ABSTRACT

SEARCH FOR CHARGED HIGGS BOSONS IN THE $\tau + \ell$ FINAL STATE WITH 36.1 fb^{-1} OF pp COLLISION DATA AT $\sqrt{s} = 13$ TeV WITH THE ATLAS EXPERIMENT

Elliot Wesley Parrish, Ph.D.
Department of Physics
Northern Illinois University, 2022
Dhiman Chakraborty and Jahred Adelman, Director

This dissertation uses 139 fb^{-1} of pp collision data collected at a center of mass energy of $\sqrt{s} = 13$ TeV by the ATLAS detector to search for charged Higgs bosons decaying to a tau lepton and a neutrino ($H^{\pm} \to \tau^{\pm}\nu_{\tau}$) in association with a leptonically decaying top quark. No significant excess was found, therefore limits are set at the 95% confidence level on the charged Higgs production cross section times the branching fraction into the $\tau^{\pm}\nu_{\tau}$ ranging from XX pb to XX fb. These limits are interpreted in the hMSSM benchmark scenario as an exclusion at 95% confidence on $\tan\beta$ as a function of $m_{H^{\pm}}$. In this scenario, for $\tan\beta = 60$, the $H^{\pm}$ mass range up to XXX GeV is excluded, with all values of $\tan\beta$ excluded for $m_{H^{\pm}} \leq XXX$ GeV.

NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

DECEMBER 2022

SEARCH FOR CHARGED HIGGS BOSONS IN THE $\tau + \ell$ FINAL STATE WITH 36.1 $\mathbf{fb}^{-1}$ OF pp COLLISION DATA AT $\sqrt{s} = 13$ TeV WITH THE ATLAS EXPERIMENT

BY

ELLIOT WESLEY PARRISH

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

Dissertation Director:
Dhiman Chakraborty and Jahred Adelman

ACKNOWLEDGEMENTS

DEDICATION

To Dr. Dhiman Chakraborty. Thank you for everything.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

EVENT RECONSTRUCTION

Before any physics analysis can be performed on the raw data from the ATLAS detector and MC simulations both raw data and MC go through a reconstruction software suite called Athena. Various algorithms are employed to identify energy deposits as particles based on shower shapes, tracker hits, calculated charge to mass ratios, etc. Figure 1.1 shows the signatures of various particles within the ATLAS detector.

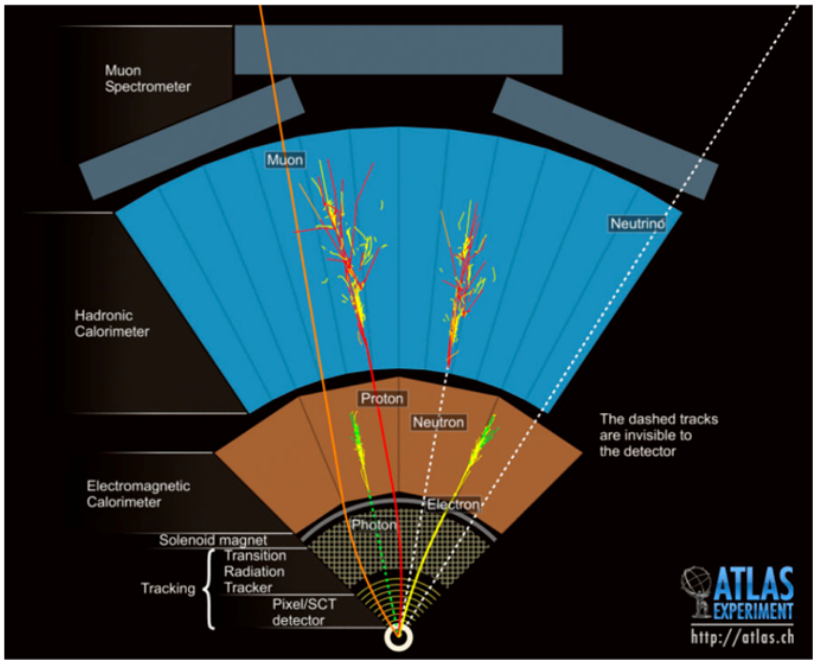

Figure 1.1: Cross section view of the ATLAS detector with subdetectors labeled. Various types of particles radial trajectories are shown.

