Czech Technical University in Prague

Faculty of Nuclear Sciences and Physical Engineering

Department of Physics

Measurement of groomed jet $\log(k_T)$ in p+p collisons with the STAR experiment

RESEARCH PROJECT

Author: Bc. Georgij Ponimatkin

Supervisor: RNDr. Jana Bielčíková, Ph.D.

Year: 2020

České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská
Katedra fyziky

Studium $\log(k_T)$ v jetech měřených v p+p srážkách v experimentu STAR

VÝZKUMNÝ ÚKOL

Vypracoval: Bc. Georgij Ponimatkin
Vedoucí práce: RNDr. Jana Bielčíková, Ph.D.
Rok: 2020

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
FAKULTA JADERNÁ A FYZIKÁLNĚ INŽENÝRSKÁ
PRAHA 1 - STARÉ MĚSTO, BŘEHOVÁ 7 - PSČ 115 19

Katedra: fyziky **Akademický rok:** 2019/2020

VÝZKUMNÝ ÚKOL

***Student:* Bc. Georgij Ponimatkin**

***Studijní program:* Aplikace přírodních věd**

***Obor:* Experimentální jaderná a částicová fyzika**

***Vedoucí úkolu:* RNDr. Jana Bielčíková, Ph.D.**

Název úkolu (česky/anglicky):

Studium log(kT) v jetech měřených v p+p srážkách v experimentu STAR
Measurement of groomed jet log(kT) in p+p collisions with the STAR experiment

Pokyny pro vypracování:

1. Rešerše aktuálních výsledků týkajících se produkce jetů na urychlovačích RHIC a LHC se zaměřením na jejich vnitřní strukturu.
2. Větvící funkce v QCD, Lund rovina pro jety.
3. Popis experimentu STAR na urychlovači RHIC a jeho detektorů.
4. Studium log(kT) v tzv. „groomed" jetech v p+p srážkách v experimentu STAR
5. Dekonvoluce signálu pomocí Bayseovské metody a porovnání výsledků s Monte-Carlo simulacemi (např. PYTHIA).

Výzkumný úkol bude vypracován v anglickém jazyce.
Součástí zadání výzkumného úkolu je jeho uložení na webové stránky katedry fyziky.

Literatura:

1. K. Yagi et al: Quark-gluon plasma, Cambridge University Press, 2005
2. M. Cacciari, G. P. Salam and G. Soyez: FastJet User Manual, Eur. Phys. J C72 (2012) 1896
3. A. J. Larkoski, S. Marzani, G. Soyez and J. Thaler: Soft Drop, JHEP 05 (2014) 146
4. F. A. Dreyer, G. P. Salam and G. Soyez: The Lund Jet Plane, JHEP 1812 (2018) 064
5. G. D'Agostini: A multidimensional unfolding method based on Bayes' teorem, Nucl. Inst. Meth. A362 (1995) 487

Datum zadání: 25.10.2019

Datum odevzdání: 26.06.2020

....
vedoucí katedry

Prohlášení

Prohlašuji, že jsem svůj výzkumný úkol vypracoval samostatně a použil jsem pouze podklady (literaturu, projekty, SW atd.) uvedené v přiloženém seznamu.

Nemám závažný důvod proti použití tohoto školního díla ve smyslu § 60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne ....

Bc. Georgij Ponimatkin

Acknowledgment

I would like to thank Jana Bielčíková for advising me through this project and also for allowing me to have a research freedom this early. I also would like to thank Raghav Kunnawalkam Elayavalli and Joern Putschke for sharing their knowledge on different aspects of high energy physics data analysis for jet substructure. Last, but not least, I would like to thank my family for supporting me through my studies.

Bc. Georgij Ponimatkin

Název práce:
Studium log($\mathbf{k_T}$) v jetech měřených v p+p srážkách v experimentu STAR

Autor: Bc. Georgij Ponimatkin

Studijní program: Aplikace přírodních věd
Obor: Experimentální jaderná a částicová fyzika
Druh práce: Výzkumný úkol

Vedoucí práce: RNDr. Jana Bielčíková, Ph.D.
Ústav jaderné fyziky AV ČR, v.v.i.

Abstrakt: Měření jetové substruktury je jedním ze základních nástrojů pro precizní studium vlastností kvantové chromodynamiky (QCD). Jednou z takových technik je měření vlastností groomed jetové substruktury, umožňující studium vlastností partonové spršky, která stojí na počátku evoluce jetu. Většina současných poznatků týkající se groomed jetové substruktury dnes ovšem pochází z měření, která nezahrnují vzdálenost ΔR mezi měřenými subjety v proceduře SoftDrop. Tato práce navazuje na předešlá měření jetové substruktury na experimentu STAR a rozšiřuje je o měření současně differenciální v ΔR. Získané výsledky jsou následně porovnány s Monte-Carlo generátorem.
Klíčová slova: jetová substruktura, Lund rovina, Bayesovská dekonvoluce

Title:
Measurement of groomed jet log($\mathbf{k_T}$) in p+p collisons with the STAR experiment
Author: Bc. Georgij Ponimatkin

Abstract: Jet substructure measurements are one of the basic tools that allow precision studies of Quantum Chromodynamics (QCD). One of such techniques are groomed jet substructure measurements, that allow to study properties of parton showers, that lie at the begining of the jet evolution. Majority of current groomed jet substructure measurements is not taking into account the distance ΔR between measured subjets in the SoftDrop procedure. This work extends previous measurements of the jet substructure in the STAR experiment by a measurement that is also differential in ΔR. Results are also compared with a Monte-Carlo generator.
Key words: jet substructure, Lund plane, Bayesian unfolding

Contents

List of Figures

Introduction

The jet substructure is currently one of the most researched topics in the field of high energy physics. The particular interest rises from the fact, that those measurements bring us to the QCD precision era, where substructure measurements can potentially unveil a lot of information about internal workings of QCD, which are inaccessible without substructure measurements.

This research project deals with the measurement of a double-differential jet substructure in p+p collisions at $\sqrt{s} = 200$ GeV in the STAR experiment. Previously, the STAR collaboration has published [11] results of inclusive jet substructure measurements in p+p collisions. This project expands this measurement by explicitly studying substructure variables dependent on transverse momentum (p_T) as well as distance between sub-jets ΔR. Such measurement will provide unprecedented constraint on the models and will serve as an important baseline for measurements in heavy-ion collisions.

Chapter 1

Splitting Functions, Jet Substructure and Lund Plane

High-energy QCD processes give rise to jets - collimated showers of particles that arise as a consequence of high-Q^2 (with Q^2 being the square of the momentum transfer) processes. In those processes, a highly virtual parton pair is being produced. As a consequence of high virtuality, these partons start to radiate and hence a parton shower is created. For $Q^2 \sim \Lambda_{QCD}$ the non-perturbative transition to hadronic state happens. At this stage we obtain experimentally measurable jets, which are thus a consequence of parton shower evolution.

1.1 Parton Showers and Lund Plane

The probability for a parton to radiate another parton(s) with momentum fraction z is described by the **QCD splitting functions** [13, 14, 15], which are at leading order (LO) given by

$$P_{q\to qg}(z) = C_F \left(\frac{1+z^2}{1-z}\right), \tag{1.1}$$

$$P_{q\to gq}(z) = C_F \frac{1+(1-z)^2}{z,} \tag{1.2}$$

$$P_{g\to q\bar{q}} = \mathrm{Tr}\left(z^2+(1-z)^2\right), \tag{1.3}$$

$$P_{q\to gg} = 2C_A \left(\frac{1}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right). \tag{1.4}$$

Here z is the shared momentum fraction at splitting, $C_F = 4/3$ and $C_A = 3$ are colour factors. It is important to note, that the probabilistic interpretation is valid only for LO form of splitting functions. A schematic representation of $q \to qg$ splitting can be seen in the left panel of Fig. 1.1. Series of two such emissions forming a simple parton shower can be seen in the right panel of Fig. 1.1.

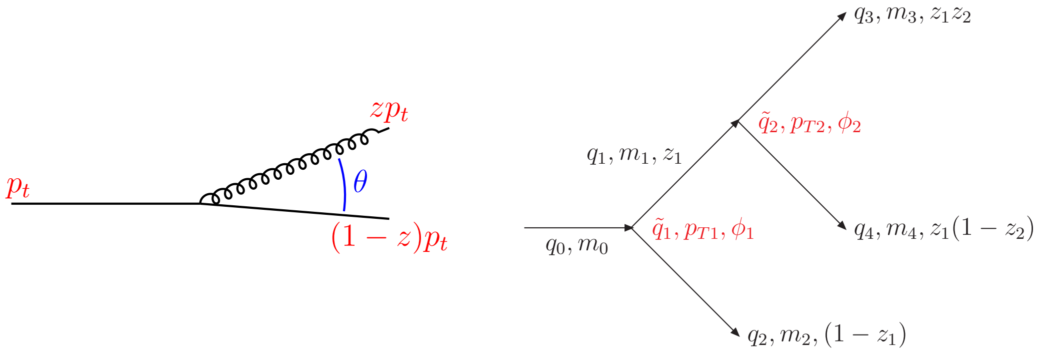

Figure 1.1: (left) Illustration of $q \to qg$ splitting. Taken from [1]. (right) Illustration of two sequential QCD splittings. Taken from [2].

As a next step, one can try to map phase space of emissions that can happen in a certain jet population. In order to do so, one needs the Lund plane [3] formalism. Assuming we have a jet that was reconstructed using the C/A algorithm [16, 17] we can traverse back the clustering history via hardest branch (i.e. splitting with the highest p_T). Assume that at each step we have two branches a and b with corresponding $p_{T,a}$ and $p_{T,b}$ such that $p_{T,a} > p_{T,b}$. Then we can construct a series of observables (here following the notation from [3])

$$\Delta_{ab} \equiv \Delta = \sqrt{(\eta_a - \eta_b)^2 + (\phi_a - \phi_b)^2}, \quad k_T = p_{T,b}\Delta_{ab}, \quad z = \frac{p_{T,b}}{p_{T,a} + p_{T,b}}, \tag{1.5}$$

here η_i is the pseudorapidity coordinate of the subjet i and φ_i is the azimuthal coordinate of the subjet i. After going through the full jet clustering history we can construct an average Lund map as a double-differential distribution of the form

$$\rho(\Delta, k_T) = \frac{1}{N_{jets}} \frac{dN}{d\log(k_T) d\log(1/\Delta)}, \tag{1.6}$$

where N_{jets} is a number of jets used to construct the emission diagram and $\frac{dN}{d\log(k_T)d\log(1/\Delta)}$ is a number of jets with particular distance between subjets $\log(1/\Delta)$ and emission strength $\log(k_T)$. An example of such map can be seen on Fig. 1.2.

For quark initiated jets at the leading order of perturbative QCD and for $\Delta \ll 1$ we can write

$$\rho \simeq \frac{\alpha_S(k_T)}{\pi} C_F \bar{z} \left(P_{gq}(\bar{z}) + P_{gq}(1-\bar{z})\right), \; \bar{z} = \frac{k_T}{p_{T,jet}\Delta}, \tag{1.7}$$

with $\alpha_S(k_T)$ being the strong coupling constant dependent on k_T and $C_F = \frac{4}{3}$ being the colour factor. Hence we see that there is a direct connection between the Lund jet plane and QCD splitting functions.

