ATLAS Note

GROUP-2024-XX

13th December 2024

Draft version 0.1

Measurement of the b-trigger efficiency with $t\bar{t}$ events with a likelihood method

The ATLAS Collaboration

The identification of jets resulting from the hadronization of a b quark plays a crucial role in multiple analyses performed with the ATLAS detector at the Large Hadron Collider. Incorporating b-jet identification algorithms at the trigger level significantly improves the signal efficiency for event selection in many of these analyses. This paper reports on the measurements of b-jet identification efficiency at the trigger level, as well as the conditional identification efficiencies where online identification must precede successful offline b-jet identification. This dual-stage approach aids analyses that utilize b-jet identification at both the trigger and offline stages. Efficiency measurements are conducted using a data set highly enriched in $t\bar{t}$ dileptonic events, with a likelihood method employed to ascertain the b-jet identification efficiency of the sample. For both the trigger-only and the conditional scenarios, the efficiency of the b-jet tagger and the corresponding scale factors (SF) are determined for jets spanning a transverse momentum range of 20 to 400 GeV. These efficiencies are then compared with simulation predictions by computing data-to-simulation scale factors.

Contents

1 Introduction

The trigger is a crucial step in the event selection of any physics analysis, so its performance must be understood and calibrated. For the Run 3, the trigger part has been upgraded compared with Run 2 and new b-tagging algorithms has been applied (DL1d for 2022) for online. This leading to a difference of b-trigger efficiency compared with Run 2's results. The study presented in this note is a measurement of the online b-jet identification efficiency, divided into a measurement of the b-trigger efficiency alone and a measurement of the conditional trigger efficiency of passing offline first. The method used to perform these two measurements is a variation of the likelihood method used for the measurement of the offline b-tagging efficiency for Run 3. In this note, aspects that overlap with the offline measurement, such as the detector description, object definition, simulated samples used, and offline b-tagging algorithms, will be either omitted or briefly addressed, with greater emphasis placed on the elements unique to this study.

A sample of b-jets with the highest possible purity is necessary for an accurate measurement of b-trigger efficiency. The measurement described in this note employs a method that selects a data sample enriched with $t\bar{t}$ dilepton events, aiming a phase space with a high b-jet purity(around 90%). The analysis uses ATLAS pp collision data at $\sqrt{s}$ =13.6 TeV for 2022.

Events are required to pass specific triggers that involve 1 light leptons or 1 muon and at least 2 jets, with the online b-tagging algorithm applied to the HLT jet without any additional cuts. To enhance the purity of the selection and minimize background, the offline selection criteria demand exactly one electron and one muon with opposite charge, both meeting tight identification criteria. Additionally, exactly two offline jets must be matched to online jets based on a geometrical requirement, and the invariant masses of jet-lepton pairs are required to be consistent with the top quark mass to further improve the purity of the selected b-jets.

The events passing the selection are categorized into various groups based on the jets' pT and the outputs of the online and offline b-tagging algorithms. Simulated events are also categorized by the predicted flavor of the hadronic jets. A maximum likelihood fit is then performed to extract the b-jet identification efficiency as a function of jet pT from the data.The results of this measurement are intended to correct the b-jet identification efficiency in simulated samples and are presented as per-jet scale factors (SFs).

This note is organized as the following:in chapter2, we show the principle to perform th calibration;in chapter3, we show the data and MC samples used in the select; in chapter 4, give a through description about the procedure of select and in the later chapters, we get the results.

2 Measurement Strategy

The b-jet trigger is always used in tandem with offline b-tagging, which is calibrated independently without making any requirements on the b-jet HLT trigger for the trigger-only. A conditional b-jet trigger efficiency is therefore calculated with respect to the offline b-tagging efficiency and defined as the number of truth b-jets that pass the offline b-tagging requirements at a specified offline working point (GN2 for 2022, 65%,70%,77%,85%,90%) and each are matched with the corresponding b-tagged HLT jet. This conditional b-jet trigger efficiency is measured in both data and simulated ttbar events. Scale factors to correct the b-jet trigger efficiency in Monte Carlo simulations to that observed in data are also derived.

The conditional efficiency, $\epsilon_b^{tri|off}$, is defined as the efficiency of a jet to be tagged as a b-jet by the online b-tagging algorithm if it has also passed the offline b-tagging. Here, 'off' represents the offline b-tagging while 'trig' represents the online b-trigger.

The results of this measurement are given in the form of per-jet scale factors of two different types: trigger-only scale factors and conditional scale factors. The trigger-only scale factors are standalone results that can be applied to measure the b-tagging efficiency at the trigger level. The conditional scale factors are intended to be combined with the existing offline b-tagging scale factors to correct simulated events in physics analyses where both trigger and offline b-tagging are used.

In order to evaluate this conditional efficiency, only events in which both jets are already tagged by the offline b-tagging are selected and the efficiency of the online b-tagging on these events is evaluated. The ratio between the conditional efficiency measured in data versus that measured in MC is the conditional scale factor defined as:

$$SF_b^{tri|off} = \frac{\epsilon_b^{tri|off,data}}{\epsilon_b^{tri|off,MC}}$$

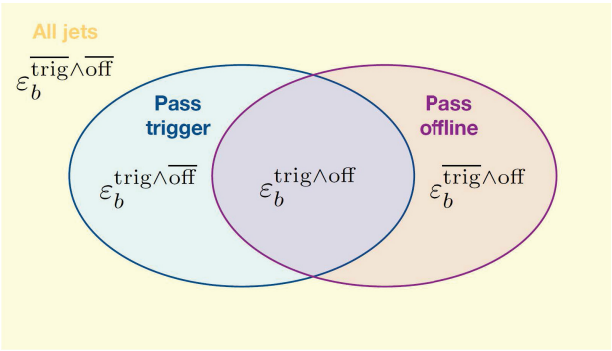


Figure 2.1: A visualisation of efficiencies should be applied in different pass/fail trigger/offline scenario

3 Data and MC samples

This study is performed using $\sqrt{s}$ = 13.6 TeV pp collision data samples taken from the year 2022 with luminosity of 26.1fb^{-1} for both trigger-only and conditional compared with offline-only with luminosity of 29.0 fb^{-1}. This is because events are required to fulfill standard data quality requirements corresponding to the 'All_Good_BjetHLT' good run list (GRL), meaning that all relevant ATLAS detector components and online b-tagging must be functioning normally,which is different from the GRL offline-only used. This GRL accounts for beamspot-related defects that are intolerable for b-jet triggers, which can result in a smaller integrated luminosity compared to the standard GRL.

Good Run List	Luminosity (fb^{-1})
offline: data22_13p6TeV.periodAllYear_DetStatus-v109-pro28-04_MERGED_PHYS_StandardGRL_All_Good_25ns_ignore_TRIGLAR.xml	29.0
trigger: data22_13p6TeV.periodAllYear_DetStatus-v109-pro28-04_MERGED_PHYS_StandardGRL_All_Good_25ns.xml	26.1

Table 3.1: Different GRL for online and offline with stricter GRL for trigger, reducing luminosity by around 10% .

Events are required to activate a lepton bofftrigger or a muon bofftrigger plus jets trigger, with lepton requirements similar to those used by the offline measurement and an additional requirement of at least 2 jets on which the online b-tagging algorithm is applied, though without making decisions. The lepton boff trigger used are listed below, with logical OR is required,as in Run2. After all requirements and selections, the analyzed data sample in 2022 contains adequate events, which are used to extract the b-jet tagging efficiencies.

Trigger Name
HLT_e26_lhtight_ivarloose_2j20_0eta290_020jvt_boffperf_pf_ftf_L1EM22VHI
HLT_mu24_ivarmedium_2j20_0eta290_020jvt_boffperf_pf_ftf_L1MU14FCH

Table 3.2: Lowest unprescaled lepton + boffperf trigger chains used for data selection in the 2022. A logical or is performed between the corresponding electron and muon trigger.

The same Monte Carlo (MC) simulated event samples used by the offline measurement [**1**] are also used here, with the same derivation format (FTAG2). The event weight in the samples is corrected to take into account the different integrated luminosities of the datasets used for this study.

4 Event Selection

Top quarks are produced in abundance at the LHC and, since the branching fraction of the top quark decay into a W boson and a b-quark is nearly 100%, selecting events with pair-produced top quarks can provide a large data sample of b-jets that can be used to study the b-jet trigger efficiency. In order to reduce the contributions from multijet and W/Z+jets backgrounds, and maximize the purity of the selection, the offline selection requires events to have exactly one electron and one muon with opposite-sign charge and satisfying tight identification criteria. Furthermore, the electron and muon provide a signature that can be used to select events at the trigger level without using a b-jet trigger such that no bias is introduced from online b-tagging. These 'single-lepton b-performance triggers' (detailed in Table 3.2) were designed and run specifically in order to study the performance of the b-jet triggers, and require the presence of an electron or muon, plus two additional jets. The b-jet trigger software is run on the jets and all associated b-tagging information is kept.

As in the offline measurement [**1**], a pseudo-top reconstruction is performed to increase the purity of the sample in bb events, i.e., events with two true b-jets. Each of the two leptons is paired to a jet in an exclusive way based on the following criterion:

$$\min_{i,j\in\{1,2\}} (m^2_{j1,li} + m^2_{j2,lj}), \tag{4.1}$$

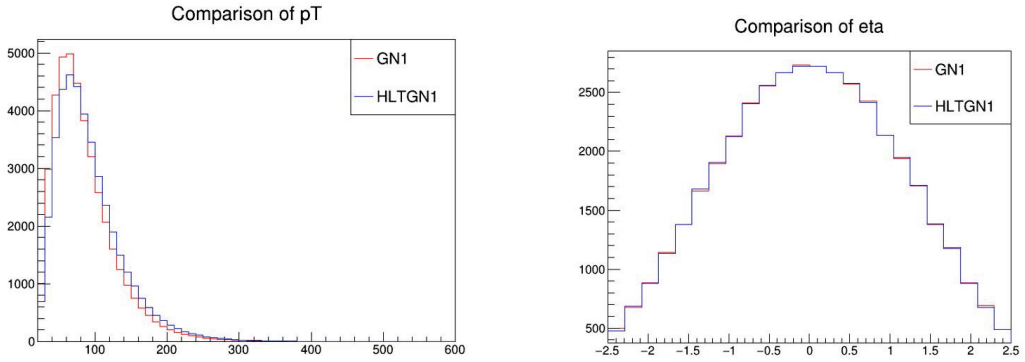
where $j1$ ($j2$) is the highest (second highest) p_T jet, $li; j$ are the two leptons, and $m_{j1,l}$ ($m_{j2,l}$) is the invariant mass of the system including the highest (second highest) p_T jet and its associated lepton. The minimum of squared masses is chosen as criterion because it is the simplest way to penalize asymmetric pairing.

The kinematics of the detector in the HLT (High-Level Trigger) section and the offline reconstruction can differ slightly for each event. This leads to minor discrepancies between the jet detection results from the trigger and the offline parts. For each event, the results obtained using offline selection versus those obtained from the trigger section will differ. Each event that satisfies the conditional requirements meets both the trigger and offline selections and has passed the offline GN2 wps filtering. Events passing the conditional requirements are compared, and the jet's η (eta) and p_T (transverse momentum) are analyzed using data from ttbar Monte Carlo simulations.

The correlation between online and offline jets was studied using a sample of approximately 200 thousand simulated $\bar{t}t$ events. Events are selected with the following criteria:

- Activate the electron trigger listed in Table 3.2.
- All offline jets are matched with one online jet.

Each jet in the selected events is used to fill 2D histograms for the p_T, η, b-tagging discriminant value, and b-tagging discriminant working point (WP), which are shown in Figs.4.2. For the two kinematic variables, the distributions are very close to the diagonals, especially the η distributions. For the b-tagging discriminant, however, the distribution is far from being diagonal. This is expected, given that two different



(a) distribution of the lead jet pT

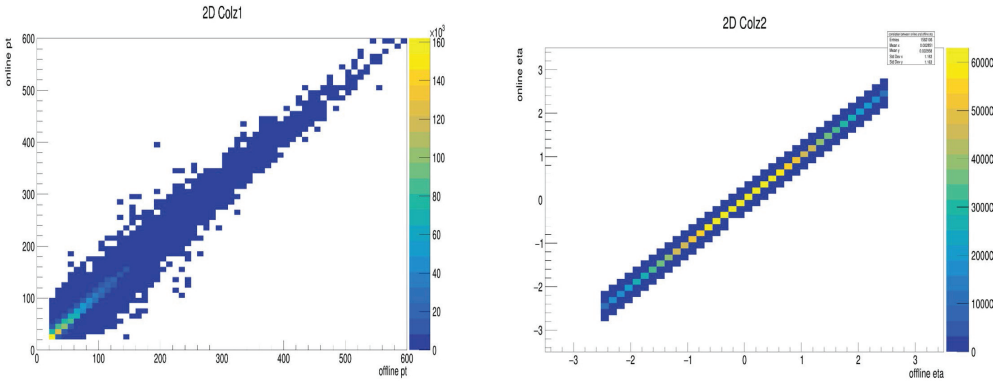
(b) distribution of the lead jet eta

Figure 4.1: plot made using ttbar MC, distribution of the lead jet kinematics for HLT jets seems very close to offline jets

versions of the b-tagging algorithm (DL1d and GN2) are used for online and offline tagging in 2022.For the kinematics,the distributions are very close to the diagonals, meaning there is high correlation for both variables.Distribution of the b-tagging discriminant values: it is far from the diagonal as expected, given that 2 different versions of the b-tagging algorithm are used.

After the selection, events are further categorized into 5 categories according to the offline b-tagging discriminant: one with no offline b-tagging applied and then one for each of the 4 working points (where the same offline working point requirements are applied to both jets). For each of the five categories, events are further classified according to the p_T of the two jets. This classification is necessary to measure the b-jet identification efficiency as a function of jet p_T. The binning used for each jet is:

20-40, 40-60, 60-140, 140-250, 250-400. GeV



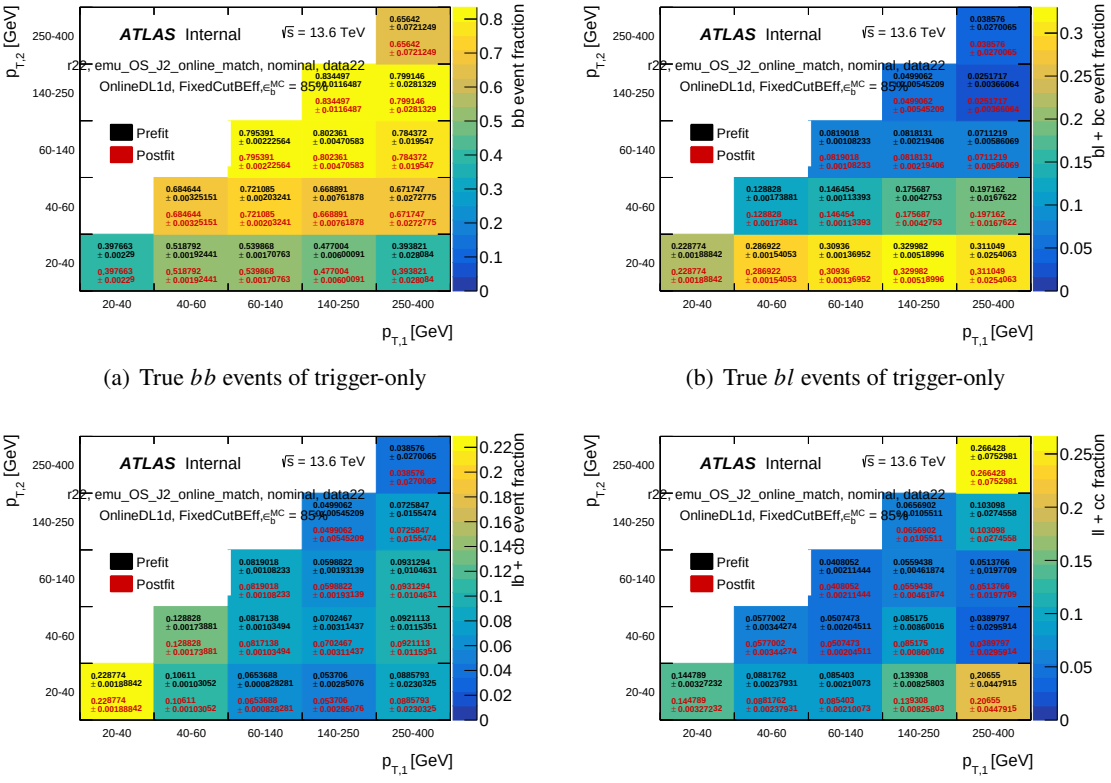
(a) pT (b) phi

(c) eta (d) b-tagging

Figure 4.2: 2D plots showing the correlation between online and offline jet kinematics

5 Backgound composition

The fraction of bb, bl, lb, and ll events for each (p_{T1}, p_{T2}) bin in the category with no online b-tagging requirement (trigger-only)is shown in Fig 5.1. As mentioned before, the select applied on the offline jets which matched with HLT jets ensure the purity of bb events in the Signal Region. As a result, the bb fraction in the selected events is low in low pt bins and increases with the increase of the leading jet and subleading jet p_T. The bl fraction is always greater than 10% for $p_T2 < 40$ GeV and is particularly high for small p_{T2} and large p_T1 values. The lb fraction is also less than 10% for $p_T2 > 60$ GeV and is greater than 10% in the bins with p_{T1} and $p_{T2} < 60$ GeV. Finally, the ll fraction shows a similar behavior to the bl one, but with a maximum of approximately 20%.