The following sections detail the identification processes of muons, electrons, photons, jets, τ leptons, and a calculated quantity called missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$). These reconstructed physics objects are the inputs to the majority of physics analyses.

1.1 Tracks

Tracks are fits connected three dimensional space-points in the ID. These space-points are created from clusters of hits in the ID. A set of three space-points are combined into one track seed, then fed into three methods in the ATLAS detector: inside-out, outside-in, and TRT-standalone. The inside-out method creates tracks by starting with a seed hit in the pixel detector, then SCT hits are added, finally the track is extrapolated out into the TRT. This method creates tracks of particles that are mostly produced in the hard pp interaction and has a requirement of $p_{\mathrm{T}} > 400$ MeV. On the other-hand the outside-in method starts with track segments in the TRT and extrapolates towards the beamline using silicon that were not used in the inside-out method. Outside-in tracking typically reconstructs secondary vertices from particles that have long enough lifetimes to decay while inside the ID, including b quarks and τ leptons. Lastly, TRT-standalone tracks are made only from seeds within the TRT and are not extrapolated to the silicon subdetectors [1]. The reconstructed tracks are used in the identification of various types of particles.

1.2 Topological Clusters

A topocluster is defined as a cluster of topologically connected calorimeter cell signals. Topological clusters in the ATLAS detector's calorimeters are vital to the identification of hadronic final states, meaning jets (Section 1.5), isolated hadrons, and hadronically decaying τ leptons (Section 1.6). Topoclusters are also included in the calculation of missing transverse energy discussed in Section 1.7, as they represent the direction and energy of softer particles in a collision event.

A topocluster is created via a growing volume algorithm that operates based on a set of three thresholds. These thresholds are defined using the calorimeter cell significance ξ_{cell} [2].

$$\xi_{cell} = \frac{E_{cell}}{\sigma_{noise,cell}} \tag{1.1}$$

Where E_{cell} is the energy in the calorimeter cell and $\sigma_{noise,cell}$ is the average expected noise of a given calorimeter cell. An in-depth review of how the $\sigma_{noise,cell}$ value is calculated for TileCal is given in A. A topocluster starts with a seed cell that has a significance greater than the seed threshold S. From the seed cell, all three-dimensionally neighboring cells with a significance greater than the growth threshold N are added to the topocluster. This is done repeatedly until there are no more neighboring cells that pass the requirement $|\xi_{cell}| > N$. If a neighboring cell also passes the $|\xi_{cell}| > S$ threshold, then the topocluster corresponding to the neighbor cell is merged into the original topocluster. Finally, a last layer of the topocluster is added from all neighboring cells passing a threshold of $|\xi_{cell}| > P$. In the ATLAS experiment, the threshold values are set at $(S, N, P) = (4, 2, 0)$.

1.3 Muon Identification

Muons are identified using a combination of information from the ID and the MS. Within the ID, muons leave tracks identical to any other charged particle; however, in the MS tracks are identified within the MDTs through a straight-line fit in a single layer and by doing a combinatorial search of CSC hits in the $\eta - \phi$ plane. [3] Muons are identified through five strategies, each using the information from the ID, MS, and calorimeter (in one case).

- Combined (CB): Match ID and MS tracks. Perform a combined track fit on ID and MS hits. Takes into account energy loss in calorimeters

- Inside-Out (IO): Extrapolate ID tracks, look for at least three loosely aligned MS hits. Calorimeter energy loss is accounted for.
- Muon Spectrometer Extrapolated (ME): Extrapolate MS tracks back to the beamline. No ID hits are taken into account.
- Segmented-Tagged (ST): Extrapolate ID tracks and match to MS segments with tight angular requirement. Muon parameters are taken directly from the ID.
- Calorimeter-Tagged (CT): Extrapolate ID tracks into the calorimeters. Look for energy deposits consistent with minimum ionizing particles. Tag as muon, take parameters from ID.