1.2 Jet Grooming

In experiments, the jet measurements are often contaminated by soft contribution, which is a consequence of the underlying event as well as non-perturbative transition

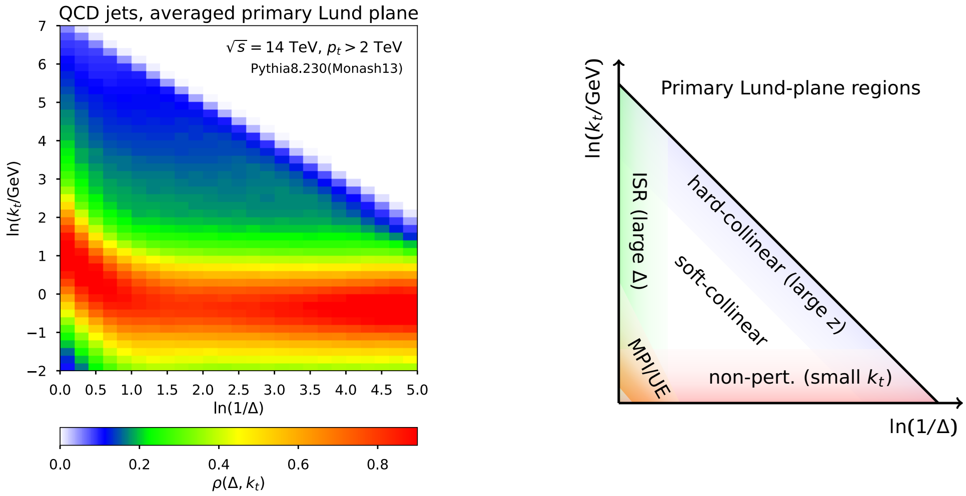

Figure 1.2: (left) An average Lund map for QCD jets with $R = 1.0$ reconstructed with C/A algorithm and $p_{T,jet} > 2$ TeV in p+p collisions at $\sqrt{s} = 14$ TeV. (right) Explanation of different regions in the Lund plane. Taken from [3].

that happens during hadronization. This background can significantly change the outcome of the jet substructure measurement, hence it is a good idea to eliminate such effects. One of such techniques is called Soft Drop [18], which is used in the analysis outlined further on.

This technique starts with the assumption that the jet was clustered with the anti-k_T algorithm [19]. As a next step, we recluster the jet with the C/A algorithm, in order to restore angular ordering, that is natural to QCD (since the anti-k_T algorithm is not "physical" in a sense that it's clustering happens around a hard core). After this we start traversing clustering history following the hardest branch. At each split we impose the condition

$$z_g = \frac{\min(p_{T,a}, p_{T,b})}{p_{T,a} + p_{T,b}} > z_{cut} \left(\frac{\Delta_{ab}}{R} \right)^{\beta}, \tag{1.8}$$

here $p_{T,a}$ and $p_{T,b}$ are transverse momenta of jet branches under consideration, Δ_{ab} is the distance between those branches in $\eta - \phi$ plane, z_{cut} is a minimal shared momentum fraction and R is a jet radius under consideration and β angular exponent. The first split that satisfied the imposed condition is accepted as a Soft Dropped (SD) split. Jet that does not have such split is discarded. Since Soft Drop is a cut on the jet emission phase space, it can be represented as a cut in the Lund plane. This can be seen in Fig. 1.3.

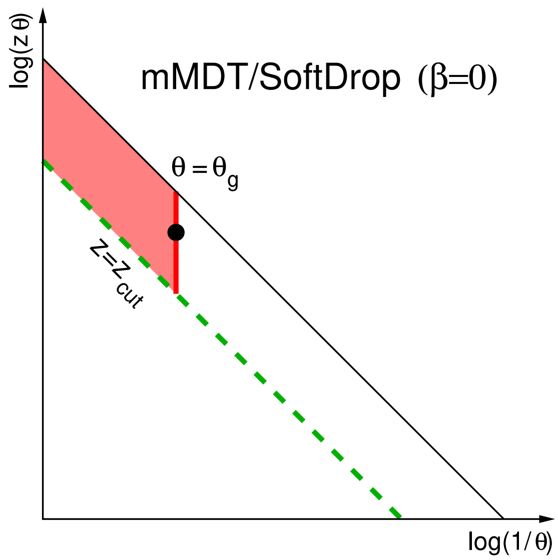

Figure 1.3: Representation of Soft Drop condition with $\beta = 0$ with respect to Lund plane. Taken from [4].

Chapter 2

STAR Experiment and RHIC

The RHIC particle accelerator, which is located at the Brookhaven National Laboratory (BNL) is a multi-purpose collider facility, whose primary designation are studies of nuclear matter properties at high energies as well as proton spin structure. One of the key properties of this machine is possibility to collide different kinds of systems, such as Au+Au, U+U, Cu+Cu but also asymmetrical systems such as Au+Cu, d+Au with variable energies of $\sqrt{s_{\mathrm{NN}}}$ from 7.7 to 200 GeV per nucleon pair. Currently RHIC is also the only collider worldwide that can accelerate spin-polarized protons. At the beginning there were 4 experiments at RHIC - BRAHMS [20], PHOBOS [21], PHENIX [22] and STAR [23]. Currently, the only active experiment is STAR, with sPHENIX experiment (the successor of PHENIX experiment) being constructed. The sPHENIX experiment will be devoted to high-precision measurements of hard-probes both in p+p as well as Au+Au collisions at RHIC energies [24]. In 2021 the construction of the STAR experiment forward upgrade will be completed. This upgrade will primarily serve for measurements of proton structure and is considered to be a step forward towards Electron Ion Collider physics.

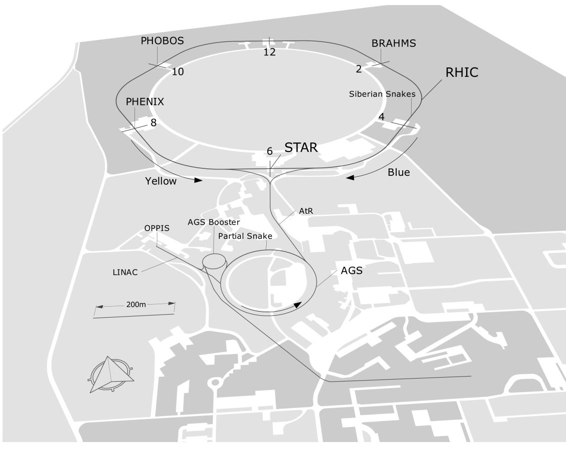

Figure 2.1: RHIC scheme. Taken from [5].

In next decade the RHIC facility will be modified and the Electron Ion Collider (EIC) will be built. The main purpose of this machine will be precise studies of nuclear structure via measuring new types of distributions - transverse momentum distributions (TMD) and generalized parton distributions (GPD). Also precise studies of hadronization are planned. More of the physics motivations can be found in [25].

2.1 STAR Experiment

The STAR experiment is a 4π multi-purpose detector built around a solenoidal magnet with magnetic induction $B = 0.5$ T which enables to measure particle momentum. The cross-section of the STAR experiment be seen on Fig. 2.2. The main sub-detector systems that are related to the jet physics studies are going to be explained bellow.

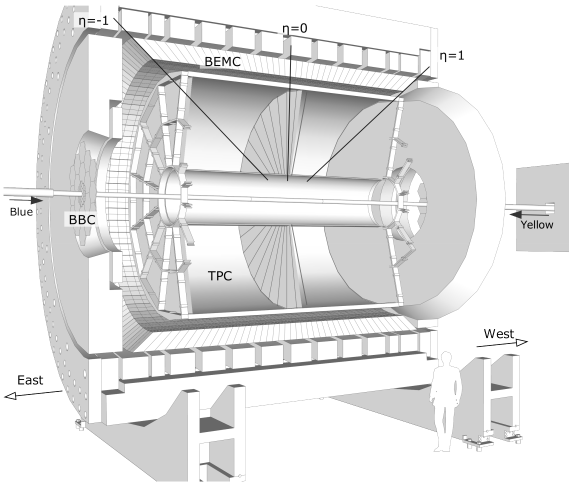

Figure 2.2: Cross-section of the STAR experiment. Taken from [5].

2.1.1 TPC

The Time Projection Chamber (TPC) [26] is a gaseous detector used for measurement and identification of charged particles. Its schematic construction can be seen on Fig. 2.3. The detector is held under a high voltage between two sectors, which are separated by a cathode membrane. When particle passes through the TPC it creates electron-ion ionization pairs. Ion pairs then drift to the central cathode while electrons drift to the end of detector, which is made of multi-wire proportional chamber pads. By construction, each of those pads provides radial $r - \varphi$ coordinates of the track, while z coordinate is obtained from the drift time. Because number of created electron-ion pairs is directly proportional to the energy of the particle, the TPC allows to measure energy loss ($\frac{dE}{dx}$) of the particle and hence provides particle identification (PID) capabilities. The original TPC had a full azimuthal coverage with pseudorapidity acceptance of $|\eta| < 1$ and minimal transverse momentum threshold of $p_T > 0.1$ GeV/c. Recently, the STAR TPC was upgraded to the iTPC [27], which has enhanced pseudorapidity coverage of $|\eta| < 1.5$ as well as lower minimal momentum threshold of $p_T > 0.05$ GeV/c.

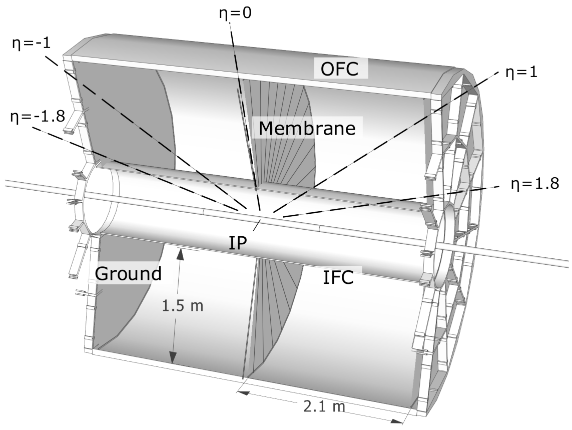

Figure 2.3: STAR Time Projection Chamber (TPC). Taken from [5].

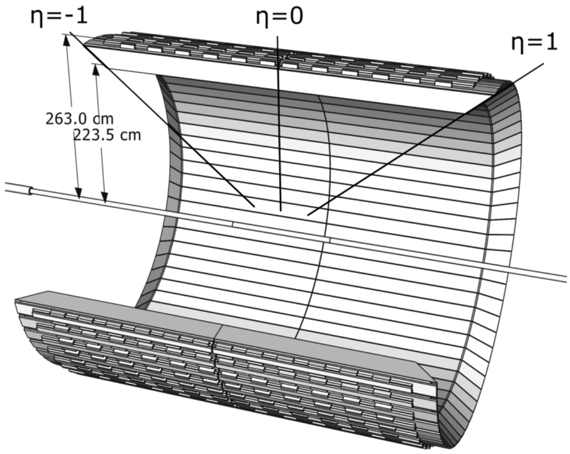

Figure 2.4: STAR Barrel Electromagnetic Calorimeter (BEMC). Taken from [5].