(a) True bb events of trigger-only

(b) True bl events of trigger-only

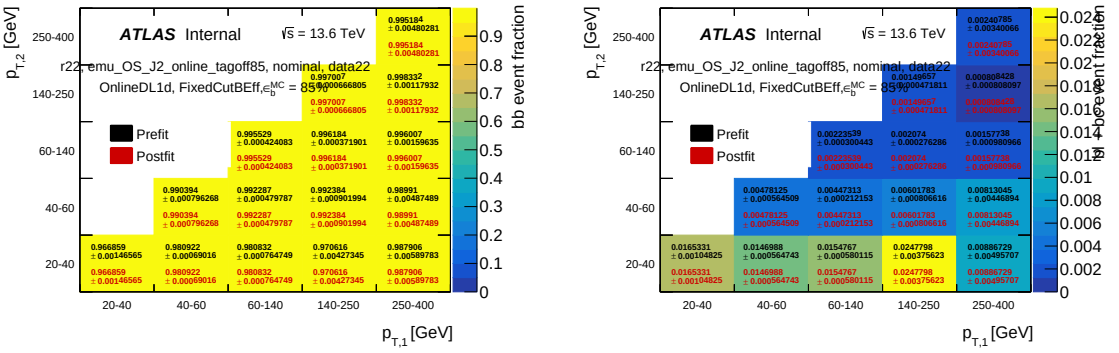
(c) True lb events of trigger-only

(d) True ll events of trigger-only

Figure 5.1: The fraction of total events in bins of the leading and subleading jet p_{T1} and p_{T2}, respectively for trigger-only, for Monte Carlo (MC) events where 'l' includes both light and charm-quark contributions.

In comparison, the conditional category has an even purer bb ratio due to the b-tagging discriminant applied on each of the two jets as shown in Fig 5.2. For the offline GN2 85% working point, the ratio of bb and bl increased compared with the trigger-only category, especially in the low p_T bin, 20–40. Among all

the p_T bins in the bb ratio, each exceeds 95%, with the maximum reaching 99%.For true ll events,in Fig 5.2.(d), few events left for the high p_T bins,lead to the empty sections in plot.



(a) True bb events of conditional offline 85% wp

(b) True bl events of conditional offline 85% wp

(c) True lb events of conditional offline 85% wp

(d) True ll events of conditional offline 85% wp

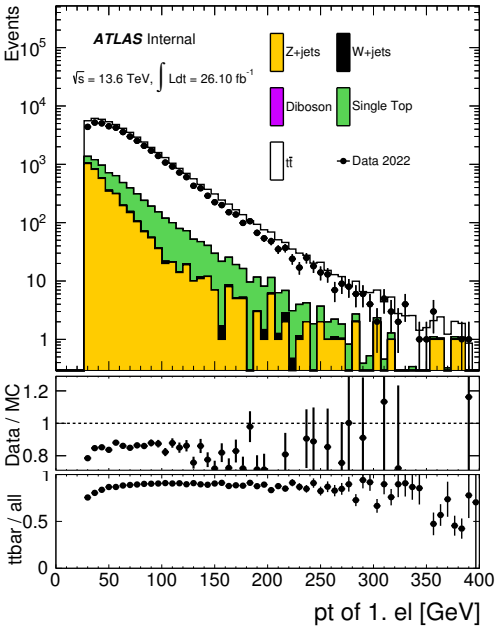
Figure 5.2: The fraction of total events in bins of the leading and subleading jet p_{T1} and p_{T2}, respectively for conditional offline 85% working point.

6 Data/Simulation comparison plots

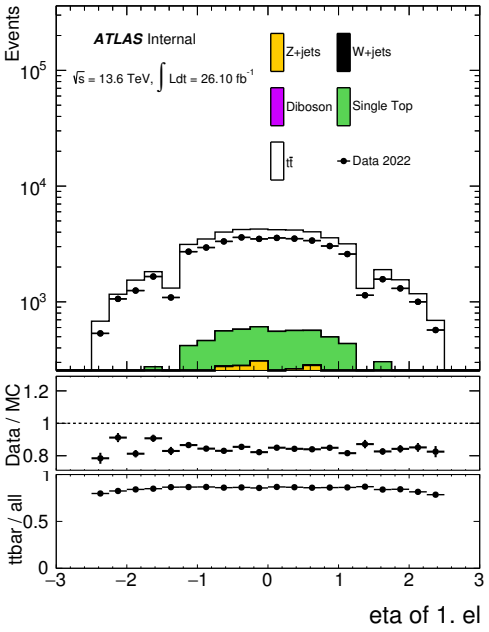
As already explained in Section 4, after the selection, events are further categorized into 6 categories according to the offline b-tagging discriminant (GN2 used for 2022): one with no offline b-tagging applied and then one for each of the 5 working points (where the same offline working point requirements are applied to both jets). This leads to 6 sets of plots that will be shown in the following pages: Figures from 6.1 to 6.3 for trigger-only, Figures from 6.4 to 6.6 for the 90% tagged category, Figures from 6.7 to 6.9 for the 85% tagged category, Figures from 6.10 to 6.12 for the 77% tagged category,Figures from 6.13 to 6.15 for the 70% tagged category, and Figures from 6.16 to 6.18 for the 65% tagged category.

For each set, data/MC-simulation agreement plots are shown for lepton and jet kinematics, for some event-wide variables and also for the online and offline b-tagging discriminants in order to get a first idea of the expected data-to-MC scale factors. The plots shown in this chapter have been produced using only the 2022 dataset.

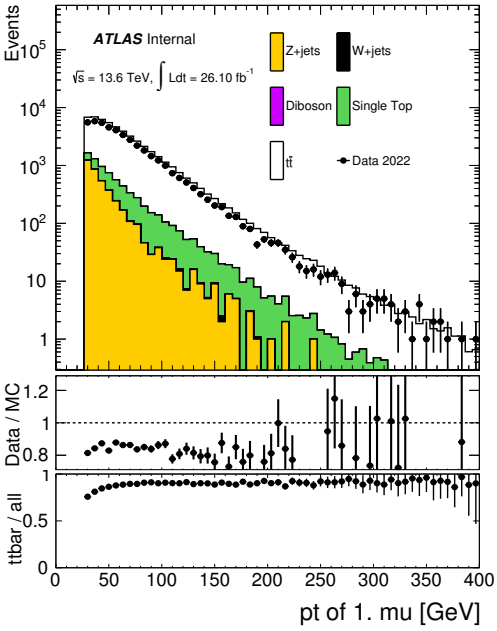
6.1 trigger-only



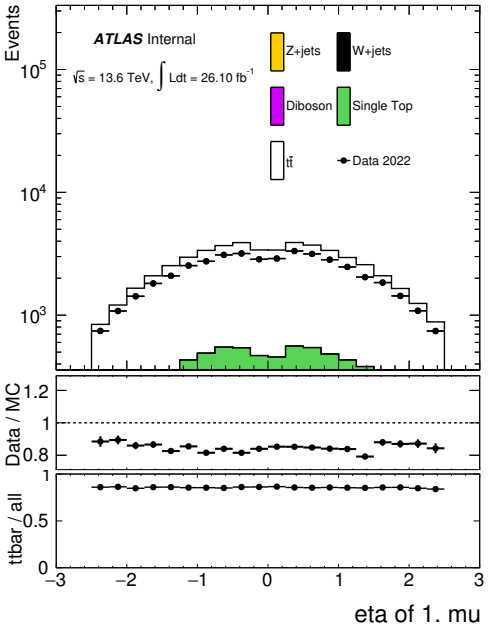
(a) Electron p_T Distribution



(b) Electron η distribution.

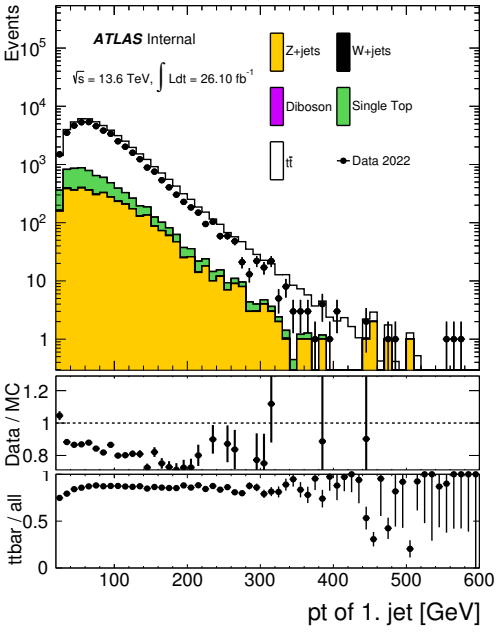


(c) Muon p_T Distribution

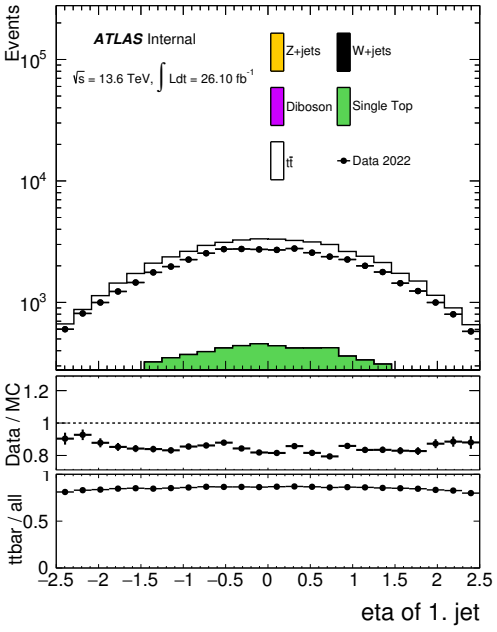


(d) Muon η distribution.

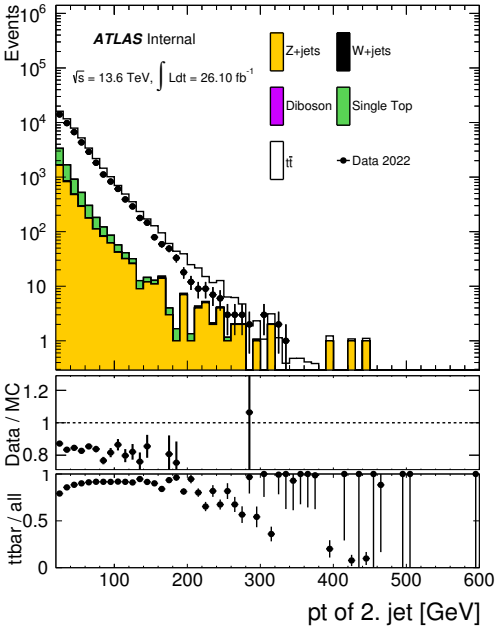
Figure 6.1: The distributions of transverse momentum p_T and pseudorapidity η of the two leptons in the selected events.



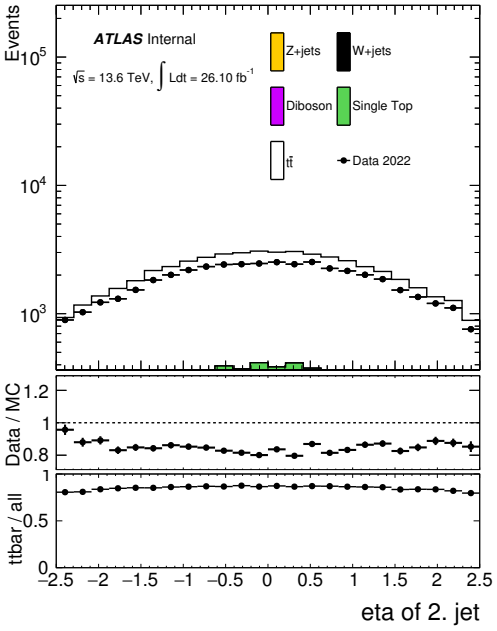
(a) Leading jet p_T distribution.



(b) Leading jet η distribution.

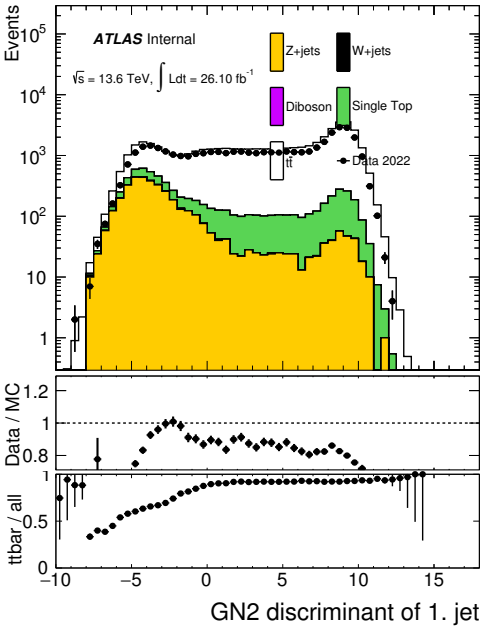


(c) Subleading jet p_T Distribution

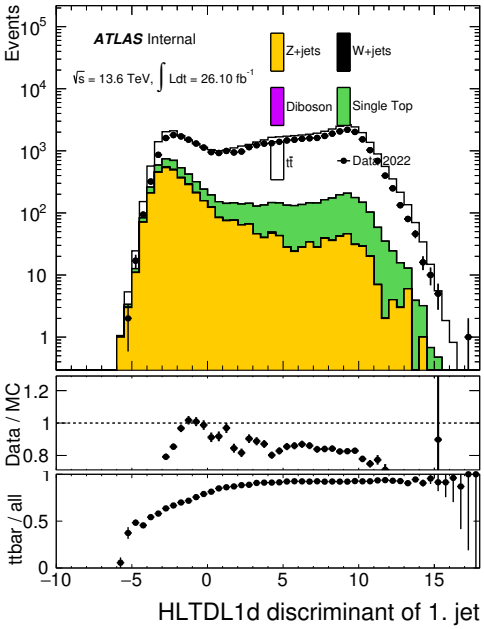


(d) Subleading jet η distribution.