All muon identification strategies have a transverse momentum cut on ID tracks of $p_{\mathrm{T}}^{track} > 2$ GeV, except for CT, which has a cut of $p_{\mathrm{T}}^{track} > 5$ GeV.

Reconstructed muons are divided into three WPs to allow analyzers a choice of purity, efficiency, and background rejection.

- Loose: Optimized for reconstruction of $H \to 4\mu$. Lowest purity and highest efficiency.
- Medium: Efficiency and purity are suitable for a wide range of analyses with small systematic uncertainties.
- Tight: High purity, slightly lower efficiency than medium WP. Significantly higher background rejection.

The analysis discussed in this dissertation uses the Loose WP for muons.

1.4 e γ Identification

Electrons and photons deposit the majority of their energy in the EM calorimeters in similar fashion. Electrons produce Bremsstrahlung photons as they interact with the EM calorimeter, the produced photons then convert into an electron-positron pair. This process repeats and produces a shower. A photon that is produced in the ID and travels to the EM calorimeters creates a very similar shower by converting into an electron-positron pair, thus creating a very similar shower. The discerning difference between an electron's signature and that of a photon is matching tracks. An electron carries an EM charge, thus leaving a track in the ID; whereas a photon does not carry an EM charge, therefore does not leave a track. The process of identifying an EM shower as either electron initiated or photon initiated is detailed in [4]. A brief algorithm flow chart of this process can be seen in Figure 1.2. If tracks in the ID are found to match a topocluster in the EM calorimeter, then it is identified as an electron, re-clustered into so called superclusters to ensure the full shower is captured, calibrated, then lastly made into an analysis object for use in physics analyses. The same algorithm is used to identify photons with the exception of matching tracks to the ID. Instead, photons are matched to conversion vertices where the initial photon first converted into an electron-positron pair. Both electrons and photons are reconstructed at three WPs. As with muons, there are three WPs, Loose, Medium, and Tight; the stricter WPs being subsets of the looser WPs.

To ensure that an electron or photon is indeed an initial particle and not part of another shower, whether it be from a converted photon in a hadron decay, electrons from heavy flavor hadrons or a light hadron mis-identified as an electron, an isolation variable is calculated. The isolation variable is based on track isolation and defined as the sum of transverse momenta

of all tracks within a cone around the electron candidate of $\Delta R = 0.2$ or in the case of high energy photons, 10 GeV/E_T. Where E_T is the tr erse energy of the electron.

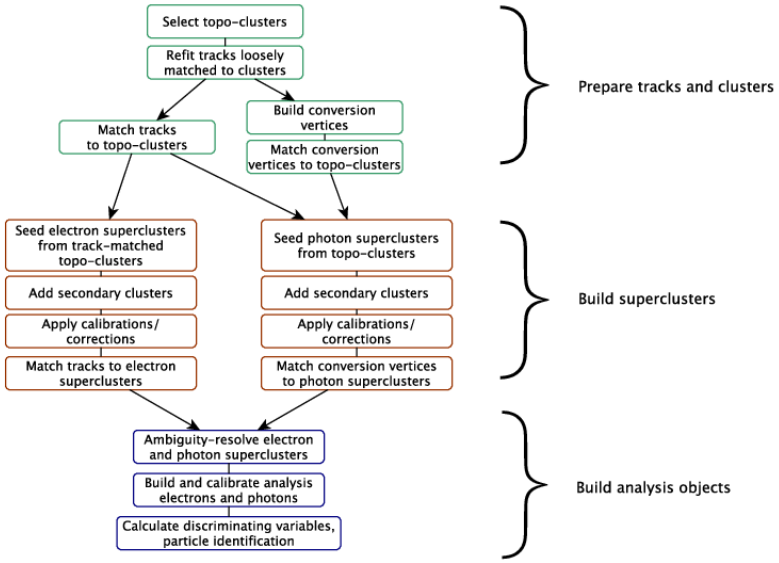

Figure 1.2: Algorithm flow diagram for the electron and photon reconstruction [4].