2.1.2 BEMC

The Barrel Electromagnetic Calorimeter (BEMC) [28], which cross-section can be seen on Fig. 2.4 is another detector in the STAR experiment, whose main purpose is measurement of neutral particles as well as event triggering. It works on the principle of a sampling calorimeter, where particle enters an absorber, which produces a cascade. This cascade is then read-out by the scintillator, afterwards particle encounters another absorber e.t.c. until the particle energy is exhausted. The BEMC is made of towers, each of the towers covers 0.05 units in η and 0.05 units in φ. An example of the BEMC module can be seen on Fig. 2.5 (left), while towers can be seen on the right.

2.1.3 STAR Forward Upgrade

In 2021 the STAR experiment will be upgraded with forward calorimeter (hadronic and electromagnetic) and forward silicon tracking systems [6]. This upgrade will

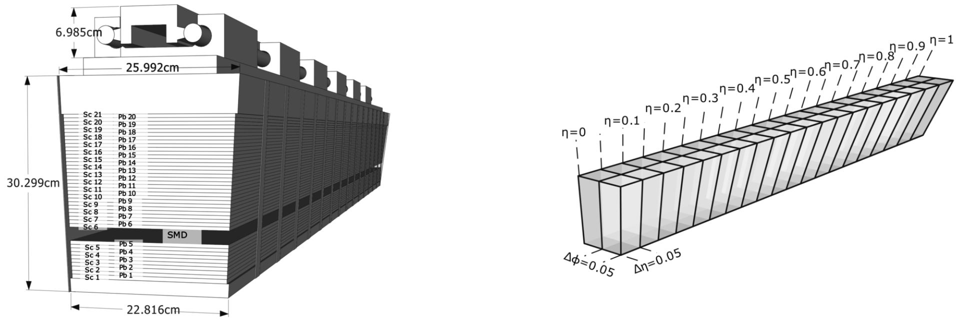

Figure 2.5: (left) BEMC module assembly. (right) BEMC towers in the module. Taken from [5].

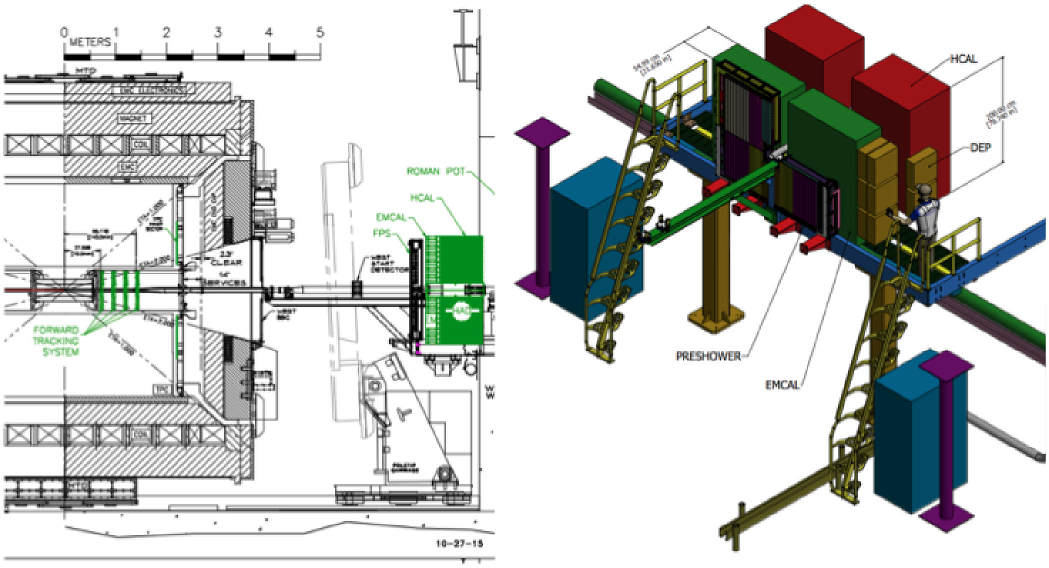

Figure 2.6: Schematics of the STAR forward upgrade. Taken from [6].

allow to conduct more thorough proton structure measurements thanks to the ability to access different x values (with x being fraction of the proton momenta carried by the quark), for example via selection of specific dijet population. The schematics of forward upgrade can be seen on Fig. 2.6. This upgrade will serve a step forward towards Electron Ion Collider (EIC) facility that is going to be built at Brookhaven National Laboratory. In the context of jet physics measurements, this upgrade will enable jet studies at forward rapidities.

Chapter 3

Recent Results on Jet Substructure Measurements

This chapter will deal with a brief review of the recent jet substructure measurements at RHIC and LHC. Currently, there exist only two published measurements of jet substructure in p+p or A+A collisions from the perspective of pure QCD studies.

3.1 Measurements of groomed shared momentum fraction (z_g)

The first measurement is a measurement by the CMS experiment [9], which is studying z_g distribution in p+p and $Pb + Pb$ collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The results can be seen of Fig. 3.1.

This measurement is qualitatively described by modern event generators (especially `PYTHIA` [7, 8]). This is a consequence of the fact, that z_g converges to the standard splitting function, which is an essential part of every event generator. In order to estimate the effect of splitting function modification due to the medium, the ratio plot between Pb+Pb distribution and p+p distribution is taken. Surprisingly, for Pb+Pb collisions with centrality above 30% there is no significant modification observed. In $0 - 30\%$ centrality region we observe slight biasing of z_g distribution towards asymmetric splitting, which might be the consequence of medium.

Another measurement comes from the ALICE experiment [10], which does similar measurement both in p+p and Pb+Pb collisions. The results can be seen of Fig. 3.2.

The p+p part of this measurement is in line with the result from the CMS collaboration, namely the fact that most event generators can describe z_g distribution with reasonable accuracy. We note that in contrast with the CMS measurement, the ALICE data are unfolded in p+p collisions. The Pb+Pb part of this measurement is different. It shows that with change of distance between subjets we see change in the behaviour of z_g distribution in the medium. This change is much more significant

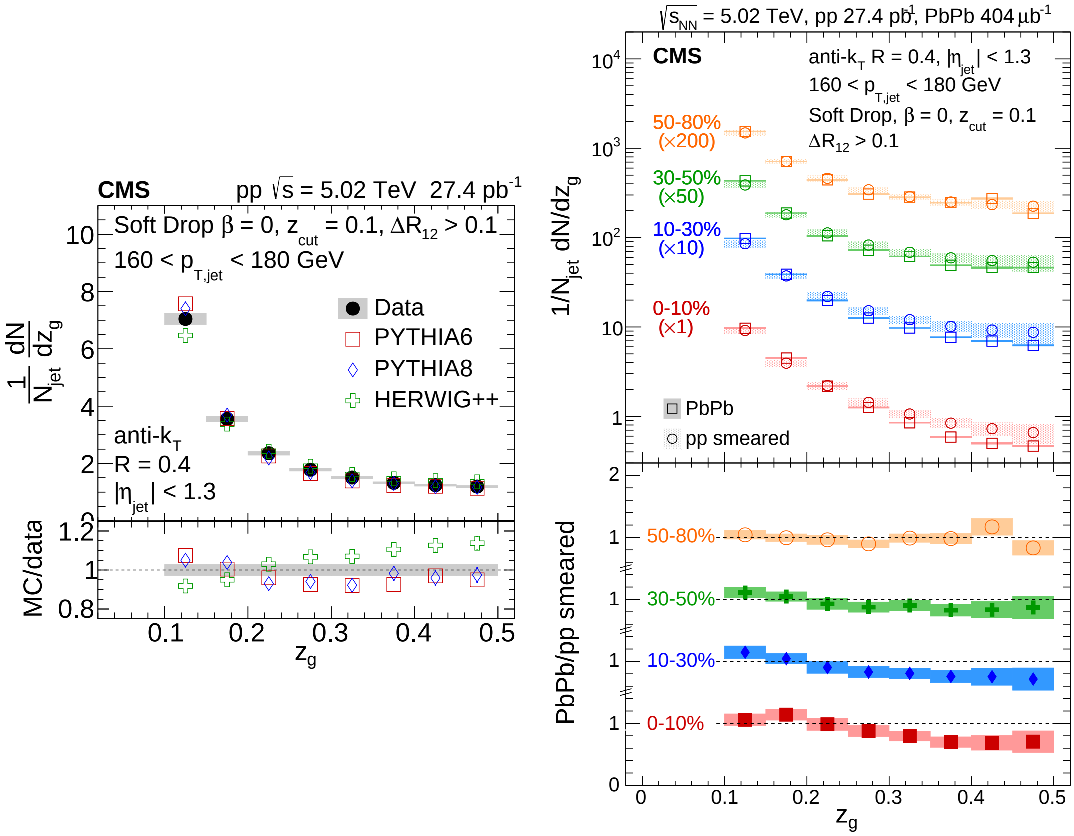

Figure 3.1: (left) z_g distribution in p+p collisions at $\sqrt{s} = 5.02$ TeV measured by the CMS collaboration which is compared with available event generators. In general, `PYTHIA6` [7] and `PYTHIA8` [8] models are providing qualitative description of measured data. (right) Detector-level z_g distribution in Pb+Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV for different centrality regions. Lower panel shows ratio of z_g in Pb+Pb collisions to z_g in p+p collisions that were smeared in order to account for deterioration of jet resolution in Pb+Pb collisions. Taken from [9].

than what is observed by the CMS collaboration in inclusive ΔR range. This hints that angular separation might play an important role for z_g, which is one of the motivation for the analysis done in this project.

Lastly, we present recent result by the STAR experiment [11], that repeats the same measurement for p+p collisions at $\sqrt{s} = 200$ GeV. One of the results can be seen on Fig. 3.3.

Again, this measurement confirms observations from the LHC, where the z_g distribution can be reasonably described by Monte-Carlo event generators (but in this case `PYTHIA8 Monash` [29] tune fails for $R = 0.2$ and $p_{T,jet} \in [15, 20]$ GeV/c).