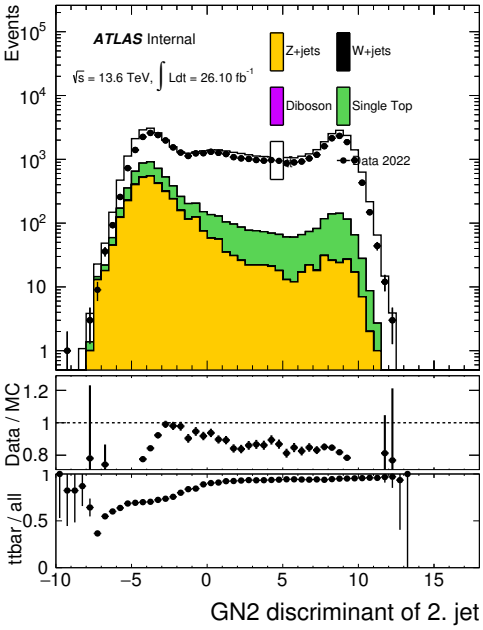
Figure 6.2: The distributions of transverse momentum p_T and pseudorapidity η of the two jets in the selected events.



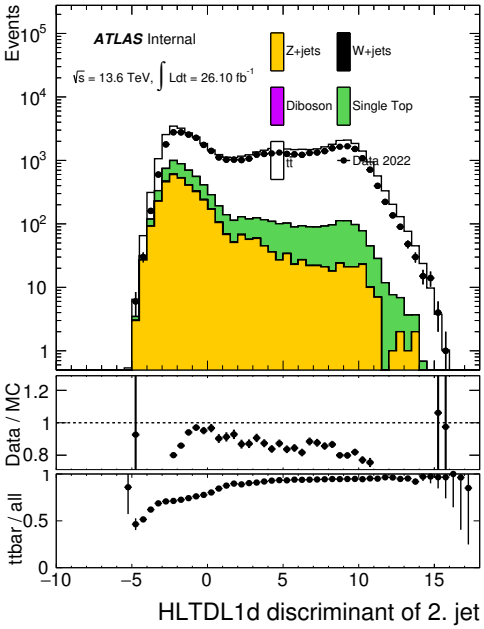
(a) Leading jet GN2 distribution.



(b) Leading jet HLTDL1d distribution.



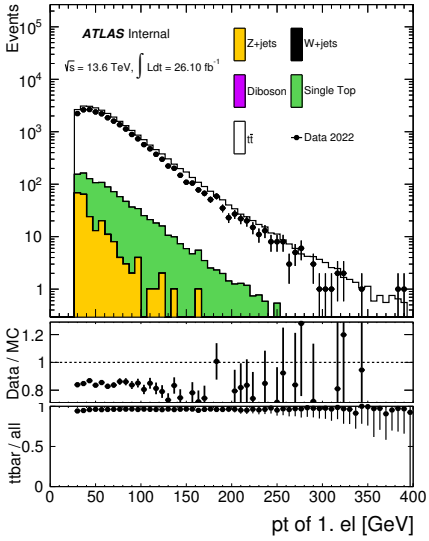
(c) Subleading jet GN2 Distribution



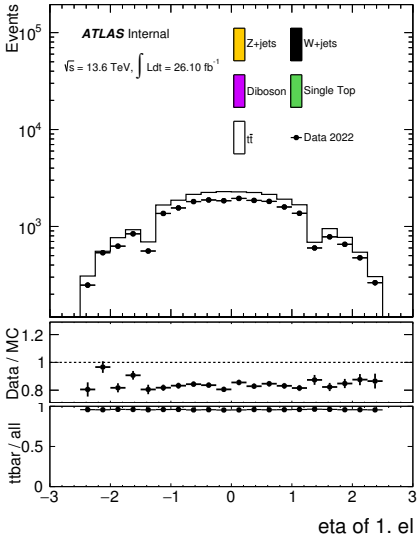
(d) Subleading jet HLTDL1d distribution.

Figure 6.3: GN2 and online DL1d b-tagging algorithms distributions of the two jets in the selected events.

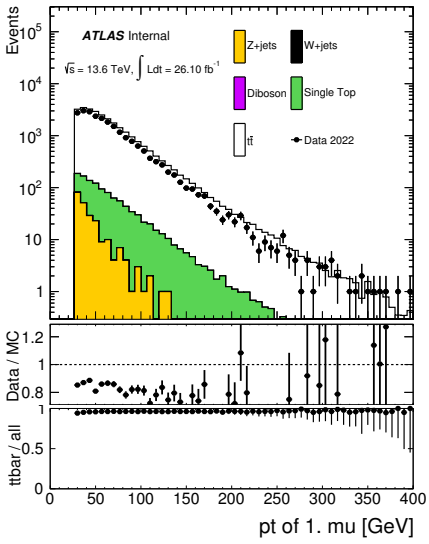
6.2 With offline b-tagging at 90 working point



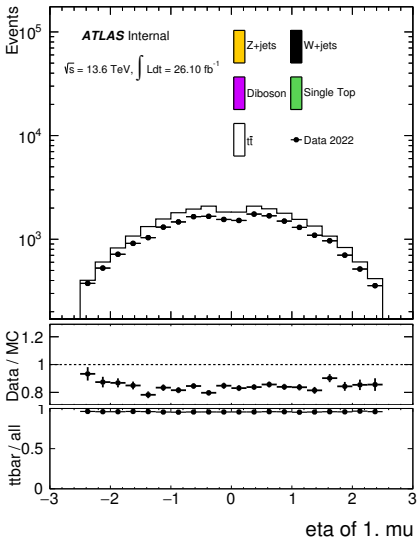
(a) Electron p_T Distribution



(b) Electron η distribution.

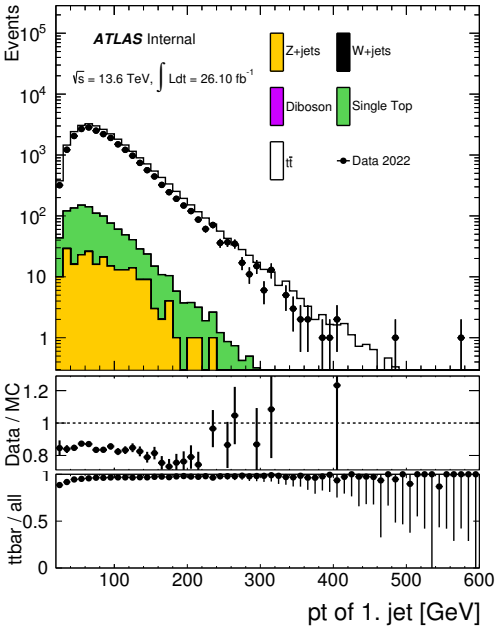


(c) Muon p_T Distribution

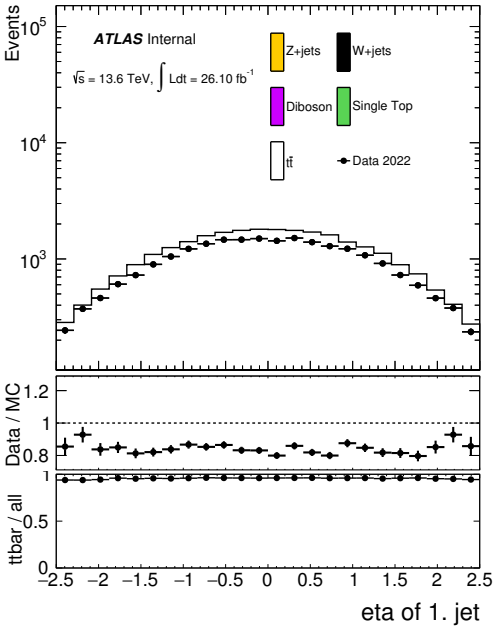


(d) Muon η distribution.

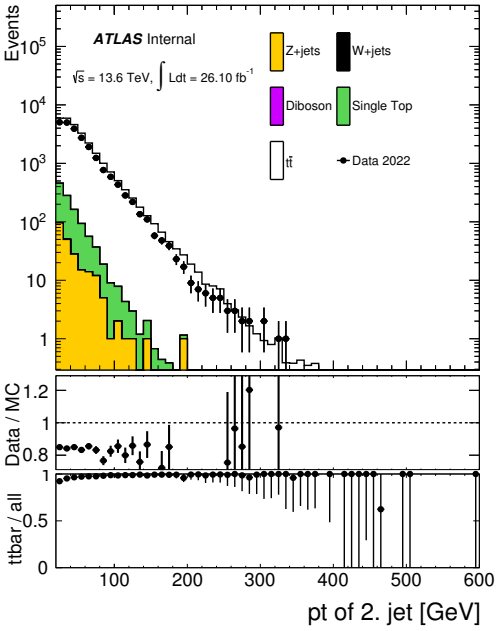
Figure 6.4: The distributions of transverse momentum p_T and pseudorapidity η of the two leptons in the selected events.



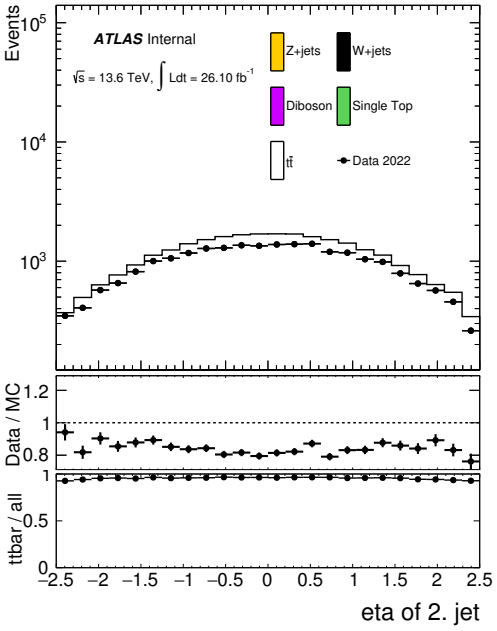
(a) Leading jet p_T distribution.



(b) Leading jet η distribution.

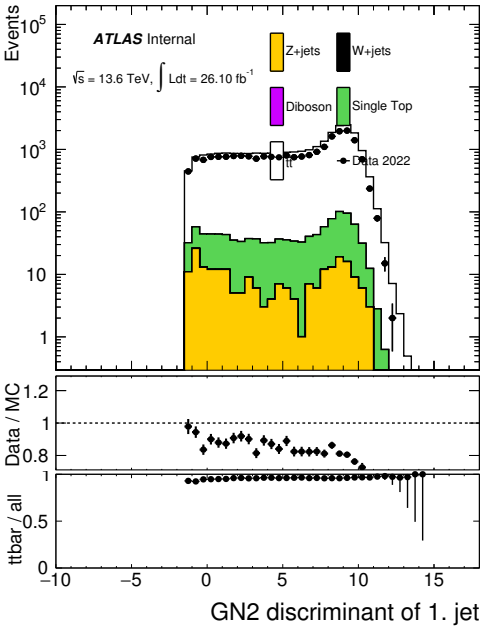


(c) Subleading jet p_T Distribution

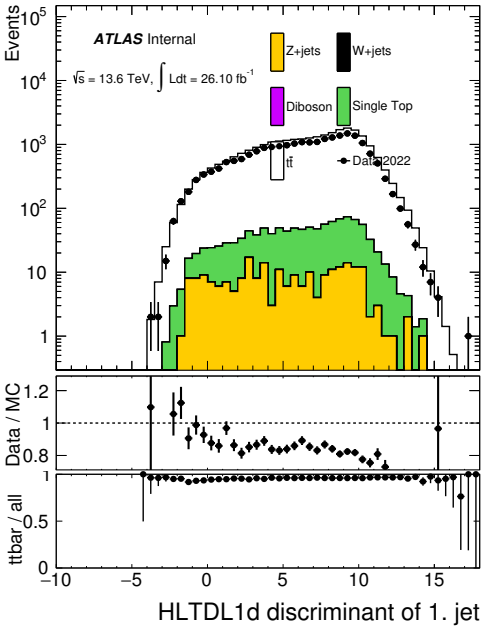


(d) Subleading jet η distribution.

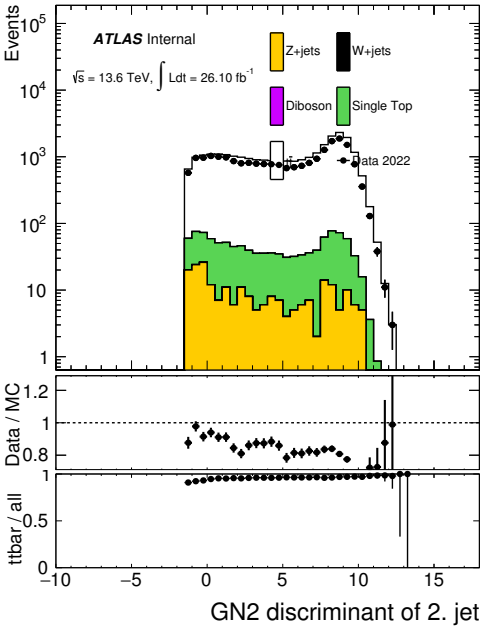
Figure 6.5: The distributions of transverse momentum p_T and pseudorapidity η of the two jets in the selected events.



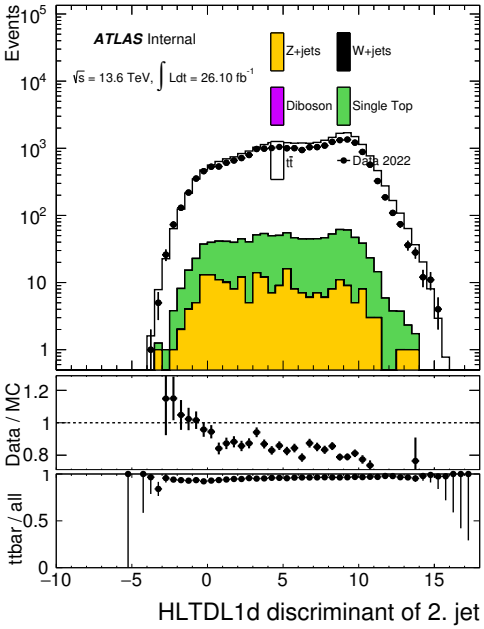
(a) Leading jet GN2 distribution.



(b) Leading jet HLTDL1d distribution.



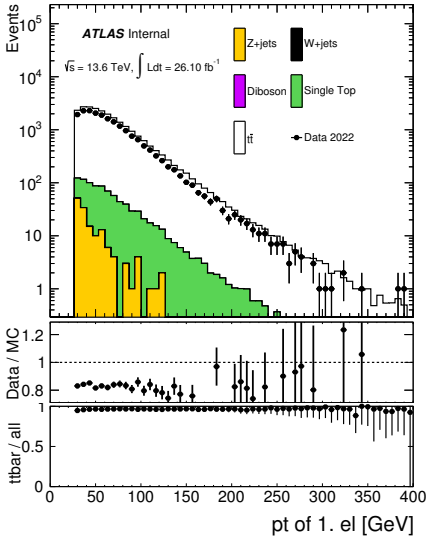
(c) Subleading jet GN2 Distribution



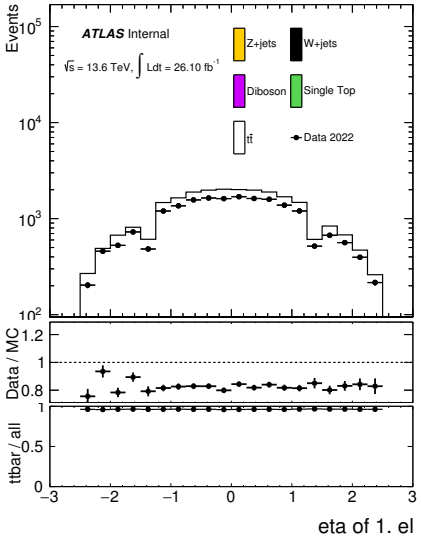
(d) Subleading jet HLTDL1d distribution.

Figure 6.6: GN2 and online DL1d b-tagging algorithms distributions of the two jets in the selected events.

