2:** Efficiency of the b-jet identification vs the average pileup density for a scenario with and without MTD at a constant light-flavor-jet efficiency of 1 % [3].

the secondary vertex information contains MTD information compared to absence of MTD refined vertices. The latter worsens remarkably with higher pileup density, in particular in the endcap.

## 4. Summary and outlook

In order to address the challenges of the high luminosity LHC due to a high pileup of 200 simultaneous collisions, a potential improvement in the efficiency of the heavy flavor tagging algorithm through an additional MIP timing detector was investigated. It was demonstrated in this context that an improved heavy flavor jet tagging performance can be achieved with the DeepCSV discriminant. This is particularly evident at tighter working points at very low misidentification probability rates. For this purpose, barrel and endcap regions in the detector were separately investigated. It was also shown that a possible pileup density dependence does not or only slightly exist if MTD information is provided. Thus the overall impact of timing information from MTD on heavy-flavor jet tagging performance looks promising. It is now under investigation how the MTD impact can possibly affect more advanced tagging algorithms such as the DeepJet [6] and the ParticleNet [7] in the prospective high luminosity LHC era.

## References

### References

[1] CMS collaboration, *The Phase-2 upgrade of the CMS tracker*, *CERN-LHCC-2017-009, CMS-TDR-014* (2017) .

[2] CMS collaboration, *HGCAL: A high-granularity calorimeter for the endcaps of CMS at HL-LHC*, *Journal of Physics: Conference Series* **928** (2017) 012025.

[3] CMS Collaboration, *A mip timing detector for the cms phase-2 upgrade*, *CERN-LHCC-2019-003, CMS-TDR-020* (2019) .

[4] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer et al., *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*, *Journal of Instrumentation* **13** (2018) P05011.

[5] M. Cacciari, G.P. Salam and G. Soyez, *The anti-$k_T$ jet clustering algorithm*, *Journal of High Energy Physics* **2008** (2008) 063.

[6] E. Bols, J. Kieseler, M. Verzetti, M. Stoye and A. Stakia, *Jet flavour classification using DeepJet*, *Journal of Instrumentation* **15** (2020) P12012.

[7] H. Qu and L. Gouskos, *Jet tagging via particle clouds*, *Physical Review D* **101** (2020) .