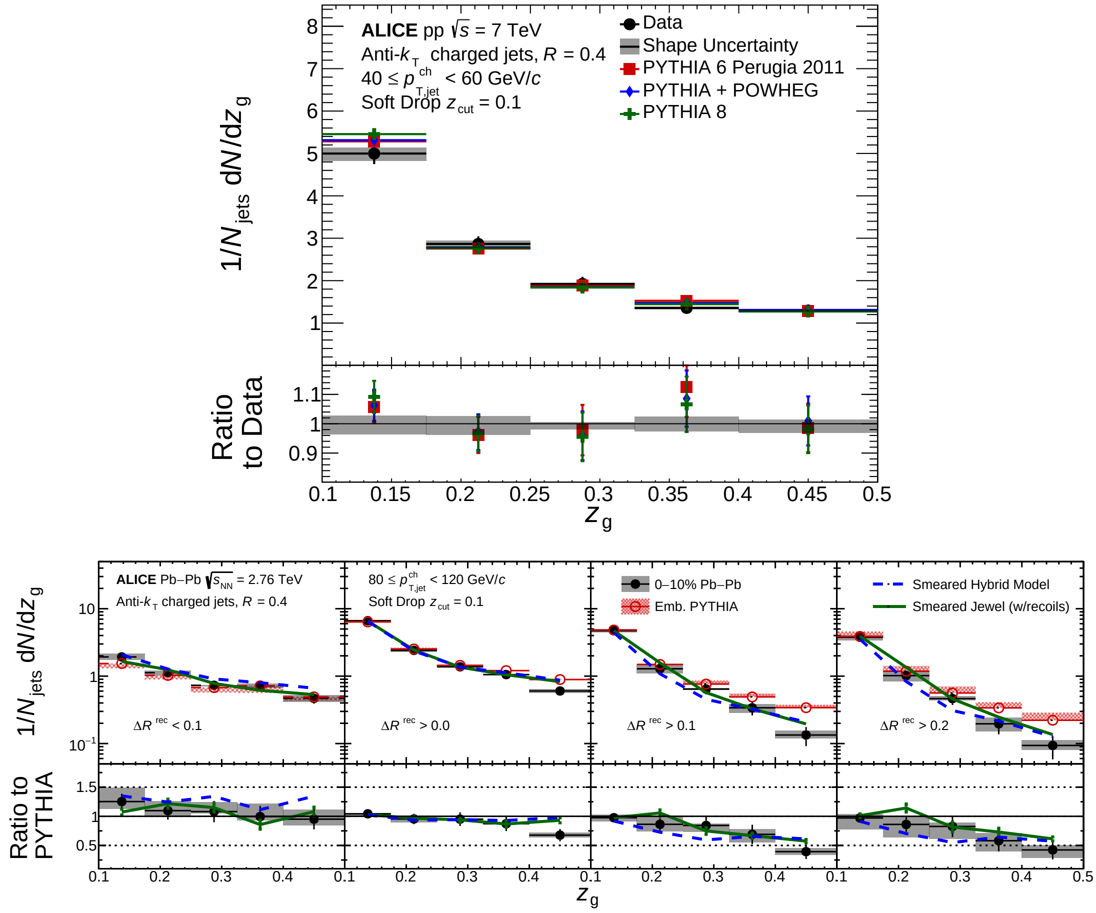

Figure 3.2: (left) Fully unfolded z_g distribution in p+p collisions at $\sqrt{s} = 5.02$ TeV measured by the ALICE collaboration which is compared with available event generators. In line with the CMS measurement, models again qualitatively describe measured data. (right) Detector-level z_g distribution in Pb+Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV for different subjet separation cut-offs ΔR^{rec}. Lower panel depicts ratio relative to the `PYTHIA` reference.. Blue and black curves represent theoretical descriptions by models. Taken from [10].

3.2 Lund Plane Measurement

To this day, there is only one published Lund Plane measurement, coming from the ATLAS experiment at the LHC [12]. The results can be seen on Fig. 3.4.

Here we can see the first measurement of Lund Plane slices in p+p collisions at $\sqrt{s} = 13$ TeV. The results are represented as $\log(1/z)$ distributions for different distances between the splits. Again, those distributions are relatively well described by Monte-Carlo event generators. This is because this measurement is, in essence, a z distribution.

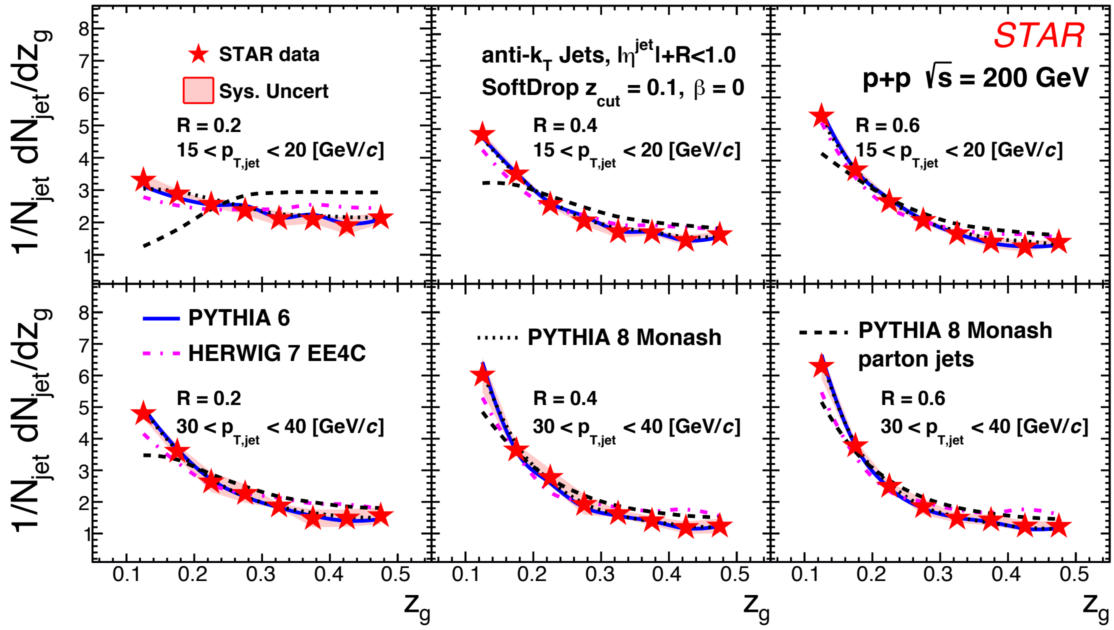

Figure 3.3: Fully unfolded z_g distribution in p+p collisions at $\sqrt{s} = 200$ GeV measured by the STAR collaboration which is compared with available event generators. Each column depicts z_g distribution for different jet resolution parameter R. Taken from [11].

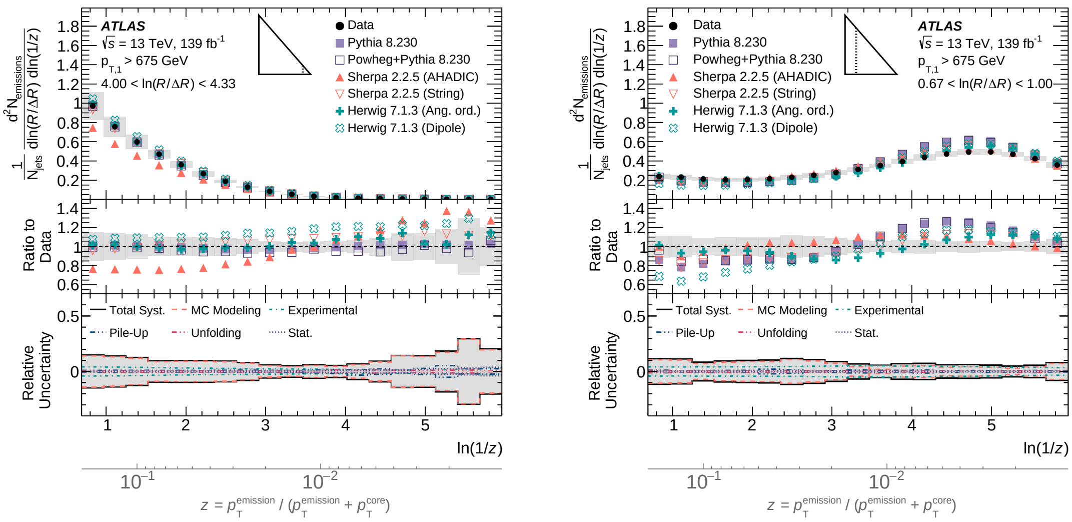

Figure 3.4: Fully unfolded $\log(1/z)$ distributions in p+p collisions at $\sqrt{s} = 13$ TeV measured by the ATLAS collaboration. Distribution on the left corresponds to the large angle emissions, while distribution on the right corresponds to the collinear splits. Taken from [12].

Chapter 4

Measuring Groomed Jet Substructure

This chapter will explain technical details of the analysis that was done in this project. This analysis uses 1.9 billion events measured in p+p collisions at $\sqrt{s} = 200$ GeV by the STAR experiment in 2012. For this analysis we used JP2 (jet patch) triggered events, which require a presence of at least 9.8 GeV/c of deposited energy in BEMC sectors. This trigger essentially allows us to select event population that guarantees to have jet-like objects.

The analysis procedure first begins with skimming over the picoDst files at the RACF computing facility at BNL. Here we do preprocessing of the available data in event-by-event fashion. First, the quality of events is checked for:

1. Event is verified to be in a good run.
2. Event is verified to have the JP2 trigger fired.
3. The correct collision vertex position is checked.

The overall event statistics after each cut can be seen on Fig. 4.1. After event quality is checked, we proceed towards event analysis:

1. All charged particles with $p_T \in [0.2, 30]$ GeV/c and $|\eta| < 1$ ranges are accepted. The histograms of accepted track η, ϕ and p_T can be seen on Fig. 4.2.
2. Next, the hits from the BEMC are extracted. In order to remove charged particle contribution from the tower energy, we project particles into each corresponding tower and apply hadronic correction, namely

$$E_{corr} = E - f_{\text{hadronic}} \sum_i |\boldsymbol{p}|_i. \tag{4.1}$$

   Here we select full hadronic correction, i.e. $f_{hadronic} = 100\%$. E here represents raw energy value extracted from the BEMC and E_{corr} is the tower energy with subtracted contribution of charged particles. If corresponding energy is

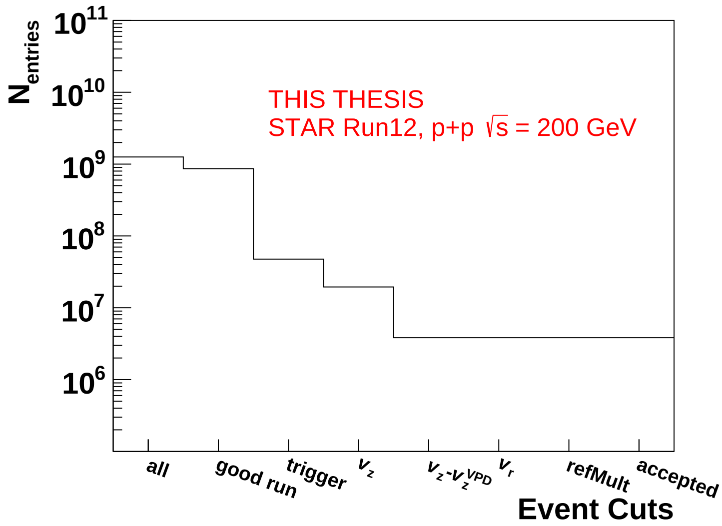

Figure 4.1: Number of events after each event quality selection steps.

negative or the tower is in the list of bad towers the corresponding tower energy is set to zero. Good hits are then converted into E_T (transverse energy) as

$$E_{T,i} = \frac{E_i}{\cosh(\eta_i)}, \tag{4.2}$$

where E_i and η_i are energy and pseudorapidity of tower i. The histograms of accepted tower η, ϕ and E_T can be seen on Fig. 4.3 and histograms of track-tower matches can be seen on Fig. 4.4.