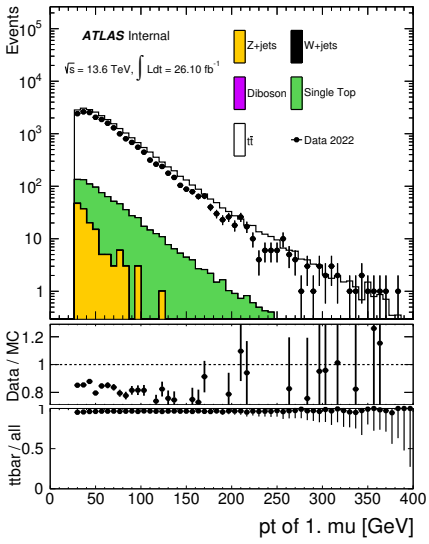
6.3 With offline b-tagging at 85 working point



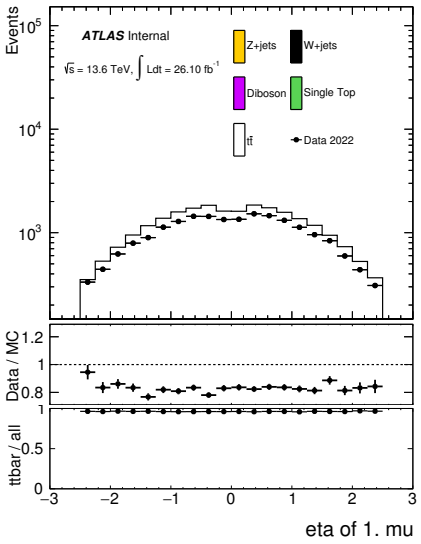
(a) Electron p_T Distribution



(b) Electron η distribution.

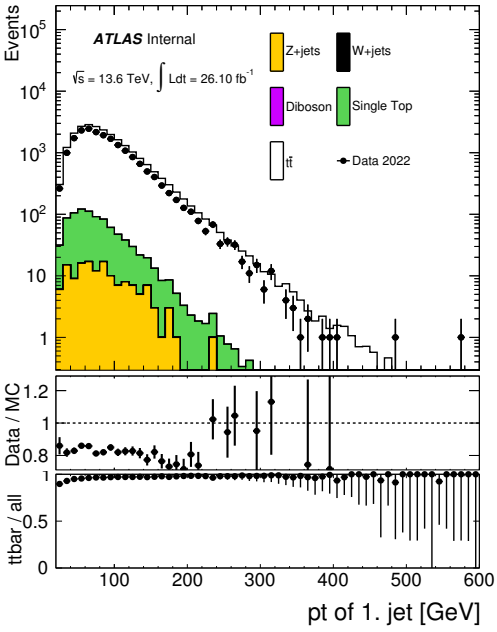


(c) Muon p_T Distribution

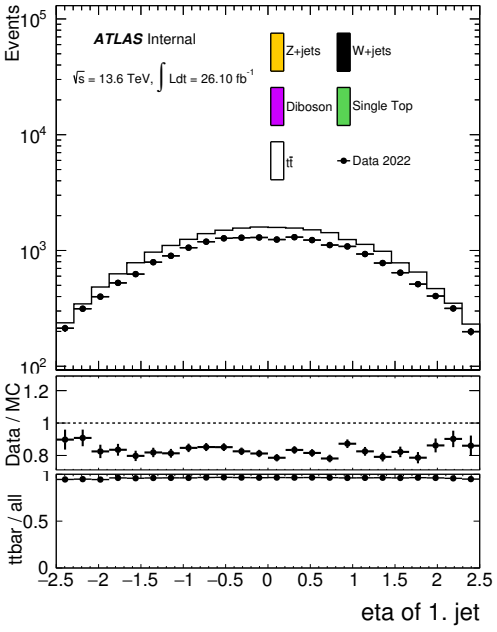


(d) Muon η distribution.

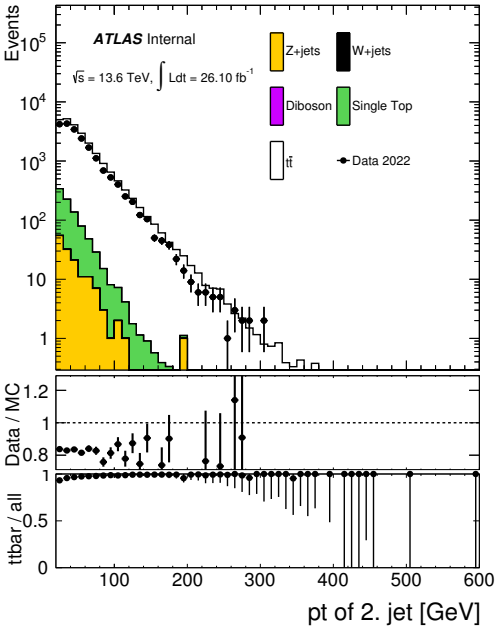
Figure 6.7: The distributions of transverse momentum p_T and pseudorapidity η of the two leptons in the selected events.



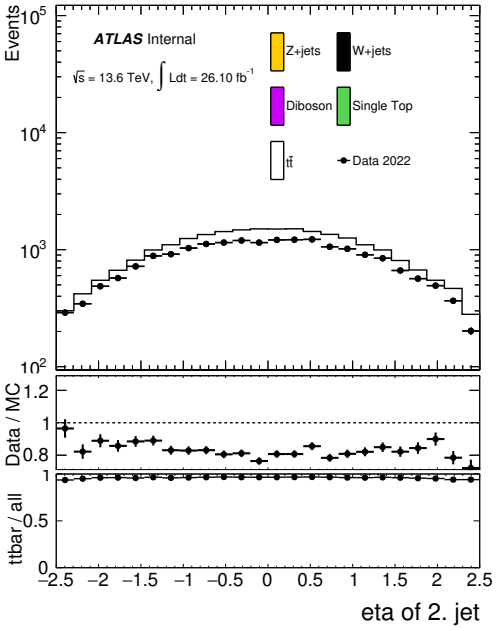
(a) Leading jet p_T distribution.



(b) Leading jet η distribution.

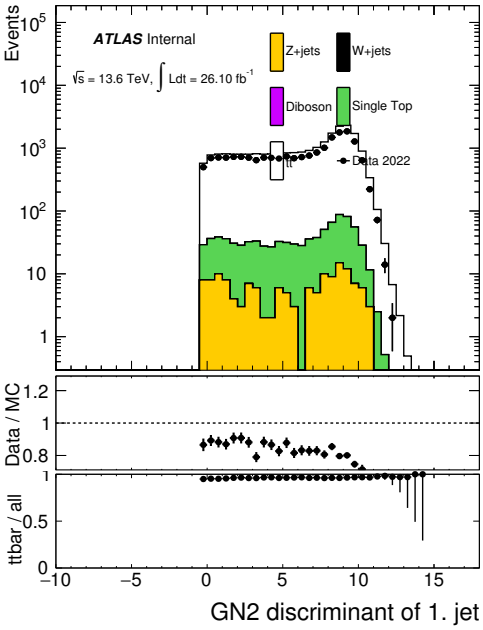


(c) Subleading jet p_T Distribution

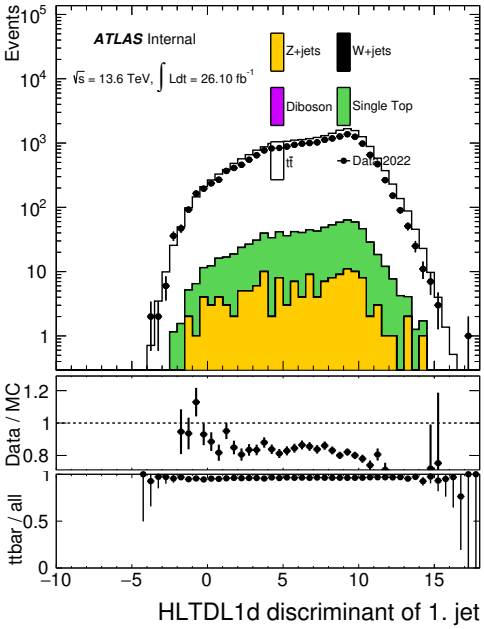


(d) Subleading jet η distribution.

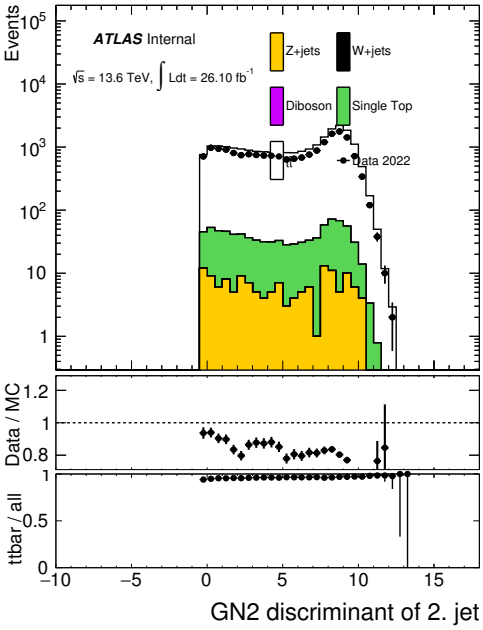
Figure 6.8: The distributions of transverse momentum p_T and pseudorapidity η of the two jets in the selected events.



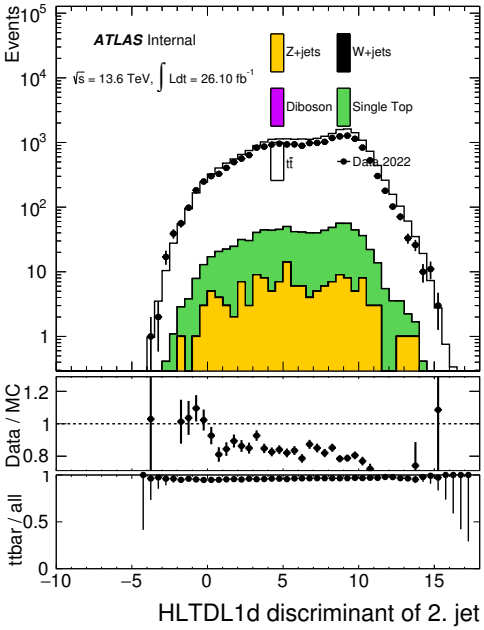
(a) Leading jet GN2 distribution.



(b) Leading jet HLTDL1d distribution.



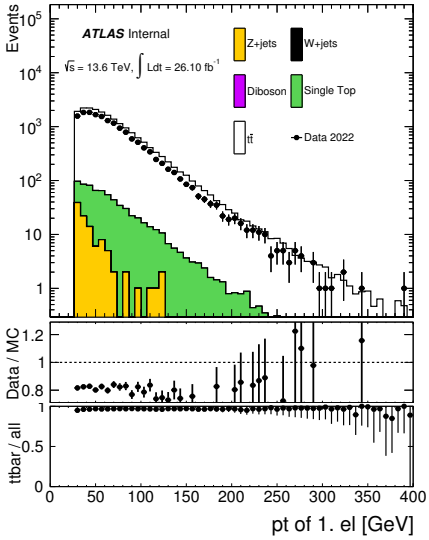
(c) Subleading jet GN2 Distribution



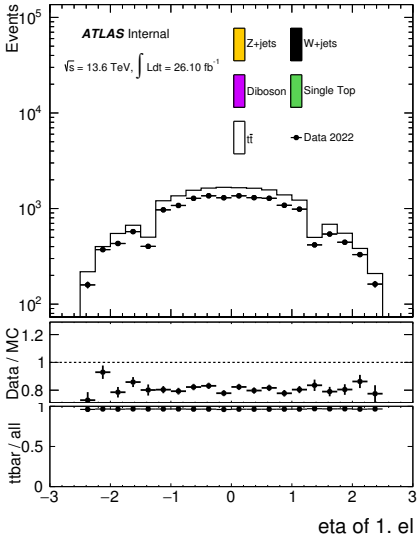
(d) Subleading jet HLTDL1d distribution.

Figure 6.9: GN2 and online DL1d b-tagging algorithms distributions of the two jets in the selected events.

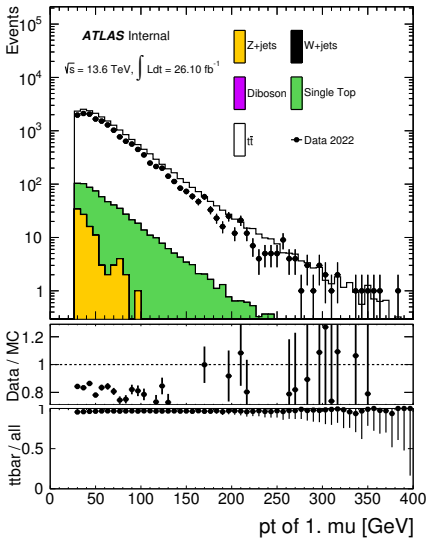
6.4 With offline b-tagging at 77 working point



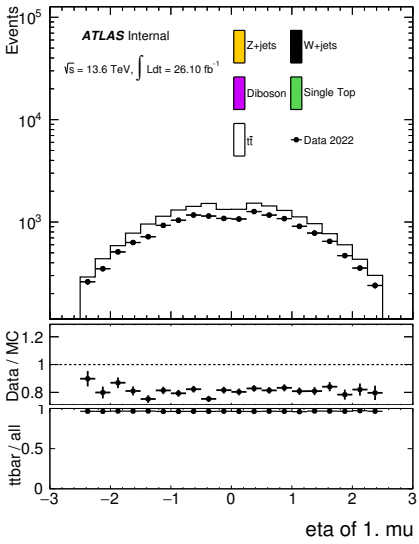
(a) Electron p_T Distribution



(b) Electron η distribution.

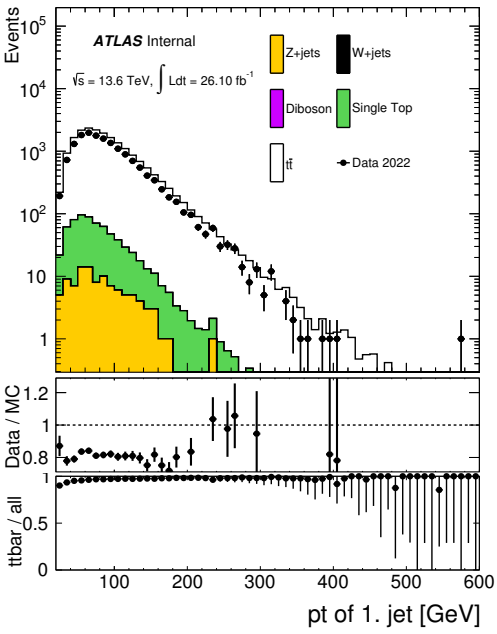


(c) Muon p_T Distribution

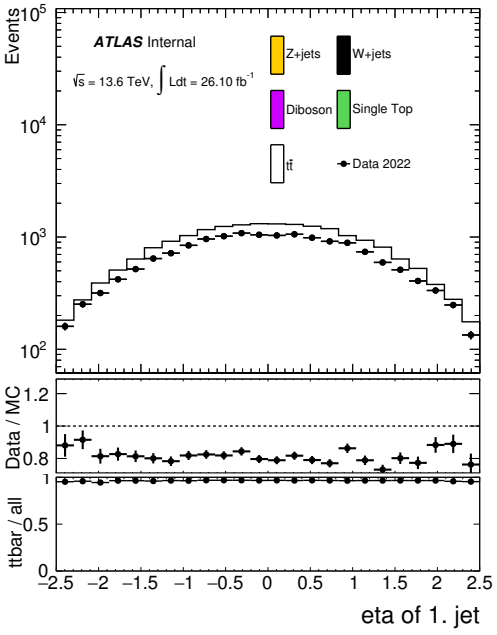


(d) Muon η distribution.

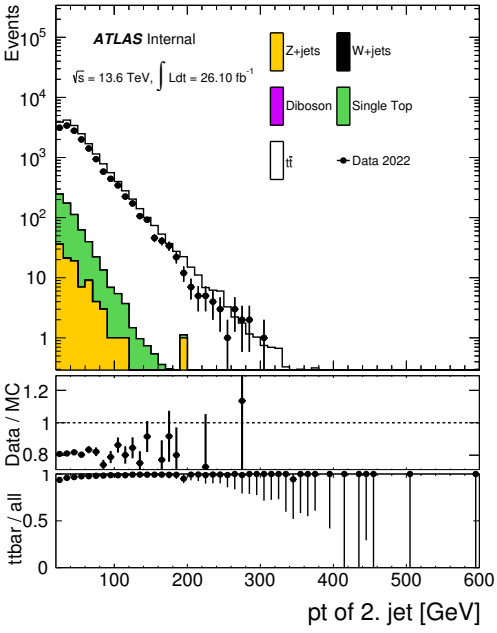
Figure 6.10: The distributions of transverse momentum p_T and pseudorapidity η of the two leptons in the selected events.