1.5 Jets

A jet is a reconstructed object of calorimeter energy that is meant to capture the energy of a hadronic shower, typically initiated from hard scatter quarks or hadrons. There are several algorithms available to perform a clustering of calorimeter topoclusters to form jets. This dissertation uses jets created from particle flow objects. The particle flow algorithm is described in detail], a flow chart of the algorithm is shown in Figure 1.3 and an idealized example of the particle flow algorithm performing the reconstruction of hadrons is shown in Figure 1.4.

The particle flow algorithm starts by matching selected tracks to a single topocluster. The expected energy of the initial particle in the calorimeter is calculated from the track

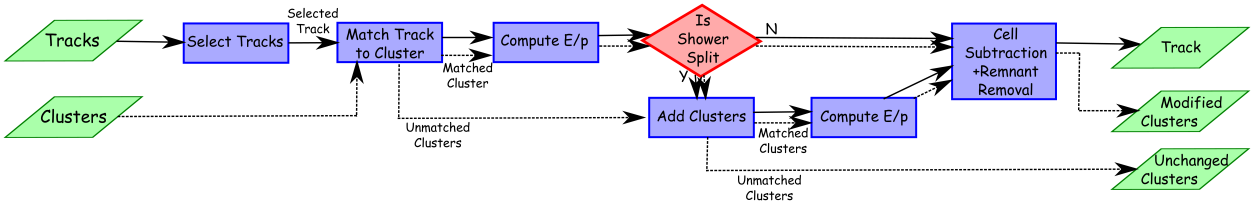

Figure 1.3: A flow chart of how the particle flow algorithm proceeds, starting with track selection and continuing until the energy associated with the selected tracks has been removed from the calorimeter. At the end, charged particles, topoclusters which have not been modified by the algorithm, and remnants of topoclusters which have had part of their energy removed remain [5].

momentum and the topocluster position. The probability of the track-topocluster system being deposited in multiple topoclusters is then calculated. The algorithm then adds in more topoclusters to the output object based on this probability. The expected energy of the initial particle is subtracted from the energy of the matched topoclusters cell by cell. If the energy of the output object is consistent with a single particle signal, then the remaining topocluster remnants are removed. The outputs of the particle flow algorithm are then fed into the anti-k_t algorithm [6] with a radius value of $R = 0.4$.

The anti-k_t algorithm is a jet finding algorithm that is collinear and infrared safe. Meaning the number of identified jets does not change due to splitting or merging of high transverse momentum particles, nor the presence of soft gluon emission between jets [7]. A jet is constructed in the anti-k_t algorithm through an iterative process using a the distance parameter defined as

$$d_{ij} = min(k_{t,i}^{-2}, k_{t,j}^{-2}) \frac{\Delta_{ij}^2}{R^2} \tag{1.2}$$

where k_t is the transverse momentum, R is an input parameter defining the radius of the jet cone, and Δ_{ij} is the distance between objects i and j defined as