3. Accepted charged and neutral parts of the event are then clustered together using the anti-k_T algorithm implemented within FastJet package [30, 31]. In this analysis the resolution parameter of $R = 0.6$ is used and all track are assumed to have pion mass. The histograms of accepted jet p_T and $\eta - \varphi$ can be seen on Fig. 4.5.

4. The clustered jets are then reclustered using the C/A algorithm also implemented within the FastJet package, and passed through the Soft-Drop condition. The substructure observables are then constructed using the accepted jets.

Unfortunately, direct measurement of the Lund plane is very complicated because of the three facts - finite resolution of detector, which forces the use of binning, the limited statistics and a need to unfold resulting distributions. Because of that the analysis strategy is modified. First, we omit $\ln\left(\frac{1}{\Delta R}\right)$ dependence in the Lund plane and change it to ΔR. This way we make our measurement more intuitive, while keeping the physics message unchanged (since $\ln(x)$ is bijective on the positive domain). Second, we add another observable, which we call $p_{T,b}$, that quantifies absolute strength of emission at soft-dropped split.

With that in mind, we thus construct the following observables:

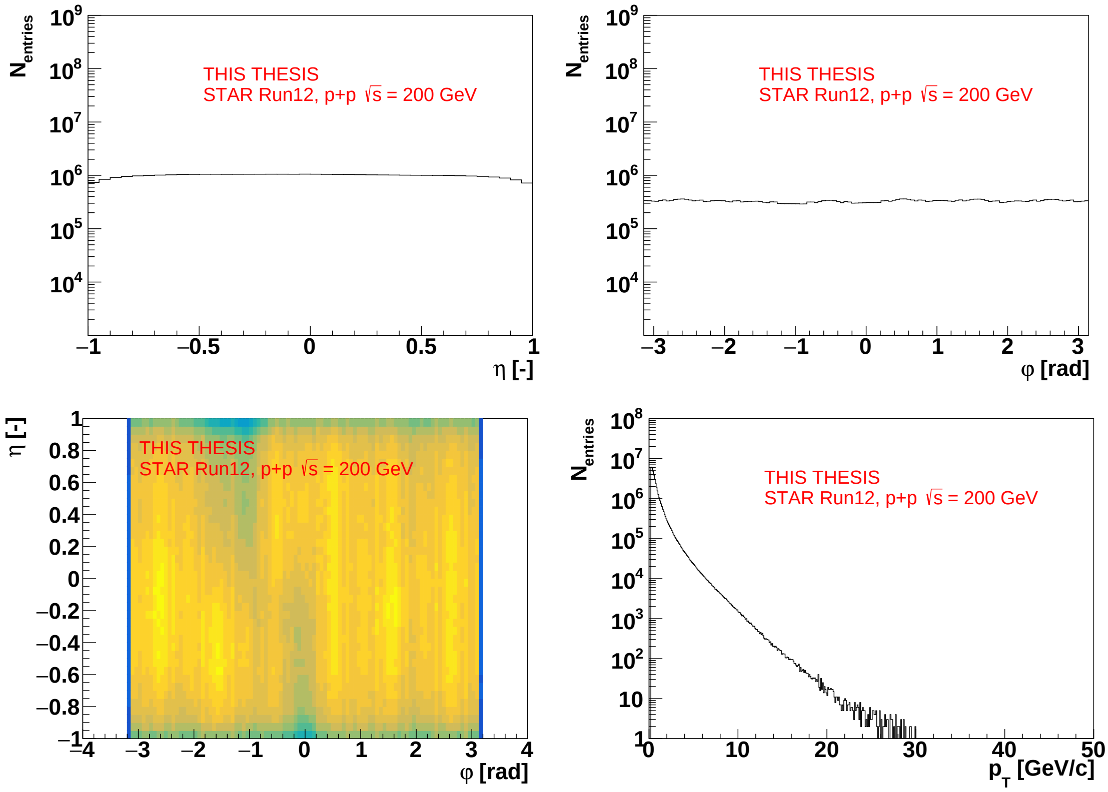

Figure 4.2: η distribution of accepted track (upper left), φ distributions of accepted tracks (upper right), $\eta - \varphi$ distribution of accepted tracks (lower left) and p_T distribution of accepted tracks (lower right).

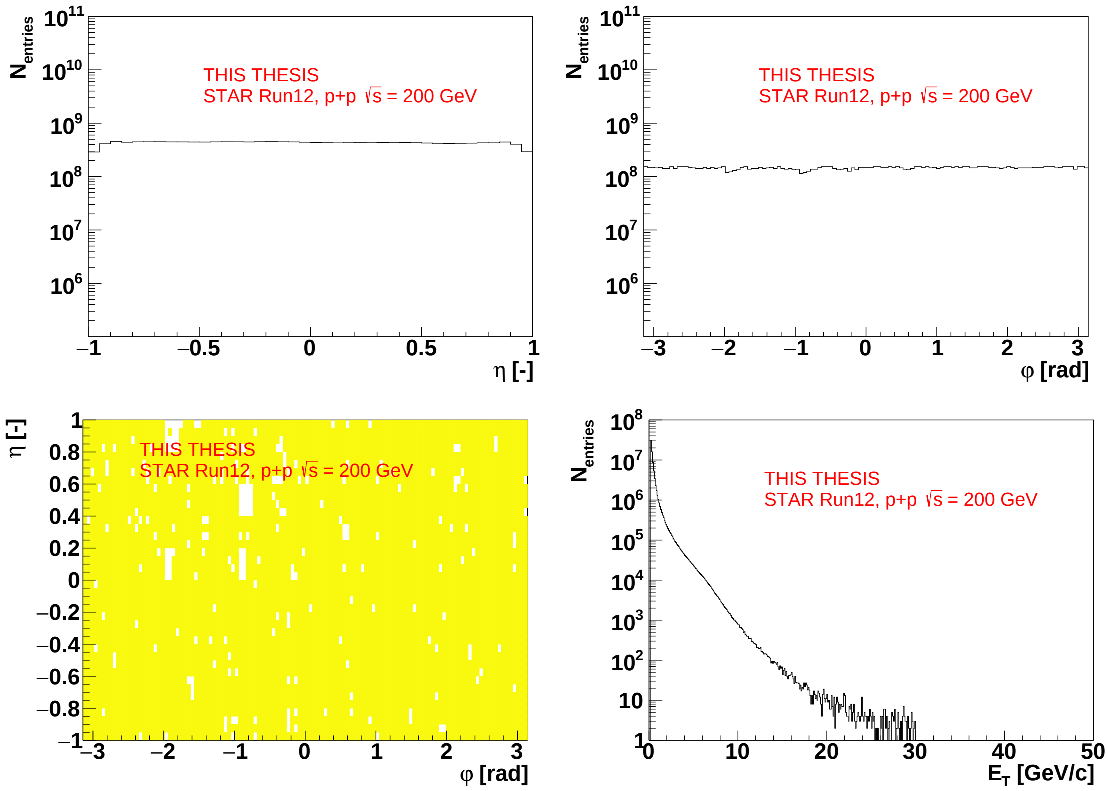

Figure 4.3: η distribution of accepted towers (upper left), φ distributions of accepted towers (upper right), $\eta - \varphi$ distribution of accepted towers (lower left) and E_T distribution of accepted towers (lower right).

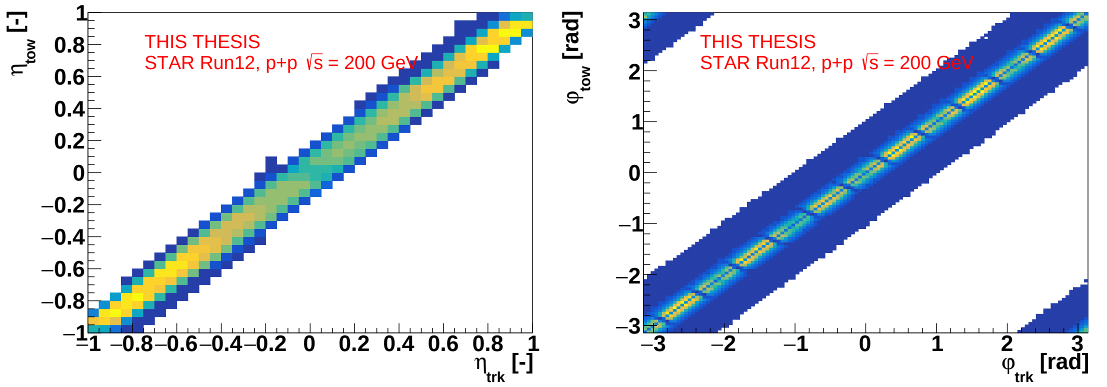

Figure 4.4: Distribution of $\eta_{trk} - \eta_{tow}$ matches (left) and $\varphi_{trk} - \varphi_{tow}$ matches (right).

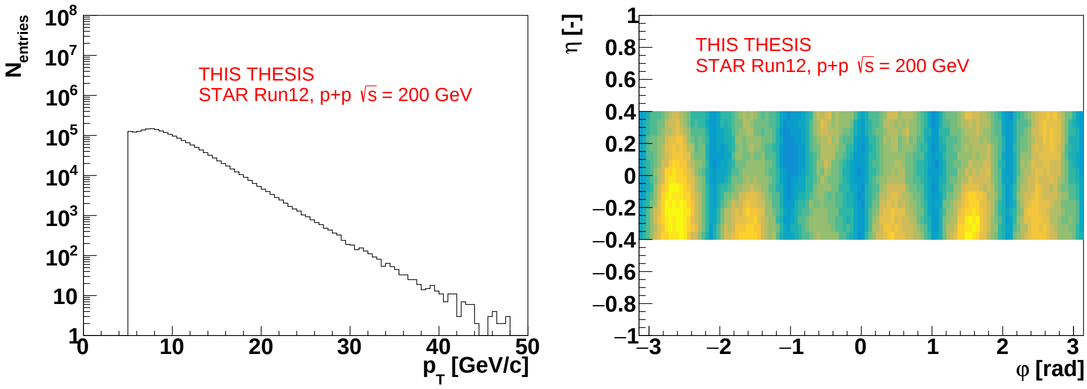

Figure 4.5: p_T distribution of accepted jets (left) and $\eta - \varphi$ distribution of accepted jets (right). Notice the effect of trigger bias on the jet p_T distribution.

- $\frac{1}{N_{jet}^{SD}} \frac{dN}{dp_{T,b}d\Delta R}$ - double-differential distribution of soft-dropped emission p_T with respect to distance between subjets.
- $\frac{1}{N_{jet}^{SD}} \frac{dN}{dz_g d\Delta R}$ - double-differential distribution of shared momentum fraction z_g with respect to distance between subjets.
- $\frac{1}{N_{jet}^{SD}} \frac{dN}{d\log(k_T)d\Delta R}$ - double-differential distribution of transverse emission strength $\log(k_T)$ with respect to distance between subjets.

In all instances the N_{jet}^{SD} is the number of jets that passed the SoftDrop condition.