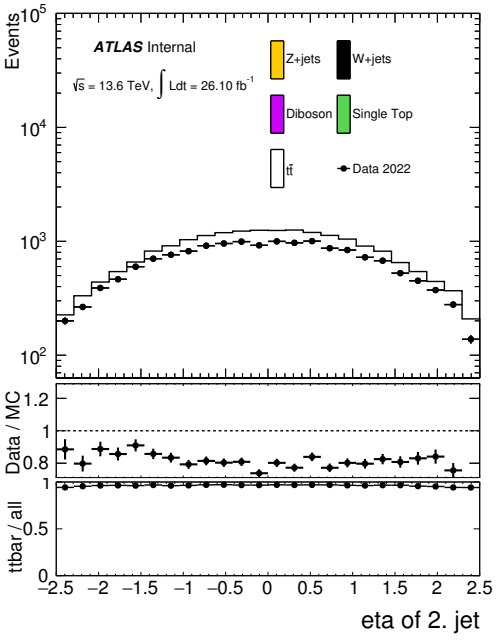
(a) Leading jet p_T distribution.



(b) Leading jet η distribution.

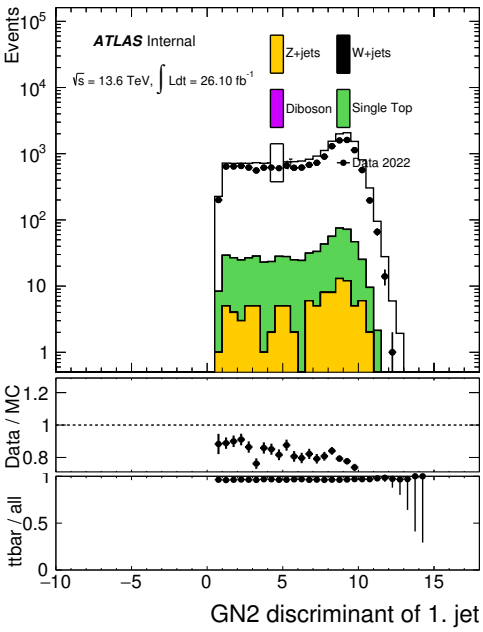


(c) Subleading jet p_T Distribution

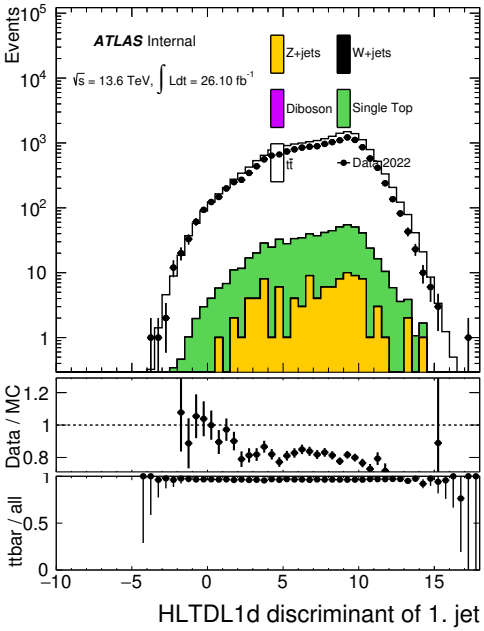


(d) Subleading jet η distribution.

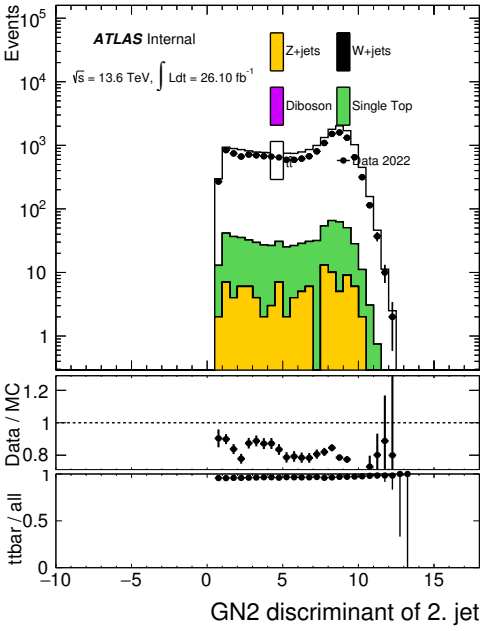
Figure 6.11: The distributions of transverse momentum p_T and pseudorapidity η of the two jets in the selected events.



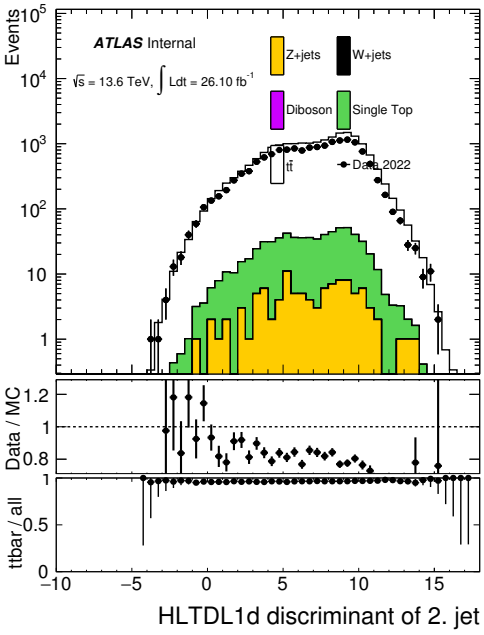
(a) Leading jet GN2 distribution.



(b) Leading jet HLTDL1d distribution.



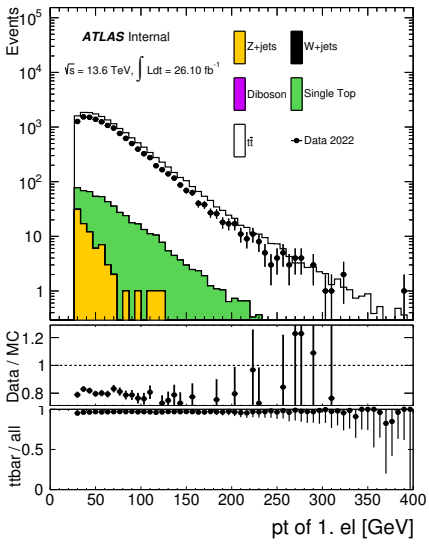
(c) Subleading jet GN2 Distribution



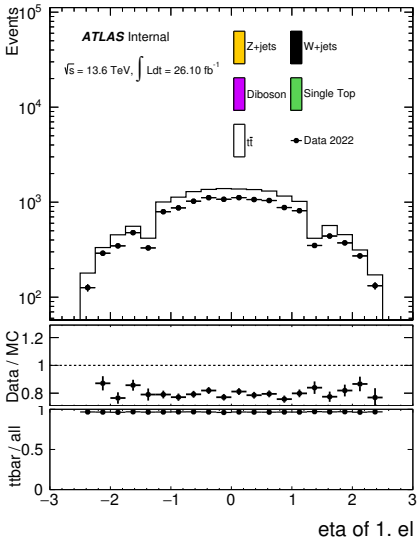
(d) Subleading jet HLTDL1d distribution.

Figure 6.12: GN2 and online DL1d b-tagging algorithms distributions of the two jets in the selected events.

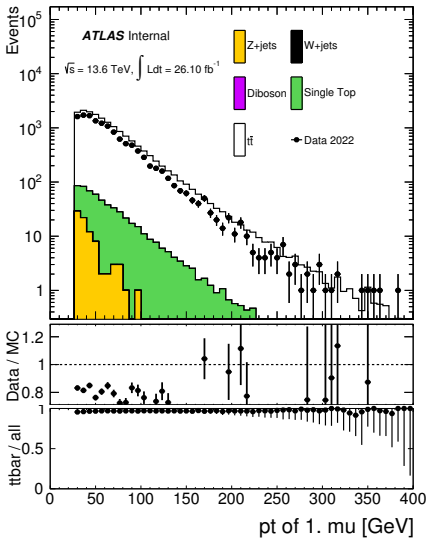
6.5 With offline b-tagging at 70 working point



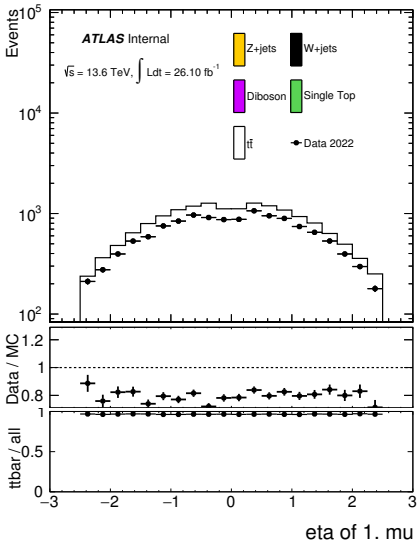
(a) Electron p_T Distribution



(b) Electron η distribution.

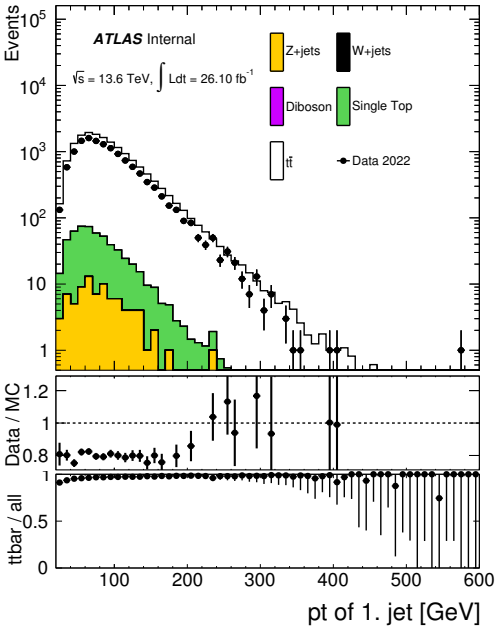


(c) Muon p_T Distribution

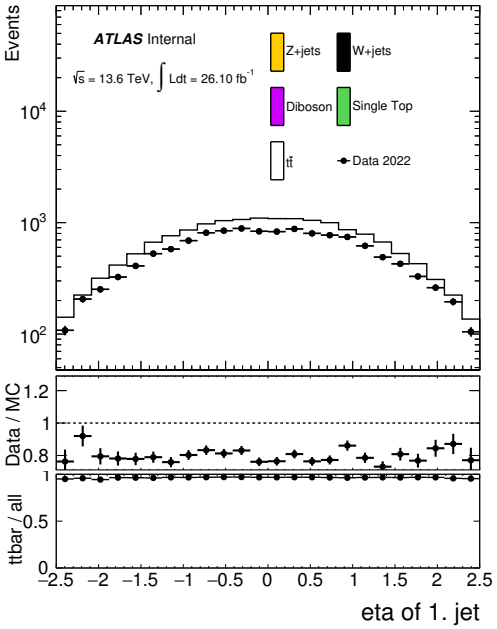


(d) Muon η distribution.

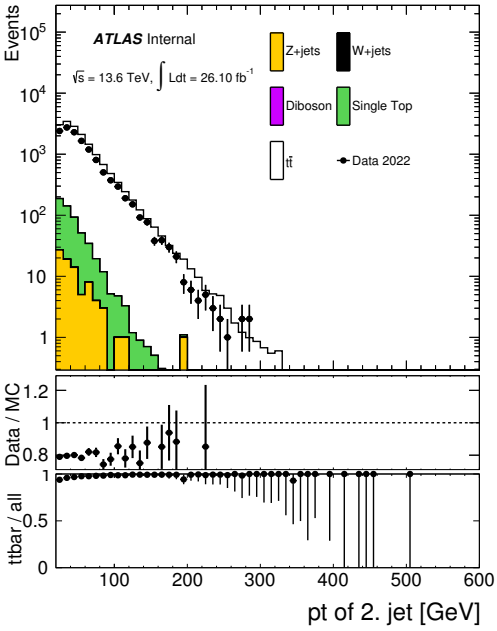
Figure 6.13: The distributions of transverse momentum p_T and pseudorapidity η of the two leptons in the selected events.



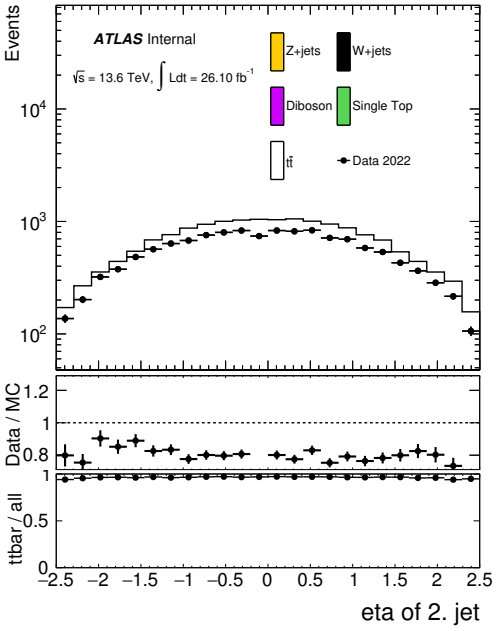
(a) Leading jet p_T distribution.



(b) Leading jet η distribution.

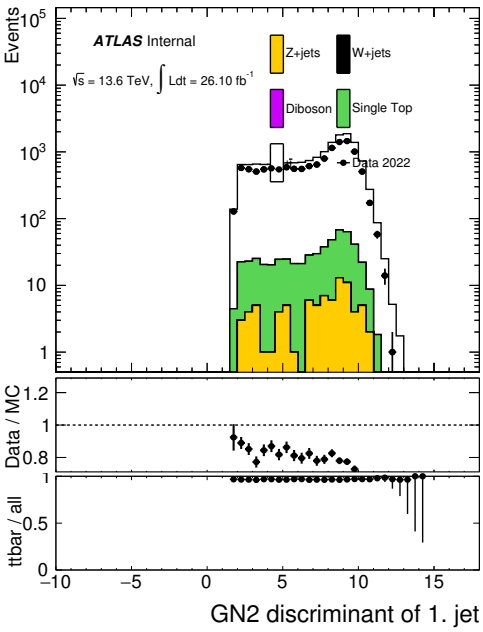


(c) Subleading jet p_T Distribution

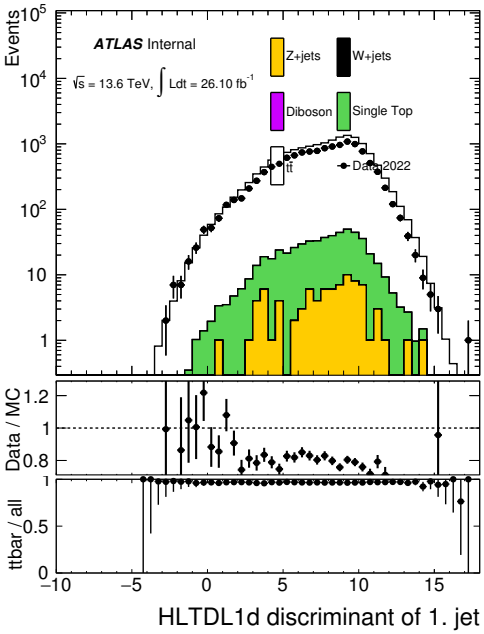


(d) Subleading jet η distribution.