$$\Delta ij = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \tag{1.3}$$

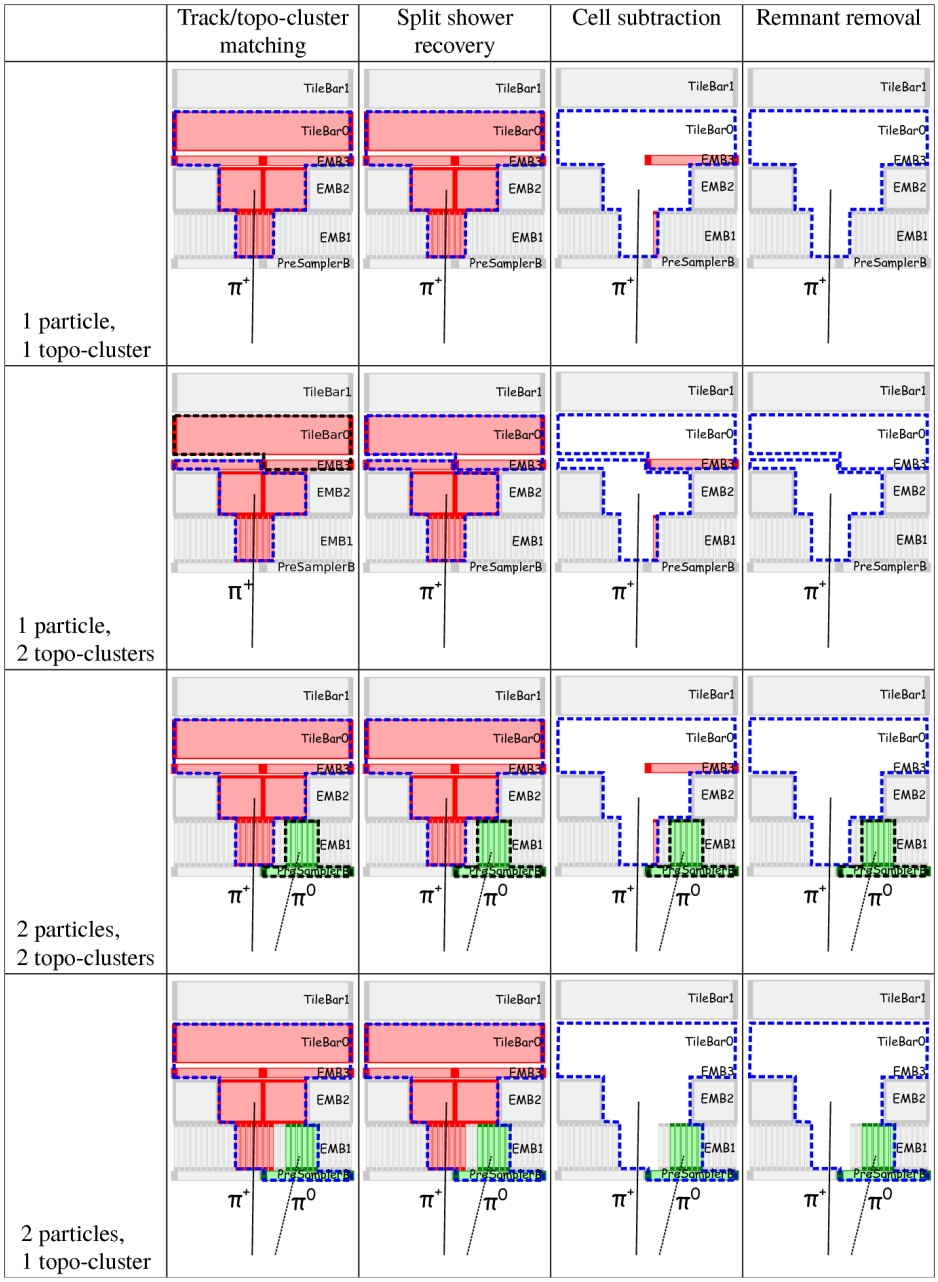

Figure 1.4: Idealized examples of how the algorithm is designed to deal with several different cases. The red cells are those which have energy from the π^+, the green have cells energy from the photons in the π^0 decay, the dotted lines represent the original topocluster boundaries with those outlined in blue having been matched by the algorithm to the π^+, while those in black are yet to be selected. The different layers in the EM calorimeter (Presampler, EMB1, EMB2, EMB3) are indicated. In this sketch only the first two layers of the Tile calorimeter are shown (TileBar0 and TileBar1) [5].

The anti-kt algorithm first identifies the smallest d_ij and clusters the particle flow objects if $d_{ij} > k_{t,i}^{-2}$. If $d_{ij} > k_{t,i}^{-2}$ then the particle flow object is discarded. This process continues iteratively until there are no more objects to consider. Objects with $\Delta > R$ are still considered, making the R input parameter an energy cut-off for clustering and not a direct radius value. Figure 1.5 shows the anti-k_t algorithm's performance compared to other jet finding algorithms. The anti-k_t algorithm results in a more conical shape than other jet finding algorithms; better encapsulating the shower profile of jets.