The measurement is done for jet $p_T \in [20, 40]$ GeV/c, because at this p_T range the jet population is unbiased by the JP2 trigger. In order to achieve acceptable statistics we bin the distance between subjets ΔR by $[0.05, 0.15]$, $[0.15, 0.25]$ and $[0.25, 0.6]$. Since this measurement will require unfolding of the measured quantities, we also generate Monte-Carlo event samples using `PYTHIA6` model with STAR tune [32]. The generated events are then also passed through GEANT3 [33] simulation in order to bring effects of JP2 triggering as well as detector smearing. The measured data and its comparison with particle and detector level spectra can be seen in the figure 4.6 bellow. We can notice, that there is a subtle difference between measured, particle and detector level distributions. This could greatly help during the unfolding procedure.

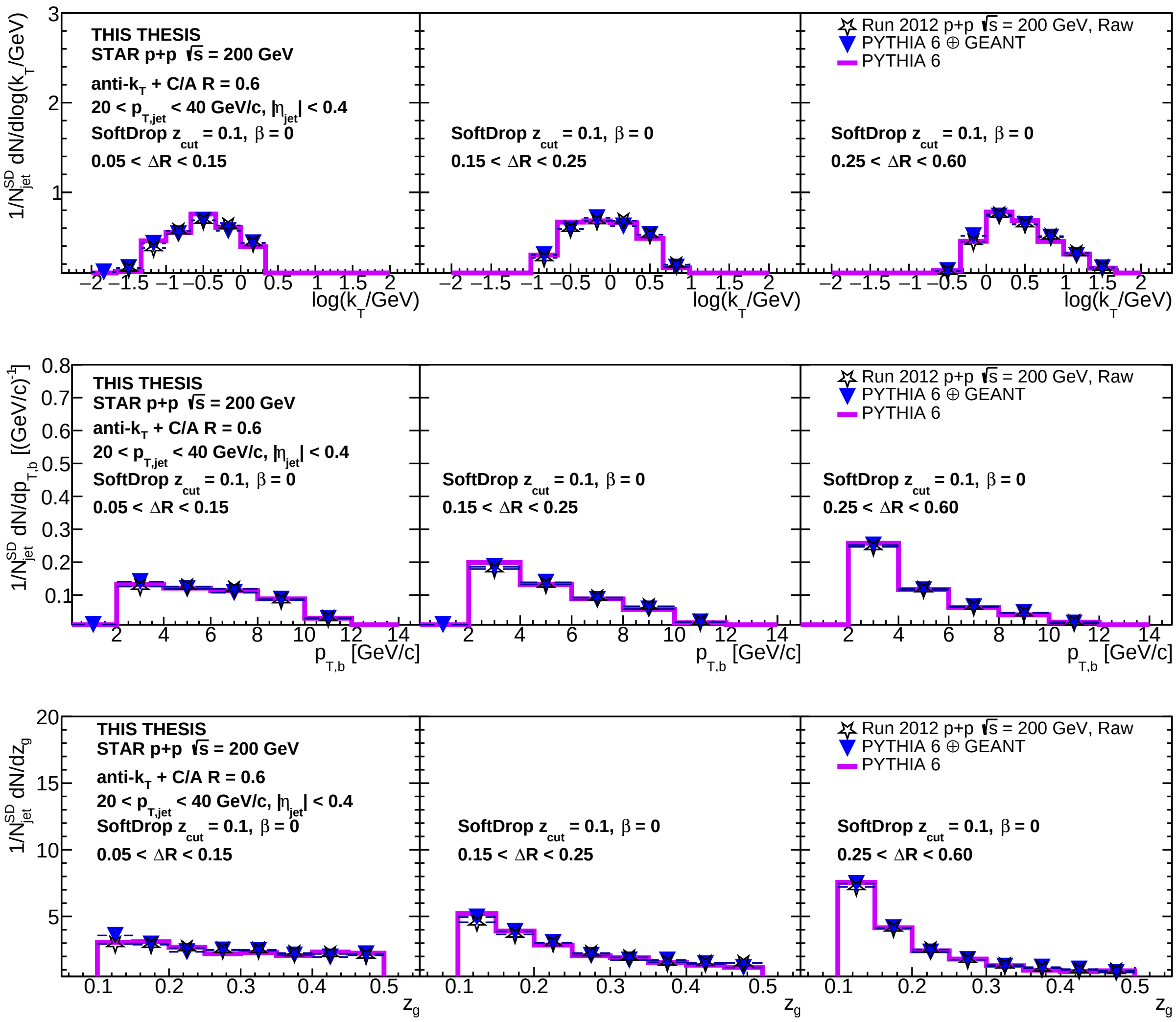

Figure 4.6: Raw double differential jet substructure observables measured in p+p collisions at $\sqrt{s} = 200$ GeV. (top row) Raw $\log(k_T)$ distributions with respect to the increasing angle between subjets ΔR. (middle row) Raw $p_{T,b}$ distributions with respect to the increasing angle between subjets ΔR. (bottom row) Raw z_g distributions with respect to the increasing angle between subjets ΔR.

Chapter 5

Unfolding of the Groomed Jet Substructure Observables

Every measurement is affected by the finite resolution of the instrumentation. Hence, the measurements done by such instrumentation are not representing what one might expect from the theory, since theoretical predictions are not taking such issues into the account. In the context of jet physics measurements, we have an effect of detector smearing on our observables. Example of this can be seen of Fig. 4.6. Here the green curve represents a prediction given by the Monte-Carlo event generator at particle level and red triangle is representing the same prediction but as it would be represented by the detector. The goal for our next steps is to unfold the raw detector level measurements represented by the open stars, into the particle level representation of such measurements.

In this analysis the unfolding was done using the iterative Bayesian method [34] implemented in the RooUnfold framework [35]. As a first step, we begin by constructing a response matrix between observables. Since in this analysis we use fixed jet p_T range of $[20, 40]$ GeV/c we apply only the 2D unfolding method. The procedure for construction is as follows:

1. For identical events we reconstruct the jets at the detector level and particle level separately.
2. Then the jets are matched based on the condition that $\Delta R(\text{particle}, \text{detector}) < 0.6$.
3. The jets without a match are then accounted for as missed jet (particle level) and fake jet (detector level).
4. After this step the response between $(O_{\text{det}}, \Delta R_{\text{det}})$ and $(O_{\text{part}}, \Delta R_{\text{part}})$ is constructed, where O is the observable in consideration - i.e. in our case z_g, $p_{T,b}$ and $\log(k_T)$.

The examples of weighted responses between pairs of measured variables can be seen on Fig. 5.1.

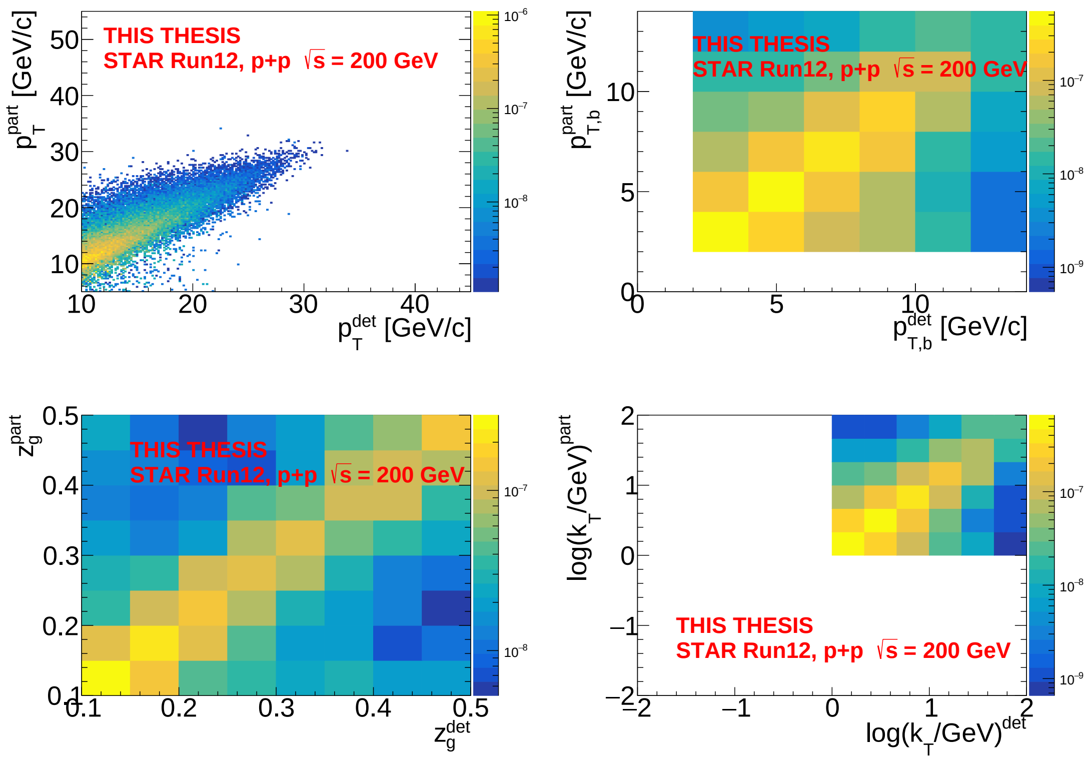

Figure 5.1: (upper left) Response matrix between measured and true jet p_T. (upper right) Response matrix between measured and true $p_{T,b}$. (lower left) Response matrix between measured and true z_g. (lower right) Response matrix between measured and true $\log(k_T)$. Substructure observables have coarser response matrices due to the lower statistic of observables.

Before unfolding we utilize observables generated with `PYTHIA6` events (the green line on Fig. 4.6) as prior distributions. Then the measured spectra are iteratively unfolded.

After the unfolding is done, one needs to ask a question about validity of such procedure, which is measured by closure testing. There are following types of closure tests:

1. **Refold closure** - after application of the unfolding step, the validity is assured through inversion of response and application of the inverted response to the unfolded data. After that the ratio of refolded data to measured data is taken. If such procedure is valid the resulting ratio should be as close to unity as possible with maximum deviation of 20%.

2. **Test and validation closure** - data used to construct response matrix for full unfolding is split into two halves, one is used to construct a response matrix and another is used to construct observables to unfold. Then the data from the first half is unfolded and the ratio of unfolded data to particle level data is taken. This procedure is called **validation closure** and it tells us how good our unfolding procedure is on the seen data. Second half, which is not used in the response matrix construction, is also unfolded and again the ratio is taken with respect to the unseen particle level data. This is called **test closure** and it tells us how valid our unfolding is on the unseen data. In both cases the ratios should again be as close to unity as possible with maximum deviation of 20%.

The resulting closure ratios can be seen on Figs. 5.2, 5.3 and 5.4. Since in almost all cases the resulting value are close to unity within 20%, the resulting closure procedure can be seen as valid in this setting. The cases of divergent first and last points can be omitted, since they arise from the bin edge effects. Fully unfolded spectra can be seen of Fig. 5.5.