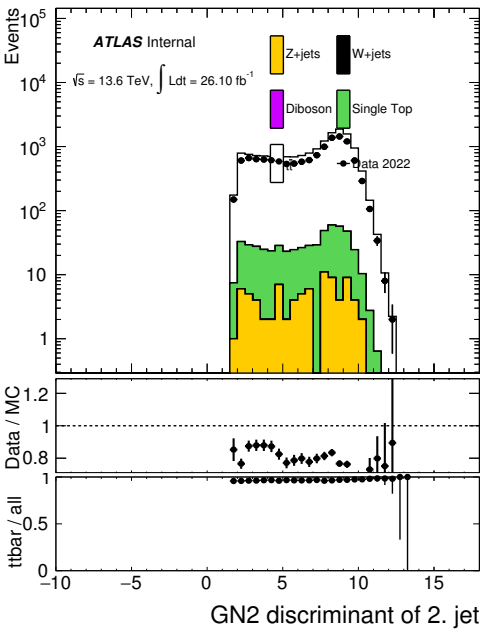
Figure 6.14: The distributions of transverse momentum p_T and pseudorapidity η of the two jets in the selected events.



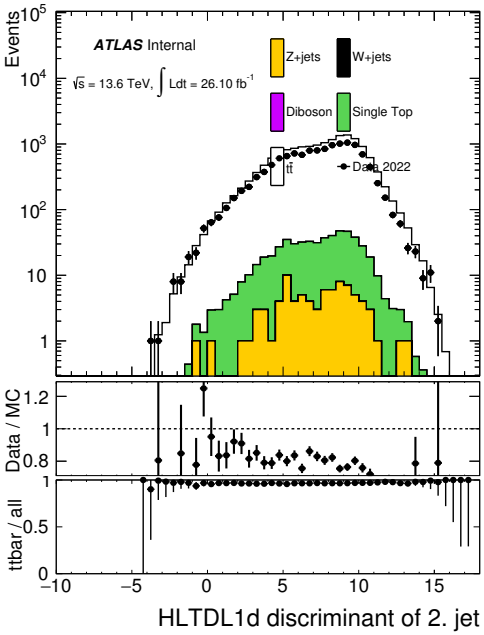
(a) Leading jet GN2 distribution.



(b) Leading jet HLTDL1d distribution.



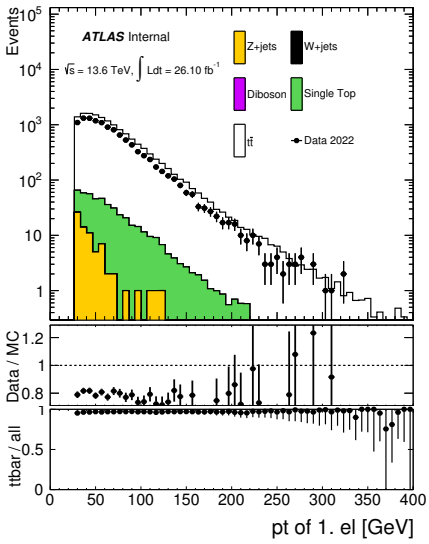
(c) Subleading jet GN2 Distribution



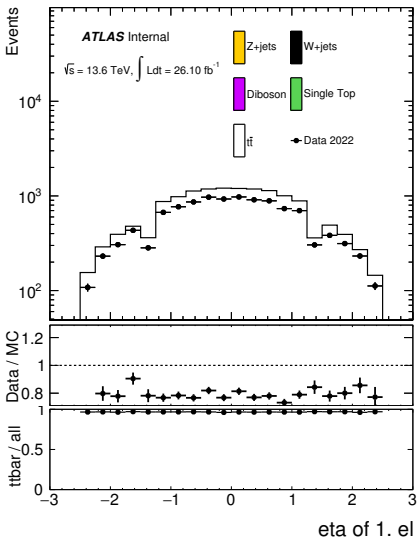
(d) Subleading jet HLTDL1d distribution.

Figure 6.15: GN2 and online DL1d b-tagging algorithms distributions of the two jets in the selected events.

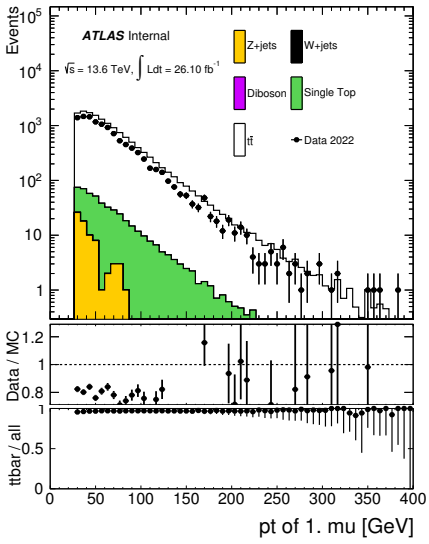
6.6 With offline b-tagging at 65 working point



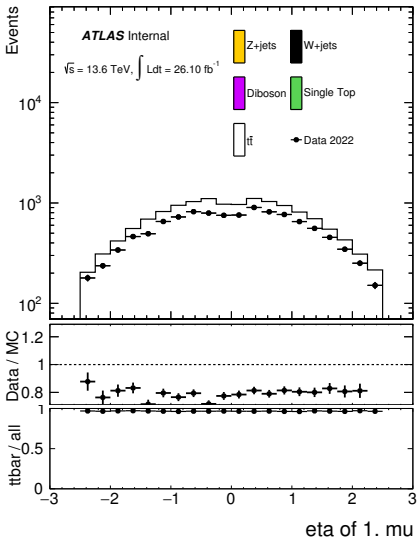
(a) Electron p_T Distribution



(b) Electron η distribution.

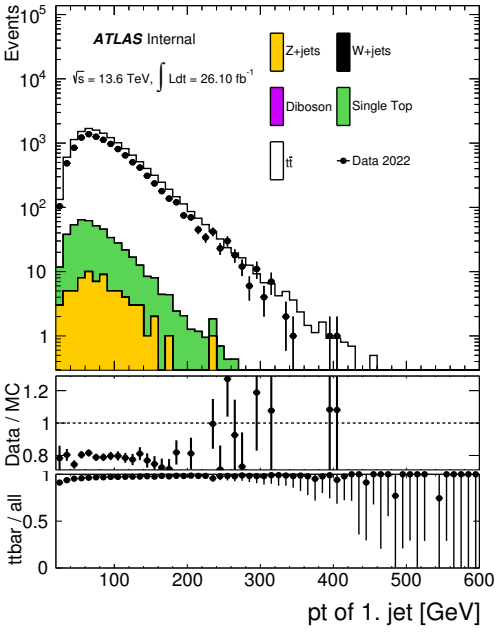


(c) Muon p_T Distribution

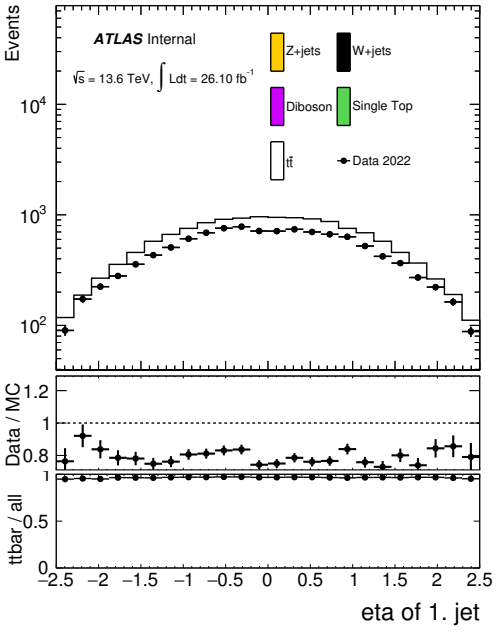


(d) Muon η distribution.

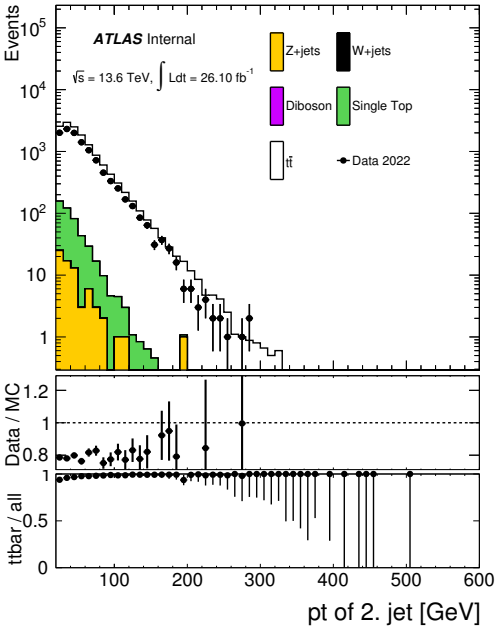
Figure 6.16: The distributions of transverse momentum p_T and pseudorapidity η of the two leptons in the selected events.



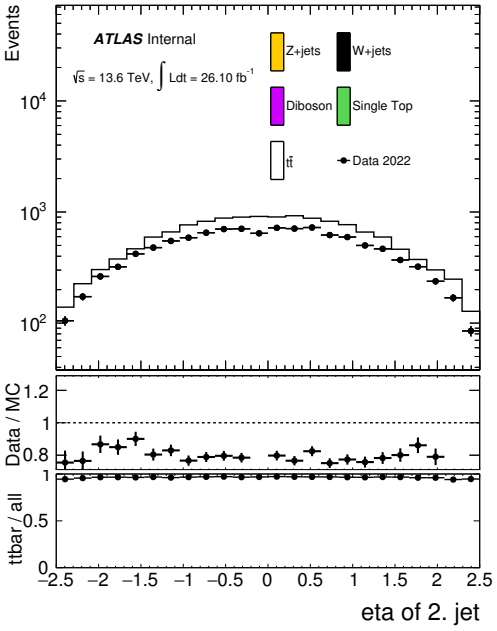
(a) Leading jet p_T distribution.



(b) Leading jet η distribution.

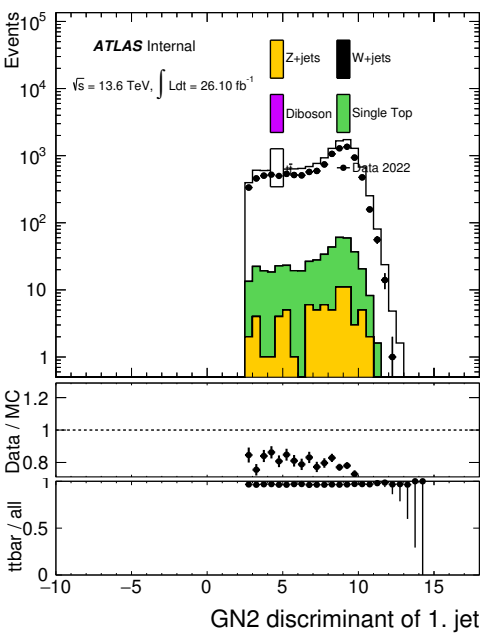


(c) Subleading jet p_T Distribution

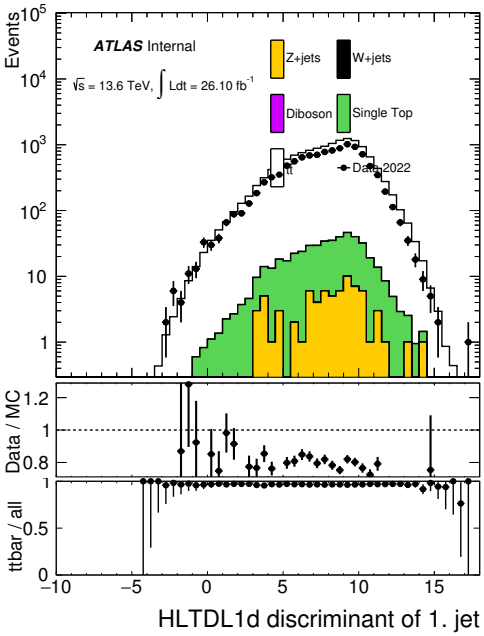


(d) Subleading jet η distribution.

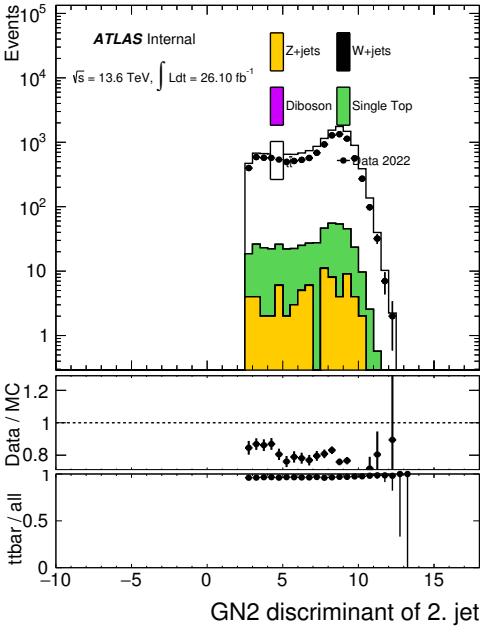
Figure 6.17: The distributions of transverse momentum p_T and pseudorapidity η of the two jets in the selected events.



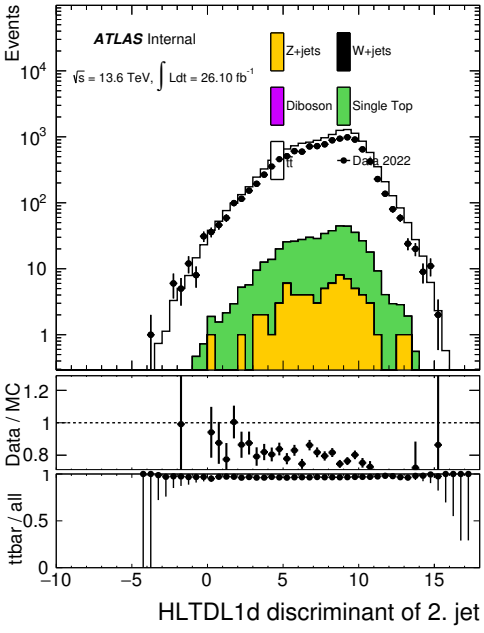
(a) Leading jet GN2 distribution.



(b) Leading jet HLTDL1d distribution.



(c) Subleading jet GN2 Distribution



(d) Subleading jet HLTDL1d distribution.

Figure 6.18: GN2 and online DL1d b-tagging algorithms distributions of the two jets in the selected events.

7 Uncertainty

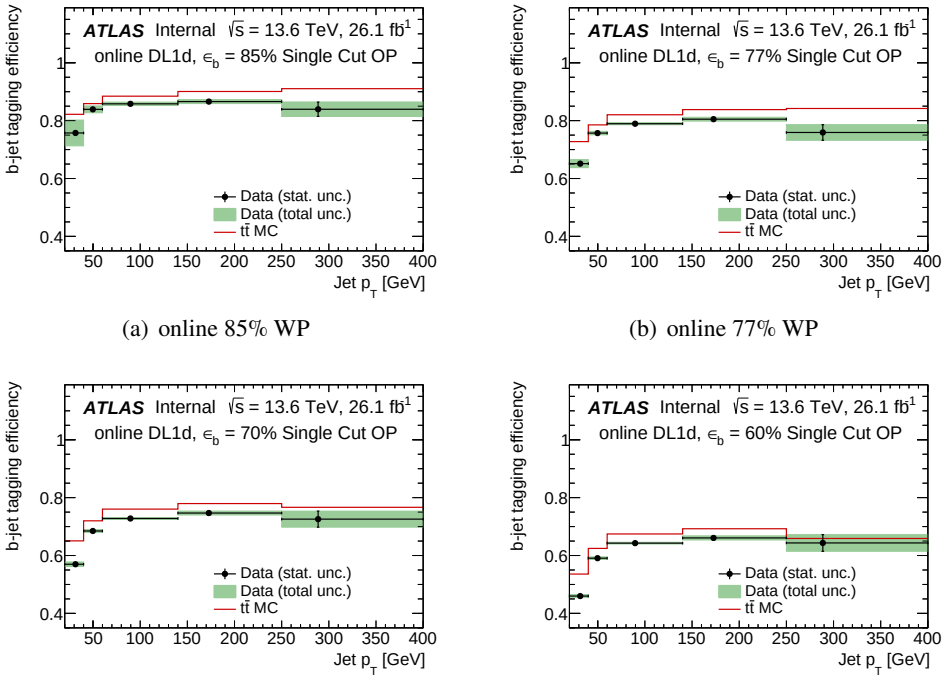
The data statistical uncertainties are obtained as in the offline measurement [**3**] with a principal component analysis method. The error matrix is diagonalized and the data statistical uncertainty is split into 20 uncorrelated components (5 p_T bins, $5-1=4$ independent b-tagging discriminant bins due to probability conservation, $5 \times 4 = 20$). No components are pruned and the total data statistical uncertainty for a given parameter is obtained by summing in quadrature the 24 components.