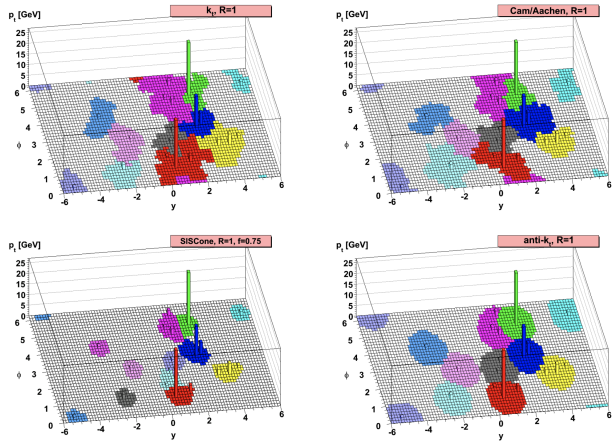

Figure 1.5: Comparison between several jet finding algorithms [6].

1.5.1 *b*-jet Tagging

Jets originating from hard scatter b quarks are an important signature in high energy physics colliders, especially so in the analysis discussed in this dissertation. An initial state b quark hadronizes into B-hadrons which have a relatively long lifetime. Due to the relativistic speeds and long lifetime of the B-hadrons they travel a distance away from the IP before decaying and creating a hadronic shower. This leads to a secondary vertex that can be

measured. A pictorial representation is shown in Figure 1.6. The impact parameter d_0 shown is the minimum distance between the tracks from the secondary vertex and IP.

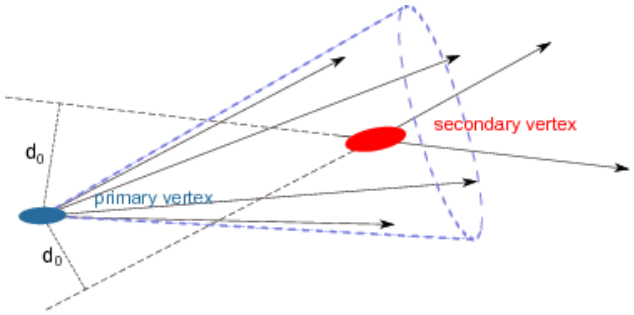

Figure 1.6: Schematic view of the tracks in a b-jet [8].

There are several methods used to tag a jet as coming from quark; this analysis uses the DL1 high level tagger [9] that is based on an Artificial Deep Neural Network. Neural networks are discussed in detail in Section ??. DL1 not only tags b-jets, but also outputs the probability for a jet to be initiated from a charm or light flavor quark. The DL1 tagger has over 20 input variables, including the p_T and η of jets [10]. The analysis discussed in this dissertation uses a fixed cut working point that corresponds to an average efficiency of 70% for b-jets in $t\bar{t}$ events.

1.6 τ Identification

One of the most important particles in the final state of the search discussed in this dissertation is the τ lepton. The decay channels of τ leptons make them difficult to reconstruct. A τ lepton can decay to hadrons, an electron, or a muon; in each decay mode, at least one neutrino is also present. The analysis discussed in this dissertation only considers τ leptons that have decayed hadronically. The hadronic decay mode consists of 1 or 3 charged hadrons ($\pi^{\pm}$), a neutrino, and possible neutral hadrons (π^0). A τ lepton decaying in this manner within the ATLAS detector leaves 1 or 3 tracks in the ID and collimated showers of energy

in the calorimeters; the neutrino does not interact with the ATLAS detector, therefore no direct signature is left behind. Reconstruction is done on the visible part of the hadronically decaying τ lepton, further referred to as $\tau_{had-vis}$ in the rest of this dissertation.

The $\tau_{had-vis}$ candidates start with an anti-k_t jet seed with $E_T > 10$ GeV in the calorim tracks and topoclusters within $\Delta R = 0.2$ are added to the $\tau_{had-vis}$ candidate. The axis of the original jet seed is redefined in the direction of the $\tau_{had-vis}$ candidate and calibration is done at the $\tau_{had-vis}$ scale. An lap removal is done to ensure the $\tau_{had-vis}$ is isolated from electrons and muons. Tracks are required to have $E_T > 30$ GeV, $|\eta| < 2.3$ and strictly either 1 or 3 tracks. A $\tau_{had-vis}$ candidate is referred to -prong or 3-prong based on the associated number of tracks. To discern $\tau_{had-vis}$ objects from quark and gl initiated jets a recurrent neural network (RNN) is . The search described in this analysis uses a medium WP that corresponds to 75% identification efficiency for 1-prong and 60% for 3-prong in $\gamma \to \tau\tau$ collision events. Only the highest p_{T} $\tau_{had-vis}$ candidate in an event is considered in this analysis.