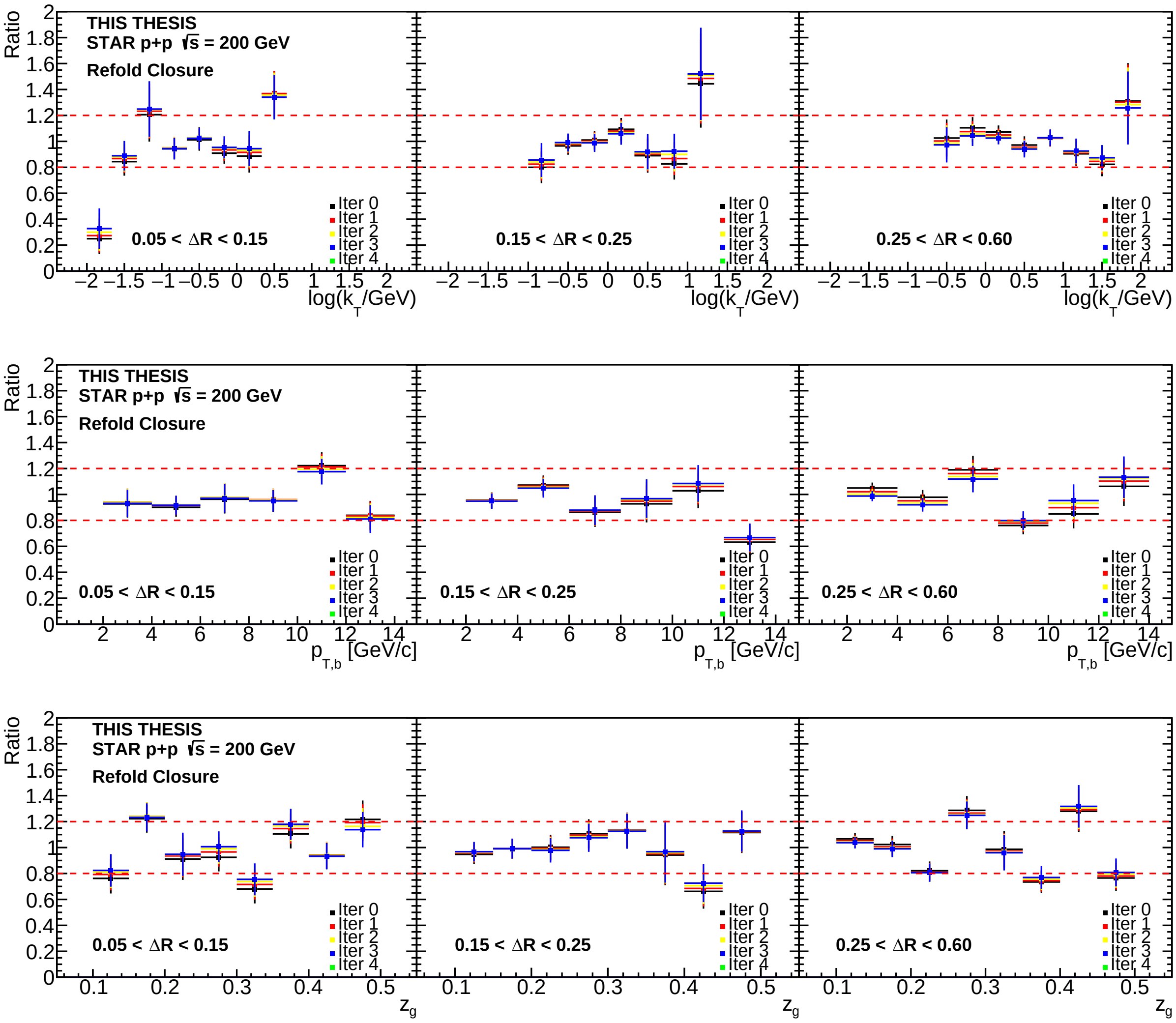

Figure 5.2: Refold closure for different ΔR distance between subjets (left-to-right) and different observables (top-down).

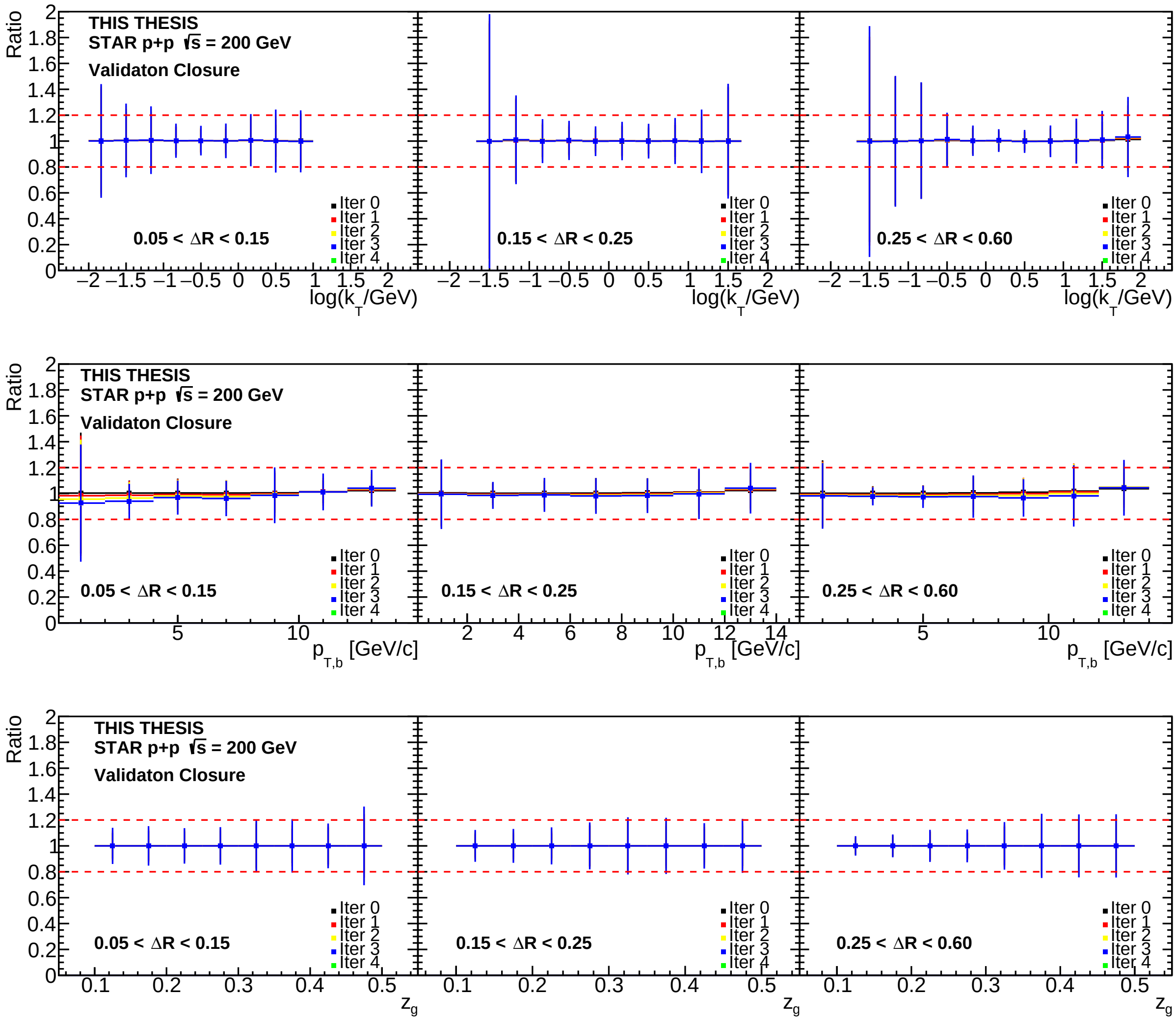

Figure 5.3: Validation closure for different ΔR distance between subjets (left-to-right) and different observables (top-down).

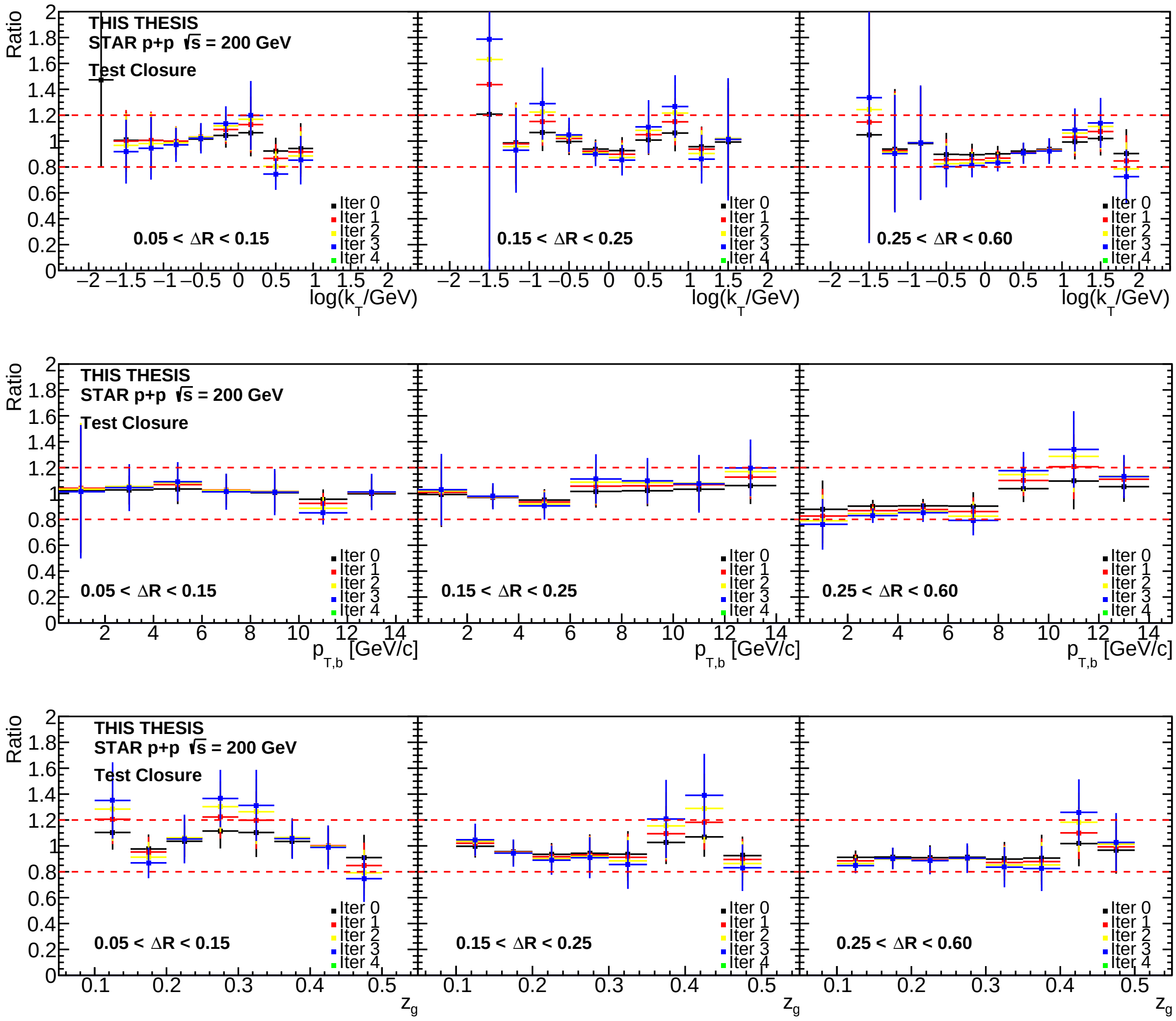

Figure 5.4: Test closure for different ΔR distance between subjets (left-to-right) and different observables (top-down).