No differences with respect to the offline measurement [**1**] are introduced in the treatment of the systematic uncertainties coming from physics process modeling and detector modeling.

8 Results

The b-tagging efficiency measurements in data taken during 2022 for the trigger-only and conditional are presented in Figs.8.1, Figs.8.2-Figs.8.6, respectively. The efficiencies are shown for all online working points, in the case where no offline b-tagging requirements are applied for trigger-only and conditional. Data-to-MC scale factors, deriving from the b-tagging efficiency measurements in 2022 are also presented separately for trigger-only and conditional.

8.1 trigger-only



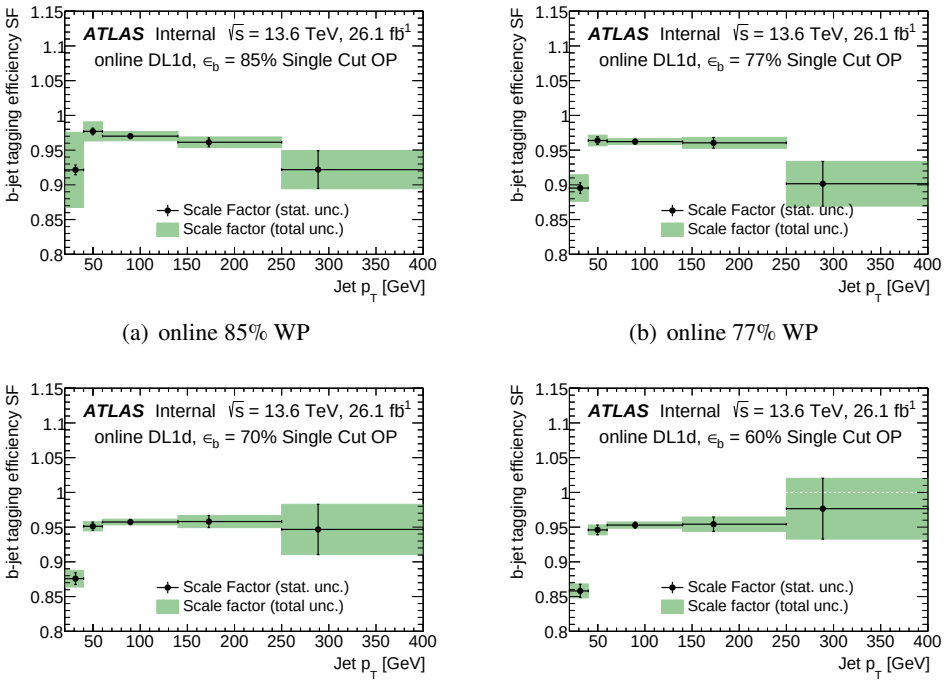
(a) online 85% WP

(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.1: Measured *b*-tagging efficiencies for the online DL1d tagger in 2022 data for all online working points.

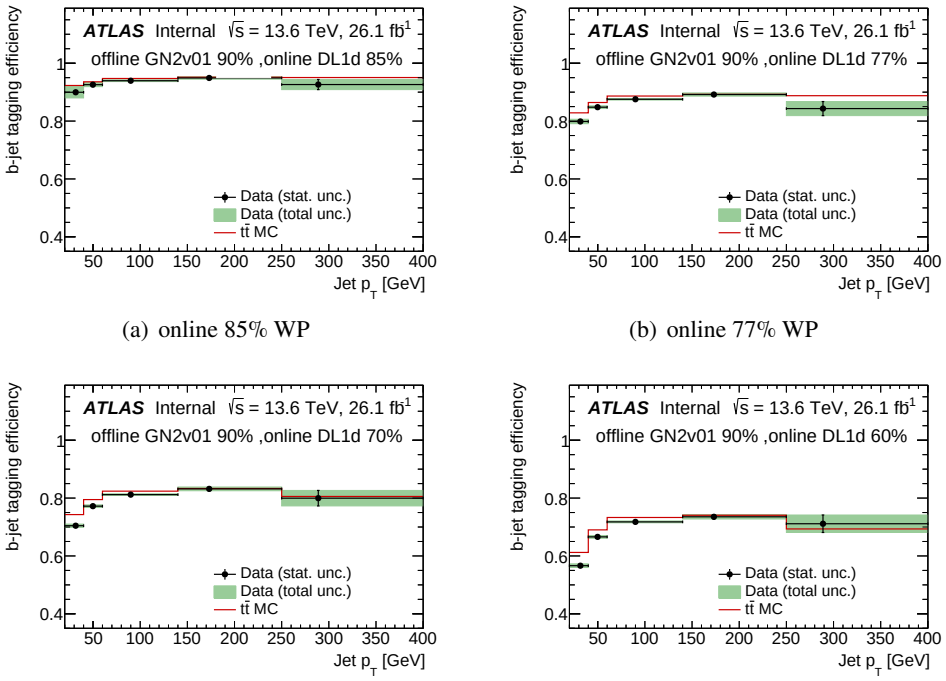


(a) online 85% WP (b) online 77% WP

(c) online 70% WP (d) online 60% WP

Figure 8.2: Measured scale factor for the online DL1d tagger in 2022 data for all online working points.

8.2 With offline b-tagging at 90 working point



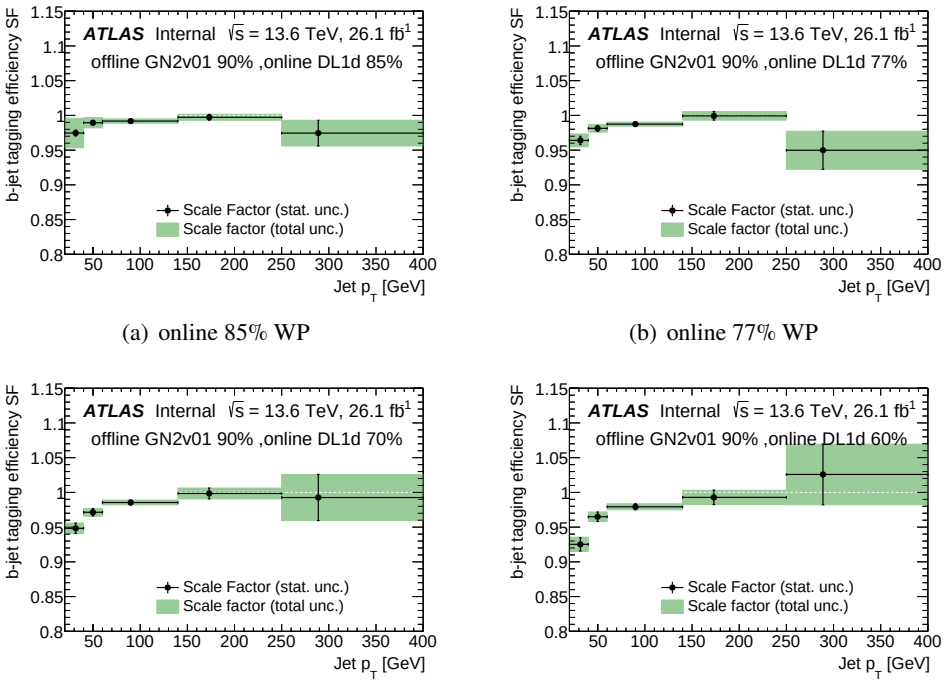
(a) online 85% WP

(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.3: Measured *b*-tagging efficiencies for the online DL1d tagger in 2022 data for all online working points.



(a) online 85% WP

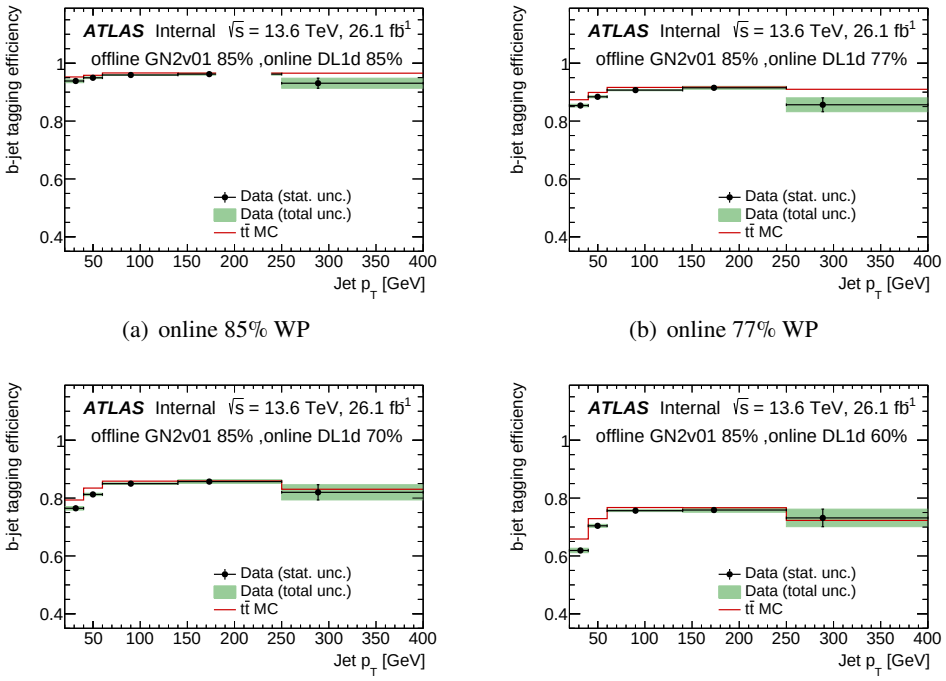
(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.4: Measured scale factor for the online DL1d tagger in 2022 data for all online working points.

8.3 With offline b-tagging at 85 working point



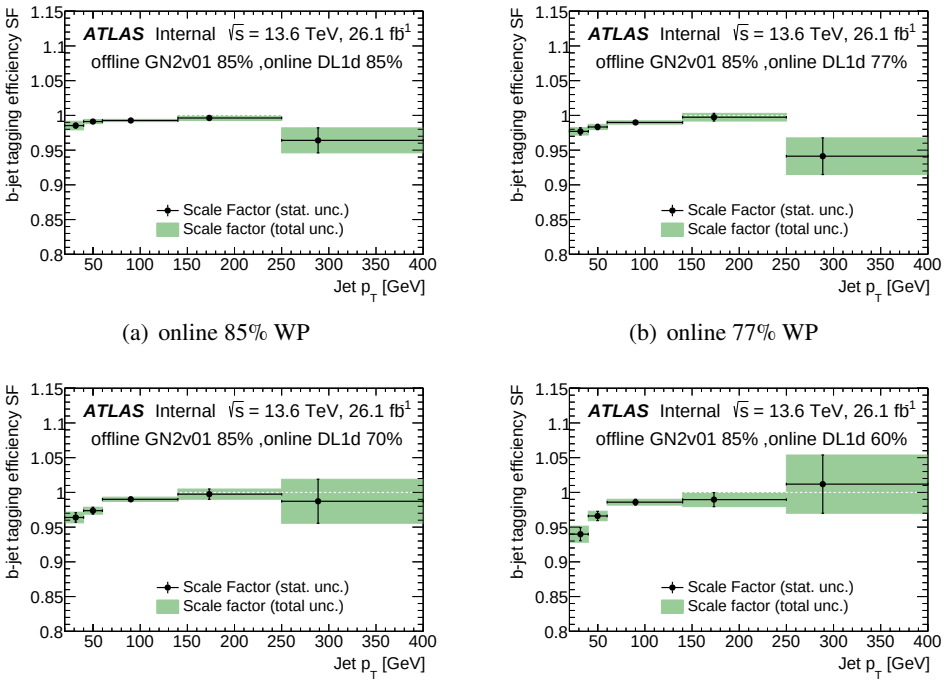
(a) online 85% WP

(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.5: Measured *b*-tagging efficiencies for the online DL1d tagger in 2022 data for all online working points.



(a) online 85% WP

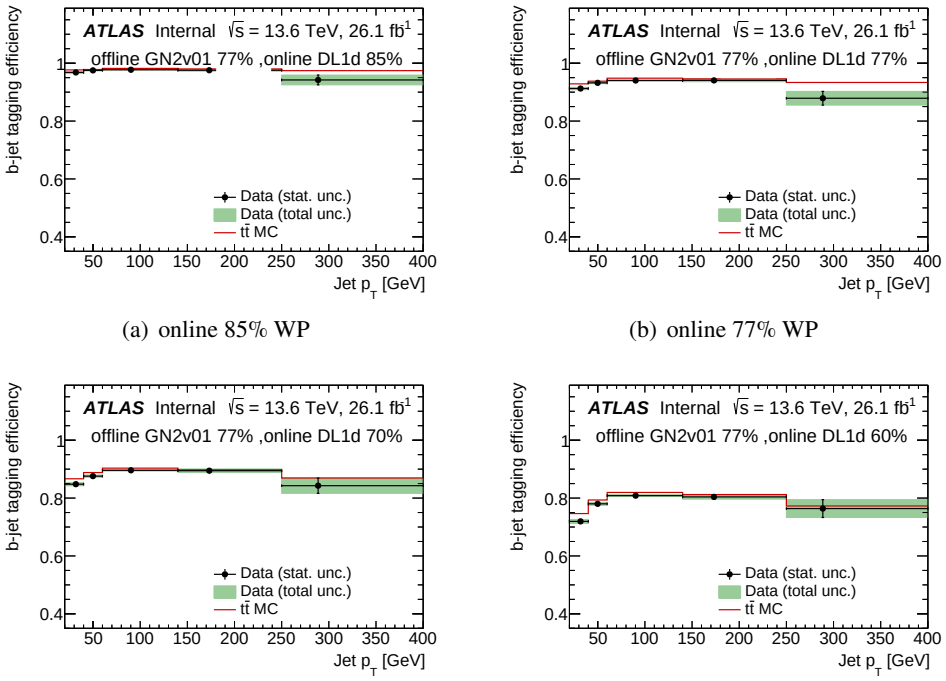
(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.6: Measured scale factor for the online DL1d tagger in 2022 data for all online working points.

8.4 With offline b-tagging at 77 working point



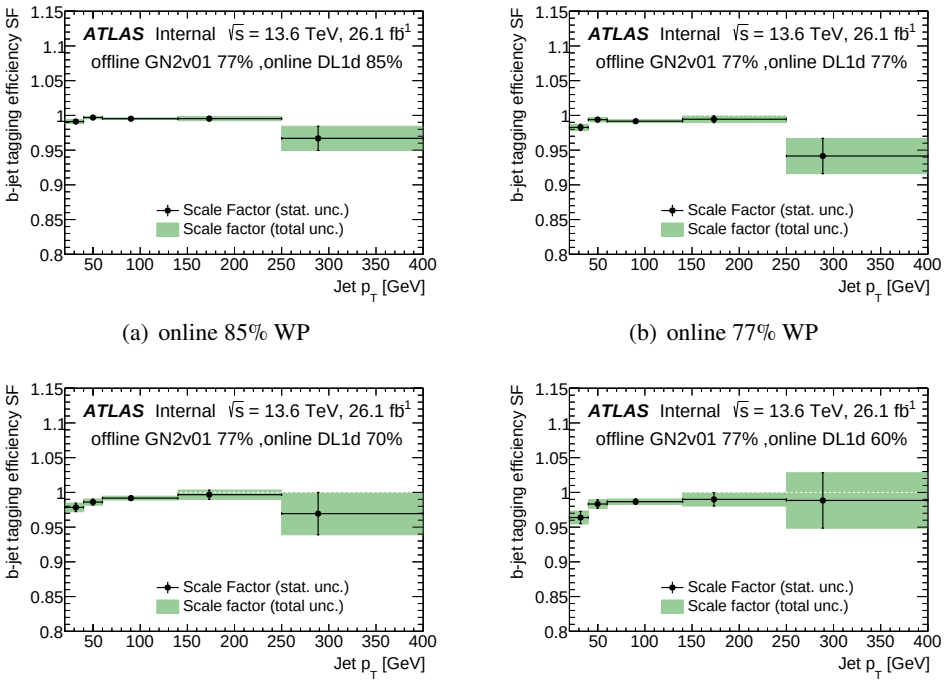
(a) online 85% WP

(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.7: Measured *b*-tagging efficiencies for the online DL1d tagger in 2022 data for all online working points.