1.7 $E_{\mathrm{T}}^{\mathrm{miss}}$

The final SM particle in the reconstruction scheme is the neutrino[1]. The presence of a neutrino, or another minimally interacting particle, can be inferred through the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$[2]; which takes advantage of the initial collision having a small momentum in the transverse plane ($p_T \simeq 0$). The initial momentum in the z direction (along the beamline) cannot be known due to the composite nature of the colliding protons and the associated PDFs of their components.

[1]The W, Z and gluon do not have long enough lifetimes to leave signatures within the ATLAS detector volume. Instead, their presence is inferred through their decay products

[2]The choice of $E_{\mathrm{T}}^{\mathrm{miss}}$ to represent missing transverse momentum is a common nomenclature. Other choices include p_T^{miss}, MET, and et miss.

The calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the ATLAS detector is defined as

$$E_T^{miss} = -\sum E_T = \sum p_T^{\mu} + \sum p_T^{e} + \sum p_T^{\gamma} + \sum p_T^{\tau} + \sum p_T^{jets} + \sum p_T^{soft} \tag{1.4}$$

where the p_T^{soft} term comes from soft tracks that are not associated with any physics objects [11]. The analysis discussed in this dissertation uses $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers in one of the subchannels to select events and is described in chapter **??**.

Appendices

APPENDIX

TILECAL DATA QUALITY

BIBLIOGRAPHY

[1] Garelli, Nicoletta on behalf of the ATLAS Collaboration.
"Performance of the ATLAS Detector in Run-2". *EPJ Web Conf.* 164 (2017), p. 01021.
DOI: 10.1051/epjconf/201716401021.
URL: https://doi.org/10.1051/epjconf/201716401021.

[2] T. A. Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". *The European Physical Journal C* 77.7 (2017), p. 490.
DOI: 10.1140/epjc/s10052-017-5004-5.
URL: https://doi.org/10.1140/epjc/s10052-017-5004-5.

[3] T. A. Collaboration. "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 *pp* collision data set at $\sqrt{s} = 13$ TeV".
Eur. Phys. J. C 81.7 (2021), p. 578. DOI: 10.1140/epjc/s10052-021-09233-2.
arXiv: 2012.00578 [hep-ex].

[4] T. A. Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data".
Journal of Instrumentation 14.12 (Dec. 2019), P12006–P12006.
DOI: 10.1088/1748-0221/14/12/p12006.
URL: https://doi.org/10.1088/1748-0221/14/12/p12006.

[5] T. A. Collaboration.
"Jet reconstruction and performance using particle flow with the ATLAS Detector".
The European Physical Journal C 77.7 (2017), p. 466.
DOI: 10.1140/epjc/s10052-017-5031-2.
URL: https://doi.org/10.1140/epjc/s10052-017-5031-2.

[6] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-k_T jet clustering algorithm".
Journal of High Energy Physics 2008.04 (Apr. 2008), pp. 063–063.

DOI: 10.1088/1126-6708/2008/04/063.

URL: https://doi.org/10.48550/arXiv.0802.1189.

[7] M. Cacciari, G. P. Salam, and G. Soyez. "The catchment area of jets". *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 005–005.

DOI: 10.1088/1126-6708/2008/04/005.

URL: https://doi.org/10.1088/1126-6708/2008/04/005.

[8] P. O. Hansson Adrian. *The ATLAS b-Jet Trigger*. Tech. rep.

Comments: 4 pages, 6 figures, conference proceedings for PIC2011.

Geneva: CERN, Nov. 2011. arXiv: 1111.4190.

URL: https://cds.cern.ch/record/1397942.

[9] *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-013. Geneva: CERN, July 2017.

URL: http://cds.cern.ch/record/2273281.

[10] *Expected performance of the ATLAS b-tagging algorithms in Run-2*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-022. Geneva: CERN, July 2015.

URL: https://cds.cern.ch/record/2037697.

[11] *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-023. Geneva: CERN, June 2018.

URL: https://cds.cern.ch/record/2625233.