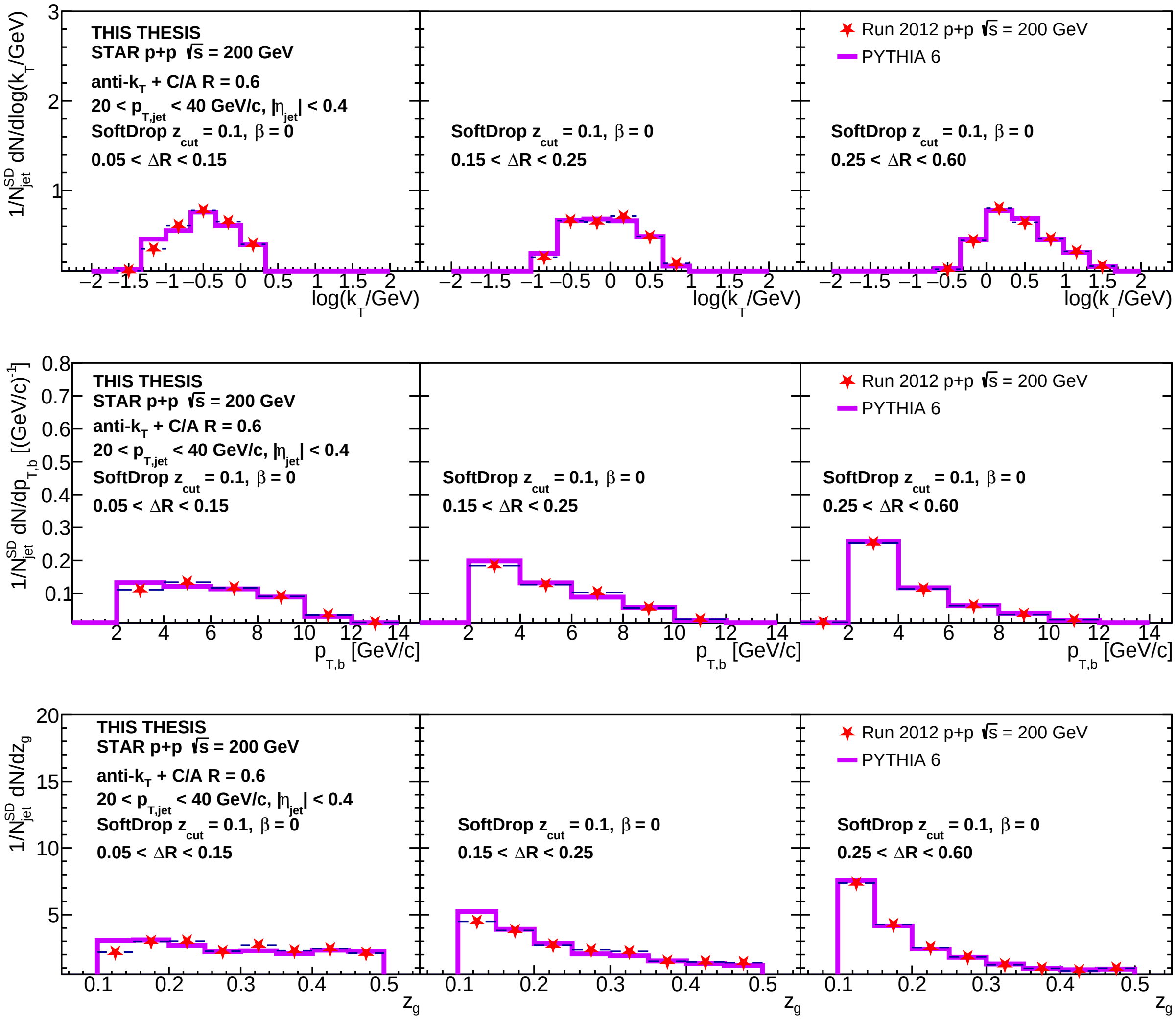

Figure 5.5: Fully unfolded double differential jet substructure observables measured at $\sqrt{s} = 200$ GeV. (top ronw) Fully unfolded $\log(k_T)$ distributions with respect to the increasing angle between subjets ΔR. (middle row) Fully unfolded $p_{T,b}$ distributions with respect to the increasing angle between subjets ΔR. (bottom row) Fully unfolded z_g distributions with respect to the increasing angle between subjets ΔR.

Conclusion

The goal of this project was to investigate double-differential jet substructure in p+p collisions at $\sqrt{s} = 200$ GeV measured by the STAR experiment.

In the first chapter the topic of jet emission and jet substructure was briefly introduced. The second chapter then describes the Relativistic Heavy Ion Collider (RHIC) and the STAR experiment, with emphasis on the detectors that are used for jet physics measurements. The third chapter summarizes the current measurements of jet substructure both at RHIC and LHC.

The fourth chapter describes the process of event selection, jet reconstruction and measurement of substructure observables. Most importantly a set of QA histograms is presented, to assure, that the analysis procedure is adhering to the required parameters. At the end of the fourth chapter the raw spectra of substructure observables are presented.

The last, fifth chapter, deals with the unfolding of the measured observables. The unfolding procedure is done in the $[20, 40]$ GeV/c p_T range due to the fact that the JP2 trigger is fully efficient at this range. Since all closures lie close to the unity within of the 20 % error, one can assume that the unfolding procedure is valid. Figure 5.5 presents the main result of this research project, the first fully unfolded double-differential jet substructure observables in p+p collisions at $\sqrt{s} = 200$ GeV at RHIC.

The results for the $\log(k_T)$ observable show, that with the increasing distance between measured subjets the intensity of the $\log(k_T)$ increases. The results for z_g and $p_{T,b}$ observable show that collinear splits have the hardest emissions, while splits with large distances between subjets are becoming softer (soft wide angle radiation). All those results lie in general within QCD expectations.

It is important to note, that current results should be regarded as preliminary only, since the unfolding procedure, while satisfying the needed closures, is not fully complete. This is due to the fact that we do not take into account shifts in the jet p_T due to the limited detector resolution. Hence a proper way will be to unfold the $(p_T, O, \Delta R)$, where O is an observable. Unfortunately, 3D unfolding is extremely complicated and goes beyond the scope of this undergraduate thesis project. One potential way to solve this is to use statistical correction based upon the jet energy resolution of the detector.

Bibliography

[1] S. Marzani, G. Soyez, and M. Spannowsky*Looking inside jets: an introduction to jet substructure and boosted-object phenomenology* Vol. 958 (Springer, 2019), 1901.10342.

[2] G. Bewick, S. Ferrario Ravasio, P. Richardson, and M. H. Seymour, JHEP **04**, 019 (2020), 1904.11866.

[3] F. A. Dreyer, G. P. Salam, and G. Soyez, JHEP **12**, 064 (2018), 1807.04758.

[4] S. Marzani, G. Soyez, and M. Spannowsky*Looking inside jets: an introduction to jet substructure and boosted-object phenomenology* Vol. 958 (Springer, 2019), 1901.10342.

[5] STAR Collaboration, Drupal, `https://drupal.star.bnl.gov/STAR/book/export/html/7278` [Online; 2.6.2019].

[6] STAR Collaboration, The STAR Forward Calorimeter System and Forward Tracking System, `https://drupal.star.bnl.gov/STAR/files/ForwardUpgrade.v20.pdf` [Online; 8.8.2020].

[7] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP **05**, 026 (2006), hep-ph/0603175.

[8] T. Sjöstrand *et al.*, Comput. Phys. Commun. **191**, 159 (2015), 1410.3012.

[9] CMS, A. M. Sirunyan *et al.*, Phys. Rev. Lett. **120**, 142302 (2018), 1708.09429.

[10] ALICE, S. Acharya *et al.*, Phys. Lett. B **802**, 135227 (2020), 1905.02512.

[11] STAR, J. Adam *et al.*, (2020), 2003.02114.

[12] ATLAS, G. Aad *et al.*, Phys. Rev. Lett. **124**, 222002 (2020), 2004.03540.

[13] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972), [Yad. Fiz.15,781(1972)].

[14] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).

[15] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977), [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[16] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, JHEP **08**, 001 (1997), hep-ph/9707323.

[17] M. Wobisch, *Measurement and QCD analysis of jet cross-sections in deep inelastic positron proton collisions at $\sqrt{s} = 300$ GeV*, PhD thesis, Aachen, Tech. Hochsch., 2000.

[18] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, JHEP **05**, 146 (2014), 1402.2657.

[19] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), 0802.1189.

[20] BRAHMS Collaboration, I. Arsene *et al.*, Nucl. Phys. **A757**, 1 (2005), nucl-ex/0410020.

[21] B. B. Back *et al.*, Nucl. Phys. **A757**, 28 (2005), nucl-ex/0410022.

[22] PHENIX Collaboration, K. Adcox *et al.*, Nucl. Phys. **A757**, 184 (2005), nucl-ex/0410003.

[23] STAR Collaboration, J. Adams *et al.*, Nucl. Phys. **A757**, 102 (2005), nucl-ex/0501009.

[24] sPHENIX, S. Campbell, J. Phys. Conf. Ser. **832**, 012012 (2017), 1611.03003.

[25] A. Accardi *et al.*, Eur. Phys. J. A **52**, 268 (2016), 1212.1701.

[26] M. Anderson *et al.*, Nucl. Instrum. Meth. **A499**, 659 (2003), nucl-ex/0301015.

[27] STAR Collaboration, A Proposal for STAR Inner TPC Sector Upgrade (iTPC), `https://drupal.star.bnl.gov/STAR/files/STAR_iTPC_proposal_06_09_2015.pdf` [Online; 8.8.2020].

[28] STAR Collaboration, M. Beddo *et al.*, Nucl. Instrum. Meth. **A499**, 725 (2003).

[29] P. Skands, S. Carrazza, and J. Rojo, Eur. Phys. J. C **74**, 3024 (2014), 1404.5630.

[30] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. C **72**, 1896 (2012), 1111.6097.

[31] M. Cacciari and G. P. Salam, Phys. Lett. B **641**, 57 (2006), hep-ph/0512210.

[32] STAR, J. Adam *et al.*, Phys. Rev. D **100**, 052005 (2019), 1906.02740.

[33] R. Brun, F. Bruyant, M. Maire, A. McPherson, and P. Zanarini, (1987).

[34] G. D'Agostini, Improved iterative Bayesian unfolding, in *Alliance Workshop on Unfolding and Data Correction Hamburg, Germany, May 27-28, 2010*, 2010, 1010.0632.

[35] T. Adye, Unfolding algorithms and tests using RooUnfold, in *PHYSTAT 2011*, pp. 313–318, Geneva, 2011, CERN, 1105.1160.