(a) online 85% WP

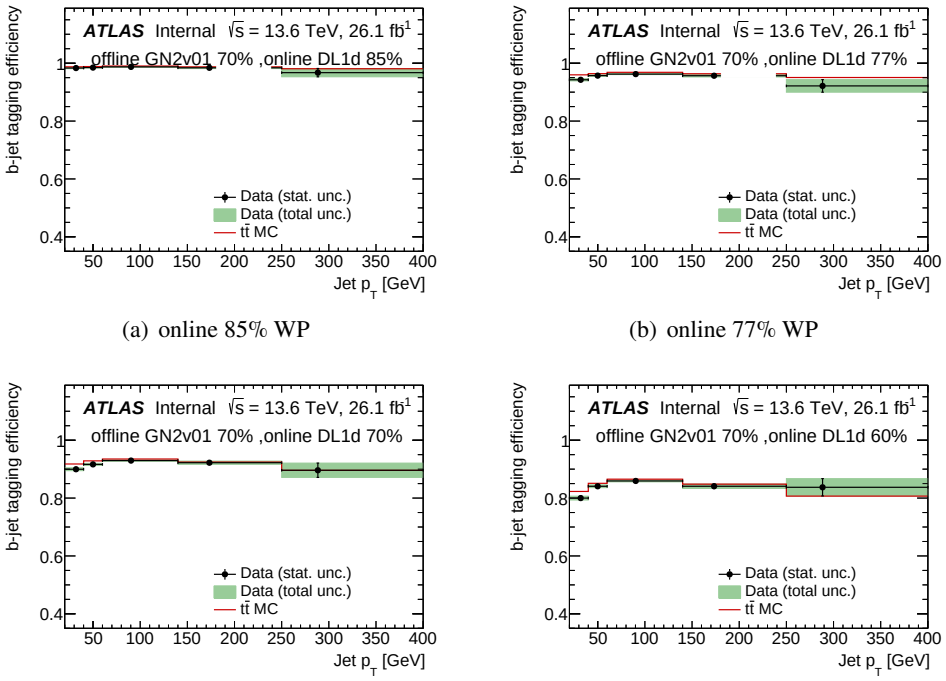
(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.8: Measured scale factor for the online DL1d tagger in 2022 data for all online working points.

8.5 With offline b-tagging at 70 working point



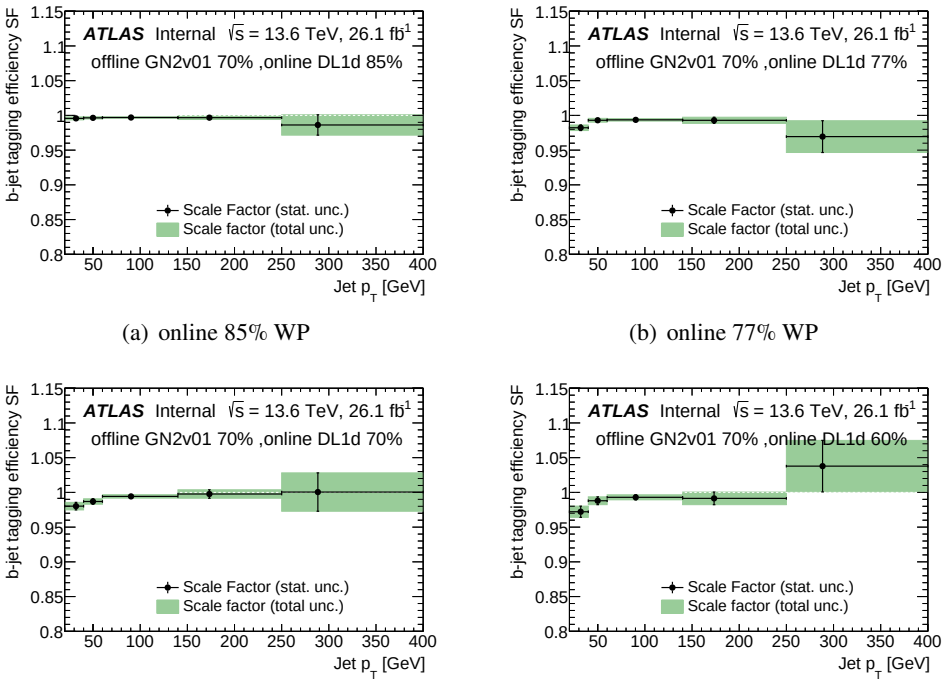
(a) online 85% WP

(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.9: Measured *b*-tagging efficiencies for the online DL1d tagger in 2022 data for all online working points.



(a) online 85% WP

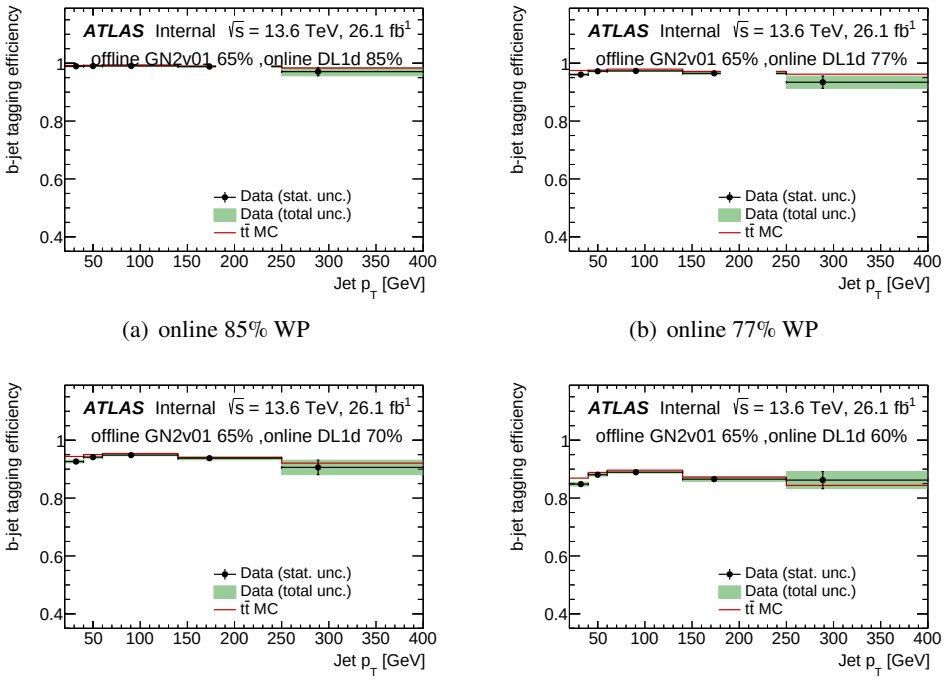
(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.10: Measured scale factor for the online DL1d tagger in 2022 data for all online working points.

8.6 With offline b-tagging at 65 working point



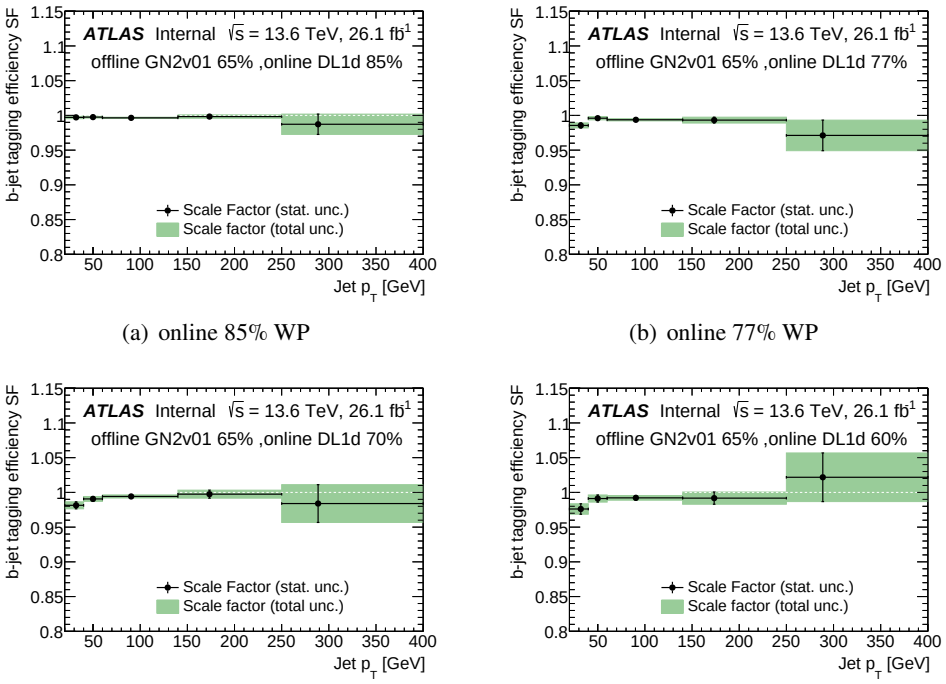
(a) online 85% WP

(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.11: Measured *b*-tagging efficiencies for the online DL1d tagger in 2022 data for all online working points.



(a) online 85% WP

(b) online 77% WP

(c) online 70% WP

(d) online 60% WP

Figure 8.12: Measured scale factor for the online DL1d tagger in 2022 data for all online working points.

9 Conclusion

A likelihood-based method to measure the online b-jet identification efficiency in a sample strongly enriched in $t\bar{t}$ di-leptonic events is presented and applied separately to the data samples recorded by the ATLAS detector in the years 2022. Results from the trigger-only b-tagging efficiency and from a conditional b-tagging efficiency are presented. The online b-tagging weight is retrieved and associated with offline jets through a geometrical matching between offline and online jet objects. This allows both tagger weights to be available on the same object, making conditional calibration possible.

The results demonstrate high precision with relative uncertainties consistently below 5% for the online b-tagging efficiency and below 3% for the conditional b-tagging efficiency. Notably, these uncertainties are predominantly driven by statistical factors. In the offline measurements, the primary source of statistical uncertainty arises from the m_{jl} control regions (CRs), which are employed to fit the flavor fractions. However, these regions are not utilized in the online measurements. For the trigger-only measurements, only trigger and geometric matching for offline jets are performed. In the conditional measurements, where events have already passed offline b-tagging, there is a marked increase in signal region purity. Consequently, the b-tagging efficiency in these conditional measurements is high, and the scale factor approaches 1, particularly when meeting the 65% working points of GN2.

Bibliography

[1] J. G. James Ferrando et al., "Measurement of the b-jet identification efficiency with $t\bar{t}$ events using an improved likelihood method," ATLAS Collaboration, ANA-FTAG-2018-01-INT2, 2018. Citations: pp. 6, 11, 13, 14, 17, 43–45, 130.

[2] ATLAS Collaboration, "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$ = 13 TeV," arXiv:1907.05120 [hep-ex], 2019. Citations: pp. 7, 9.

[3] R. Gupta and J. Alison, "b-tagging efficiency for b-jet triggers," CERN, ATL-COM-DAQ-2017-062, 2017. URL: https://cds.cern.ch/record/2271945. Citation: p. 8.

[4] M. Aaboud et al., "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s}$ = 13 TeV," arXiv:1805.01845 [hep-ex], 2018. Citations: pp. 9, 43.

[5] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k_t jet clustering algorithm," JHEP **04**, 063 (2008). arXiv:0802.1189 [hep-ph]. Citation: p. 11.

[6] ATLAS Collaboration, "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1," Eur. Phys. J. C **77**, 490 (2017). arXiv:1603.02934 [hep-ex]. Citation: p. 11.

[7] ATLAS Collaboration, "Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector," ATLAS-CONF-2015-029, 2015. URL: https://cds.cern.ch/record/2037702. Citation: p. 11.

[8] ATLAS Collaboration, "Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at $\sqrt{s}$ = 13 TeV," ATLAS-CONF-2018-006, 2018. URL: https://cds.cern.ch/record/2314418. Citation: p. 13.

List of contributions

Name	Role/Contribution
Yuhui Miao	Development of code framework. Calibrating b-tagging efficiency and SF.
Lei Zhang	Local supervisor.
Katharine Leney	Research supervisor.

Table 9.1: Contributions of team members to the project

Appendiex

- **In file `SubmitToGrid.py`:**
  - Defined a new flag `Submit22Trigger`.
  - By setting this flag to true, jobs corresponding to the trigger study with 2022 data/MC will be submitted.
  - By doing so, a pre-defined config file `PFlow/pflow.trig2022.config.mc23.txt` will be loaded.
- **Create a pre-defined config file `PFlow/pflow.trig2022.config.mc23.txt`:**
  - Defined a flag `PFlowSaveRun3Trigger` to control whether to perform Run2 or Run3 trigger studies.
  - The corresponding source files are `TriggerManagerRun2.cxx` and `TriggerManagerRun3.cxx`.
- **To run the ntuple production job for trigger study:**
  - The procedure is the same as the offline one. Please be aware of the filename changes.
  - **If you want to run this analysis locally:**
    - `source setup_local.sh`
    - `mkdir run && cd run`
    - Create a text file (e.g., "`input.txt`") that contains the full path to your input file(s).
    - Create an AnalysisTop config file (e.g., the one at `grid/PFlow/pflow.trig2022.config.mc23.txt`).
    - `top-xaod pflow.trig2022.config.mc23.txt input.txt`
  - **If you want to run this analysis on the Grid (Panda):**
    - `source setup_grid.sh`
    - `cd grid`
    - Sample lists are located at `MC_Samples.py` and `Data_Samples.py`.
    - Run with `SubmitToGrid.py`.
    - (The MC and data samples are exactly the same as the offline part.)

The structure of the code framework compared with the ofline part.

- **In file `htcondor_submit/submit_to_condor.py`:**
  - Defined a new flag `-onlineflag`.
  - By adding this flag to `run_options_for_all_jobs`, jobs corresponding to the trigger study with 2022 data/MC will be submitted (false by default).
  - With this flag, `finalSelection.py` will perform selection for the trigger part.

- When doing so, a source control function TrigfinalSelectionMC will be loaded in file. /TSelectors/source/Selectors/macros/runTSelectorMC.cxx (using finalSelectionMC do offline by default).

- **In file finalSelection.py:**
  - Added a choice of onlineflag: flag of whether to apply online calibration, false by default.
- **In file /TSelectors/source/Selectors/macros/runTSelectorMC.cxx:**
  - This is an outer control code for the source file: if onlineflag is true, using TrigfinalSelectionMC to perform the trigger part final selection, otherwise using finalSelectionMC for offline by default.
- **In file /TSelectors/source/Selectors/macros/libSelectorMC/:**
  - Defined the source file for selection.
  - TrigDilepChannel, TrigfinalSelectionMC, and TrigHistsForFitContainer files have been added.
  - If using TrigfinalSelectionMC in runTSelectorMC.cxx, this added Trig file will be used for the trigger part (will not be used when performing offline).
- **In file TrigfinalSelectionMC.cxx:**
  - The channels emu_OS_J2_online_emutrig, emu_OS_J2_online_match, and emu_OS_J2_online_tagoff60 have been saved for events that pass the bofftrigger; kinematics match between two offline jets; conditional that passing the offline WPs of GN2 first.
- **In file TrigDilepChannel:**
  - Each channel mentioned in the selection part will be saved in the format of TrigDilepChannel.
  - The TrigDilepChannel will use the functions predefined in TrigHistsForFitContainer to perform the fit.
- **In file TrigHistsForFitContainer:**
  - Saving the histograms for further fitting for the trigger part.
- **In file JetContainer:**
  - A minor modification to save the histograms of HTLGN1 and HLTDL1d, which will be used in the trigger calibration part.